

# Self-Experience and the Feeling of Being Oneself

Inauguraldissertation  
zur Erlangung der Doktorwürde  
der Philosophischen Fakultät der Universität Heidelberg,

vorgelegt von  
Mag. Mag. Philipp Schmidt

Erstgutachter: Prof. Dr. Dr. Thomas Fuchs  
Zweitgutachter: PD Dr. Ingrid Vendrell-Ferran

Datum: 22.3.2021



*Dem lebendigen Geist*



# Abstract

This dissertation aims to clarify the role of self-experience for the notion of selfhood. Against three forms of skepticism about self-experience widespread in the history of philosophy, it argues that self-experience is informative with regard to the being of selfhood and that self-experience is best understood as the very way of being of selfhood. On this view, self-experience constitutes selfhood. To support this claim, this dissertation examines the epistemology, metaphysics, and phenomenology of self-experience, notably, in its most basic forms.

In Part I, I focus on the epistemology of self-experience. I identify and discuss the three forms of skepticism about self-experience that are usually propagated and developed as a response to a Cartesian approach to selfhood. (a) The first skepticism consists in the idea that self-experience is throughout deceptive and wrongly gives rise to the belief that selfhood is real. (b) The second skepticism consists in the idea that our experience does not comprise any self-experience in a relevant sense and that therefore self-experience neither motivates the belief that selfhood is real nor provides any insights into what selfhood is. (c) The third skepticism consists in the idea that self-experience does warrant the reality of selfhood but is uninformative regarding the metaphysical nature of selfhood. I reject all three of these forms of skepticisms. This refusal motivates the thesis that self-experience may not only reveal the reality of selfhood but also the metaphysical nature of selfhood.

In Part II, I identify Zahavi's minimal self-view with such a metaphysical thesis about the relationship between self-experience and selfhood. Clarifying the ontological commitments of the minimal self-view, I lay out what I suggest to call *self-experientialism*. Self-experientialism takes self-experience to be both the *modus cognoscendi* and *modus essendi* of selfhood. On this view, the reality of selfhood, its being, consists in having self-experience and, thus, is of experiential nature. After motivating self-experientialism by discussing some accounts of pre-reflective self-awareness similar to—but, as my analysis shows, to be distinguished from—Zahavi's minimal self, I contrast self-experientialism with other recent experiential accounts of selfhood. Against these alternatives, I argue that Zahavi's phenomenologically motivated self-experientialism is to be preferred as an experiential approach to selfhood.

In Part III, I focus on the phenomenology of self-experience. Given that self-experience ought to present the being of selfhood, I argue we need to explain how a sense of being is involved in even the most basic form of self-experience. This, I will argue, will require an integration of feeling into the picture. Departing from a discussion of recent criticisms of the minimal self-view that doubt the minimal self could account for an experience of individuation, I develop the view that the minimal self includes what I propose to call a *feeling of being oneself*.



# Table of Contents

Acknowledgements	11
Introduction	15
Part I	
Ch. 1 The Hallucination Model of Self and the Deception of Self-Experience	27
1.1 Focus on self-experience and the no-self doctrine	28
Virtual self theory (VST) and the hallucination model of self	28
Hallucination vs. illusion model of self	30
Self-representation and self-experience	30
The hallucination model of self vs. illusionism	32
Skepticism about self-experience and the hallucination model	33
1.2 The neuroscientific study of self-experience	34
1.3 Metzinger's <i>self-model theory of subjectivity</i> and hallucination	36
1.4 <i>Being no one</i> : Key assumptions and arguments	39
The "phenomenological fallacy"	39
The impossibility of direct (self-)reference	41
The transparency of representation and the sense of self	42
Self-experience, indirect reference, and the lack of a "self"-referent	44
All phenomenal content is hallucination and the virtual self is unreal	45
1.5 Argumentum ad absurdum. Fending off the hallucination model	49
Global phenomenal hallucination?	50
Partial phenomenal hallucination? The incoherence of self-hallucination	55
1.6 Conclusion: The hallucination model is inapt	60
Ch. 2 What Self-Experience? No Self and the Grinning of the Sceptics	63
2.1 The Humean challenge: Is self-experience possible?	65
2.2 Strawson and the limits of Humean skepticism	67
2.3 The No-Self-Picture is phenomenologically inaccurate	75
McGilvary on James	77
Applying McGilvary's critique to the No-Self-Picture	81
2.4 Zahavi with Husserl: The stream of consciousness shows up	90
Zahavi and the minimal self against the No-Self-Picture	91
Husserl's "inner time-consciousness" as the basis of the minimal self	95
2.5 Illusionism and recent doubts about the stream of consciousness	101
Benovsky: The stream of consciousness as an illusion of the brain	102
Blackmore: The stream of consciousness as a delusion of consciousness	102
Garfield: There is no phenomenal consciousness	103
Response to the skeptics	104
2.6 Conclusion: The stream of consciousness against the No-Self-Picture	108

Ch. 3	Elusiveness, Illusion, and the Limits of Self-Experience	111
3.1	Elusiveness I: Kant and the indeterminability of the pure ‘I think’	113
	Kant on self-knowledge	113
	Kant’s motivation	121
	Conceptual ramifications of Kant’s elusiveness claim and soul-theory	123
	On what’s wrong with soul-theory and the metaphysics of elusiveness	128
3.2	Illusiveness: McClelland on the self as the bearer of self-experience	137
	Who is the bearer? Identifying the self with the brain	139
	The brain doesn’t fare better than the pure ‘I think’	142
3.3	Elusiveness II: Two takes on why the self must be elusive	145
	Howell: Non-objective self-awareness means elusiveness of the self as the bearer	145
	Evans: Identifying the elusive self with unprojected consciousness	147
3.4	Conclusion: The elusiveness view still gets the picture wrong	150

## Part II

Ch. 4	Minimal Self and the Ontological Interpretation of Self-Experience	155
4.1	The three conceptual roles of the minimal self	157
4.2	On the 1 <sup>st</sup> conceptual role: Phenomenal consciousness	160
4.3	On the 2 <sup>nd</sup> conceptual role: Self-consciousness and sense of self	162
	Guillot: Zahavi’s account is guilty of a confusion concerning ‘subjective character’	162
	Terminological clarifications: Zahavi’s response to Guillot	164
	A phenomenological model: Three levels of experience and egocentricity	167
	Phenomenological arguments: That pre-reflective self-awareness is egological	170
	Challenge I: McClelland and the elusiveness of mineness	178
	Challenge II: Peacocke’s creature at Degree 0	180
	Challenge III: Dainton’s notion of co-consciousness and the Isolation Thesis	182
4.4	On the 3 <sup>rd</sup> conceptual role: The notion of selfhood	185
	From the phenomenology of reflection to a sense of self	187
	From the phenomenology of a sense of self to a notion of selfhood and back?	190
	Minimal self and phenomenological self-experientialism	193
4.5	Conclusion: Self is sense of self	194
Ch. 5	Gaps, Identity, and the Primacy of the Minimal Self	197
5.1	Self-experientialism is not Cartesianism—or is it?	198
5.2	A modest reading of Locke: Distinguishing self from person	202
5.3	Nida-Rümelin: Pre-reflective self-awareness as ‘nature-revealing’	207
5.4	Dainton: Substrate phenomenalism	212
5.5	Strawson: The pearl view	218
	The constancy of objects implies experiential continuity	223
	The gaps are in the content, not the stream	225
	Frictions in Strawson’s phenomenological proposal and metaphysical ramifications	227



5.6	Bayne: Virtual phenomenality	229
	Virtual phenomenality is an experiential approach to selfhood	229
	Virtual phenomenality can address the issue of personal identity	231
	Self-experientialism is not the same	232
	The advantages of virtual phenomenality presuppose self-experientialism	234
5.7	Self-experientialism, identity, and ontology: The primacy of self	237
5.8	Conclusion: Self before person	241

### Part III

Ch. 6	Minimal Self and the Feeling of Being Oneself	245
6.1	Is Zahavi's thin self too thin?	249
6.2	Mick and Mack: What is missing in the minimal self?	252
6.3	Caring about oneself—a Lotzean idea on <i>Fürsichsein</i>	254
	From Lotze to a new perspective on Zahavi's notion of the minimal self and its critics	260
	Lotze's <i>Fürsichsein</i> , self-concern, and Zahavi's minimal self	265
6.4	Integrating feeling into the minimal self account	267
	Husserl's <i>lebendige Gegenwart</i> : The 'I' as feeling in affection and action	268
	Ratcliffe's existential feelings: Protention, possibility, and being	273
	Existential feelings and the stream of consciousness	277
	On the phenomenology of 'neutral' experiences	282
6.5	Minimal self and the feeling of being oneself	292
6.6	Conclusion: The feeling of being oneself	297
Ch. 7	Specifying the Feeling of Being Oneself and Depersonalization	299
7.1	Specifying the feeling of being oneself	301
	Feeling and consciousness	301
	<i>Seinsgefühl</i> : Existence, life, and the body	303
	The value of existence	307
	The intertwining of feeling, temporality, and self	309
	Openness and involvement: Egological participation through feeling	311
	Variability of the feeling of being oneself: Scope and graduality	317
7.2	Minimal self, the feeling of being oneself, and depersonalization	320
	Depersonalization as a loss of self: An anti-thesis to the minimal self-view?	321
	Billon and Guillot: What is left of the sense of self in depersonalization?	323
	Oesterreich's alternative approach: Inhibition of the feeling of life ( <i>Lebensgefühl</i> )	330
	Self-experience and the feeling of being oneself in depersonalization	333
	Zahavi's account of depersonalization and the feeling of being oneself	338
7.3	Conclusion: Depersonalization as a modification of self-feeling	341
	Conclusion: Self-Realization and How I Feel about Myself	343
	References	355



# Acknowledgements

There is a considerable number of people who have been involved in the process of writing this dissertation. They all helped me—in various ways—to find what I came to perceive and now like as my own path through the jungle of philosophical thought. I started the project in 2016 when I was working as a Research Assistant (“Prae-Doc-Assistent”) to Matthew Ratcliffe at the University of Vienna supervised by him and Thomas Fuchs at Heidelberg University, where I now submit the thesis. I am very grateful to Thomas Fuchs for taking the lead in the supervision of my thesis and for the numerous discussions in the past years in which he was nothing but encouraging and supportive. I am most grateful for his efforts in trying to provide help as much as he can. I learned a lot from our meetings and enjoyed all the fine events as well as all my stays and times in Heidelberg with the “Section”, which offers a unique atmosphere for philosophical exchange. I am also very grateful to Matthew Ratcliffe for his continuing support as my boss and supervisor in Vienna and later as external supervisor, after he had moved back to the UK in 2018. Despite all the commitments of his new job, he made time free to read, comment and accompany my work, which I greatly appreciate. I also want to thank him for his great hospitality during my two short stays in York. I enjoyed them genuinely and profited highly from our exchange. I am also grateful for all the philosophical as well as practical knowledge he shared with me and the other members of his Viennese research group, a crew that I spent so many unforgettable moments with.

Moreover, I wish to express my gratitude to Dan Zahavi for his generosity in reading and providing feedback on two chapters. I also greatly appreciate his hospitality and friendliness during my brief research stays in Copenhagen (2016, 2017) and Oxford (2019), where he was available for consultation on several topics of my thesis. I am very much indebted to Christian Tewes for his helpful suggestions to improve my dissertation proposal. Special thanks to Sophie Loidolt who has constantly been my mentor and friend already for more than a decade, and, who has always supported me, especially in moments of doubt.

Although it should go without saying, I wish to emphasize that, despite being endowed with the help of such exquisite and prominent philosophers, I am to be held fully responsible for all remaining lack of clarity, errors, misunderstandings, and idiosyncracies of my dissertation thesis as well as the positions I defend in it.

I am also grateful to Sophie Loidolt for her guidance during my time as her Research Assistant (“wiss. Mitarbeiter”) from 2018 to 2021 at TU Darmstadt, a time in which she invigorated me in developing and trusting my own ideas—and, in which she allowed me to grow with her. I am also grateful to Michela Summa at the University of Würzburg, to whom I have also been Research Assistant (“wiss. Mitarbeiter”) since 2019. Both were

highly flexible in the past year allowing me to get a lot of work done on my thesis (more than half of the work was written in 2020), despite all the interesting projects and goals we have been pursuing together. I am much obliged to the freedom they have been granting me under their leadership.

I am much indebted also to a number of members of research groups and reading circles (“Sperl-Lesekreis”) that I belonged to over longer periods of time—in Vienna, Heidelberg, Darmstadt, and Würzburg. I feel incredibly honoured and thankful for the opportunity to learn from such wonderful communities of brilliant philosophical thinkers. The following have especially shaped my thinking in one way or another over the course of time: Oren Bader, Valeria Bizzari, Maike Cram, Diego D’Angelo, Christoph Durt, Wolfgang Fasching, Peter Gaitsch, Lukas Iwer, Martin Klein, Sophie Loidolt, Oliver Lukitsch, Karl Mertens, Martin Müller, Line Ryberg Ingerslev, Graham Parkes, Johanna Schlossberger, Matthias Schlossberger, Robert Stefanek, Christian Sternad, Michela Summa, Christian Tewes, Dylan Trigg, Daniel Vespermann, and others. Apart from these research groups, I have benefitted from philosophical exchange with and/or received support in some way by Maria Chiara Bruttomesso, Andrea Cimino, Erik Norman Dzwiza-Ohlsen, Reinhold Esterbauer, Michelle Montague, Jörn Müller, Luisa Paz Rodríguez Suárez, Lanei Rodemeyer, Adrian Spremberg, Georg Stenger, Galen Strawson, Zeno van Duppen, Juan Velázquez González, Anja Weiberg, and many others.

In early 2019, I spent a couple of months at Central European University Budapest in order to write on my thesis and to visit some courses in the philosophy of mind. I thank Tim Crane for enabling this in a very unbureaucratic manner and for his very friendly welcoming. I also thank Damian Aleksiev and Jamie Elliott for our philosophical exchanges during that time. Moreover, in early 2020, I spent a couple of days at the Husserl Archives at KU Leuven in order to study unpublished manuscripts of Edmund Husserl. Thanks to Andrea Cimino, Emanuele Caminada, Julia Jansen, and Thomas Vongehr for making this possible.

I presented materials of this dissertation thesis at various occasions and benefitted a lot of the insightful comments of the participants: “Workshop der DGAP Neues aus der Phänomenologie” (Apr. 2016, Heidelberg), “Cologne-Leuven Summer School in Genetic Phenomenology” (Jul. 2016, Cologne), “Workshop Cognitive Phenomenology with Michelle Montague & Galen Strawson” (May 2017, Vienna), “Workshop Phänomenologie und Neuer Realismus” (Jul. 2017, Vienna), “Workshop Perspectives on the Minimal Self” (May 2018, Vienna), “Workshop Personhood and Selfhood” (Jul. 2018, Manchester), “Workshop Multiple Dimensions of Selfhood” (Sep. 2018, Berlin), “International Conference Phenomenology and Personal Identity” (Nov. 2018, Prague), “International Conference First-Person Science of Consciousness” (May 2019, Witten/Herdecke),

“Workshop Double Intentionality. Historical and Contemporary Perspectives” (Sep. 2019, Würzburg).

I wish to thank my family for their constant moral support in pursuing my philosophical projects and their belief in me. I especially thank my parents for recalibrating my compass each time when I seemed to have lost sight of what matters to me. Thanks go to my brother who is always there for me and helps me developing other perspectives when needed. Finally, I wish to thank my partner, Jane Boddy, for her love and acceptance of who I am. Without her, all the work wouldn't have been possible. Having her on my side, life—work or not work—all seems like a beautiful adventure and journey.



# Introduction

The notion of self—together with its relatives such as the ‘I’, ‘ego’, ‘me’, ‘person’, and ‘soul’—has always been one of the most controversial philosophical issues and continues to be so today. Despite the long tradition of discussing the notion of self in the history of philosophy and adjacent disciplines, there still exists great heterogeneity in philosophers’ and researcher’s assessment of central aspects of what constitutes selfhood and how we are to understand its ontological status, let alone, whether we should talk about selfhood in the first place. Crucially, the heterogeneity not only concerns the way in which thinkers address the question of self—but also how thinkers pose the question of self to begin with.

Richard Swinburne, for instance, asks in his recent book: *Are we bodies or souls?* (2019) and addresses the “nature of human beings”<sup>1</sup>. Christopher Peacocke begins his book *The Mirror of the World. Subjects, Consciousness, and Self-Consciousness* (2014) with the questions: “What is it to represent something as yourself? What is the nature of the thing that you refer to when you think or perceive ‘I am thus-and-so’, if indeed there is any such thing?”<sup>2</sup> In *The Phenomenal Self* (2008), Barry Dainton starts out with the question, “What kind of thing am I?”<sup>3</sup> Jonardon Ganeri’s *The Self. Naturalism, Consciousness, and the First-Person Stance* (2012) aims to answer the questions, “What is it to occupy a first-person stance? Is the first-personal idea one has of oneself in conflict with the idea of oneself as a physical thing being in the world?”<sup>4</sup> Dan Zahavi, highlighting the interdisciplinary character of the issue of self, lists the following questions at the outset of his *Subjectivity and Selfhood. Investigating the First-Person Perspective* (2005): “What is a self? Does it really exist in reality or is it a mere social construct—or perhaps a neurologically induced illusion? If something like a self exists, what role does it play in our conscious life, and how and when does it emerge in the development of the infant? What might such psychopathological and neuropsychiatric disorders as schizophrenia or autism reveal about the fragility of self-identity and self-coherence?”<sup>5</sup>

These few examples show that, although there is a shared goal in that all want to determine what selfhood is, the interest in the issue of self can evoke quite different ways of approaching the target phenomenon, which most certainly shed light on quite multifarious aspects of it. Surely, any phenomenon selected for scrutiny can be regarded from distinct angles and perspectives, give rise to different questions, and require even more or less compatible methods of investigation. Some phenomena, however, not only

---

<sup>1</sup> Swinburne 2019, 1.

<sup>2</sup> Peacocke 2014, 1.

<sup>3</sup> Dainton 2008, xii.

<sup>4</sup> Ganeri 2012, 4.

<sup>5</sup> Zahavi 2005, 1.

raise distinct questions and demand diverse approaches, they also seem to be highly shaped by the way they are addressed. The self or selfhood is one of the more prominent examples of such phenomena: Depending on how one asks about it, the guiding *notion* of self or selfhood will make a huge impact on what one will find over the course of one's investigation. Swinburne clearly moves within the boundaries of traditional Cartesian dualism when he puts the two options on the table for understanding selfhood: body or soul. The idea that selfhood must consist in some form of substance or thing also seems to be presupposed in the way Dainton and Peacocke set their agenda. For Peacocke it also seems to be already decided that the self—if it exists—is something that you *represent* as yourself, which suggests that representation is the right way to characterize and describe the manner in which we consciously relate to ourselves. Zahavi focusses on the question how the self features in our conscious life, its development and vulnerability.

Do these thinkers refer to the same notion of self of which they are eager to describe the nature? Most problematically, it is not always clear when they target a different aspect of self or address a wholly different notion of self. Accordingly, what they carve out as the nature of self often is a correlate of their notion of self: Whatever they take to be a self shapes how they describe its nature and however they describe the nature of self will likely shape their notion of self.

It is no wonder, thus, that there exist so many different—and mostly divergent—concepts of self,<sup>6</sup> which has led some to insist we should stop applying the notion of self altogether:

Once we have accounted for people, their mental features, their relation to those human animals we call their bodies, and so on, we think we need to say something about 'the self' as well. There is no good reason to think so.

...

[T]here can be no reason, other than tradition, to continue to speak of the self.<sup>7</sup>

Since Eric Olson's verdict, debate on selfhood and its nature has intensified and extended rather than anything else. After the prominent focus on linguistic and computational models of consciousness and selfhood in the second half of the 20<sup>th</sup> century, especially in the analytic philosophy of mind, in the last two decades, a new interest and focus on the experience of selfhood and the relationship between phenomenal consciousness and self-consciousness has emerged. It is in this context that the question of the reality of selfhood and what selfhood consists in now has been increasingly investigated. One—if not the central—question concerns the relationship between *self-experience* and *self*, i.e., the

---

<sup>6</sup> Cf. Olson 1998, 648ff. or Strawson 1999, 484.

<sup>7</sup> Olson 1998, 656. Crucially, for Olson, this conclusion is based on the idea that there is no problem of the self: "But no concept could be so problematic that no one could agree about *anything* to do with it. For lack of a subject matter, then, there is no problem of the self." See Olson 1998, 645.



relationship between the phenomenological question of how we experience ourselves and the metaphysical question of what the nature of selfhood consists in. It should not come as a surprise that researchers and thinkers diverge heavily when it comes to assessing this question, too.

A wide-spread assumption, especially in the philosophy of mind, is that both questions are to be kept apart. For instance, Quassim Cassam highlights that “questions about the nature of self-consciousness are epistemological questions, whereas questions about the nature of the self are metaphysical”<sup>8</sup>. For this reason, he infers that “at the very least, further argument would be required to show that introspective self-awareness does not embody an illusion about the actual nature of the subject of one’s representations”<sup>9</sup>. There lingers a great suspicion regarding self-experience in the debate.

However, others have pursued a quite different path and consider the phenomenology of self-experience precisely the starting point for any inquiry into the metaphysics of self. Galen Strawson has been a prominent defender of such a line of thinking. He writes:

Well, it is self-experience that gives rise to the problem [of the self, P.S.]—the vivid sense, delusory or not, that there is such a thing as the self. I think, in fact, that it is the whole source of the problem, in such a way that when we ask whether selves exist, what we are actually asking is: Does anything like the sort of thing that is figured in self-experience exist?<sup>10</sup>

Since, for Strawson, it is self-experience that evokes the question about self in the first place, we should first study our self-experience before we could possibly form a judgment about the ontological status of self. He emphasizes:

Before we ask the metaphysical question

(I) Do selves exist?

We must ask and answer the phenomenological question

(II) What sort of thing is figured in self-experience?<sup>11</sup>

I think Strawson’s appeal to self-experience is right and I agree that the phenomenological question is primordial to the metaphysical question in the manner he has stated it. But before one can proceed to answering questions about the ontological status of selfhood

---

<sup>8</sup> Cassam 1997, 3.

<sup>9</sup> Cassam 1997, 7.

<sup>10</sup> Strawson 2017, 44.

<sup>11</sup> Strawson 2017, 44.

based on a phenomenological description of self-experience, another very important issue needs to be addressed: What is it that we can learn about selfhood when we look into self-experience? In which sense does or does not self-experience tell us something about the nature of selfhood? And, in which way can we characterize self-experience in the phenomenologically most adequate manner?

These, I suggest, are the questions that should be addressed and should present our point of departure when we are interested in the notion of self and the nature of ourselves. Although not always in a straightforward manner, these questions have in fact been guiding thinkers in developing their conceptions of selfhood. This becomes salient when we consider the Cartesian proposal of selfhood and his famous dictum *cogito ergo sum*, which has often been taken as the position to beat. I think, therefore, I am. Few sentences in the history of philosophy have gained a similar popularity like these words, which, being at the core of the Cartesian philosophical approach, have given rise to new paradigms in epistemology and metaphysics. Mostly by drawing attention to the role of self-consciousness in grounding a body of knowledge, Descartes paved the way for the philosophy and the sciences of the modern age. In doing so, he also laid the foundation of a new set of philosophical problems. One of these problems concerns precisely the center or principle of knowledge: the self. What can be said about the self according to Descartes? The cogito-argument, as is well-known, leads to three interrelated claims:

- (1) that we have self-experience
- (2) that we can infer from self-experience that we exist
- (3) that we can infer from self-experience what our metaphysical nature is: thinking being.

To make a case for these claims, Descartes, most famously, points to the impossibility of doubt about the performing subject of doubt. While I can cast doubt on everything that is given to me to the effect that it could turn out to having been an appearance only, the same does not apply to the thinking itself involved in the doubt. Trying to cast doubt on everything consciously given to the subject, the subject will find only one aspect of consciousness impossible to put into doubt: the activity of doubt itself, i.e., the activity of the thinking subject itself. According to Descartes, the experienced impossibility of doubting one's own thinking activity not only amounts to the insight *that* one exists, but ultimately also reveals that one exists essentially as a thinking being. Put differently, the general idea is that self-experience reveals the metaphysical nature of selfhood and warrants the reality of self. This is one central line of thinking about the relationship between self-experience and self. There are three others that are important, systematically but also historically. All of them can be considered as responses—intended or not, explicit or not—to the Cartesian proposal and consist in a different form of skepticism about self-experience directed at one or more than one of the three Cartesian claims.

The first form of skepticism can be seen as being most averted from the Cartesian approach in that it takes our experience, including our self-experience, to be throughout hallucinatory and subject to doubt. On this first form of skepticism, we may be said to have self-experience but we cannot conclude anything from it with regard to our selfhood: neither what the nature of our selfhood is nor that we exist. In fact, on account of this first skepticism, we are even better off denying that selfhood is real, and that self-experience wrongly makes us believe in the self, when really self-experience refers to nothing at all. Hence, the first form of skepticism acknowledges Descartes' claim (1) but denies (2) and (3).

The second form of skepticism denies (1) and maintains that we don't have self-experience to begin with—undermining (2) and (3) by implication. This kind of skepticism, thus, consists in doubt about the phenomenology of self-experience. Although not necessarily, it can also amount to or is associated with the no-self doctrine, the notion that selfhood is not real.

The third form of skepticism is the mildest of all forms of skepticism. It acknowledges that we have self-experience and also that we are entitled to infer from that that we exist. It is only skeptical with regard to what else we can learn from self-experience about ourselves, and so denies that self-experience is revelatory with regard to the nature of selfhood. Accordingly, this form of skepticism accepts (1) and (2), but denies (3).

In this dissertation, my aim is to present and discuss these three forms of skepticism and to show that none of them present an attractive line of thinking (Part I). That is, my aim is to defend the view that we have self-experience that is informative with regard to our existence (i.e., that selfhood is real) and being (i.e., the nature of selfhood). Does this imply that we should endorse Cartesianism? No, in fact, my second aim is to present an alternative to both skepticism about self-experience *and* Cartesianism. To do so, I describe and discuss the approach of what I label *phenomenological self-experientialism*, an ontological interpretation of self-experience that can be found in thinkers of the phenomenological tradition and which I defend against alternative recent experiential approaches to selfhood. The basic idea of self-experientialism is that self-experience is both the *modus cognoscendi* and *modus essendi* of selfhood. Self-experience, on this view, is informative about selfhood because self-experience forms the very being of selfhood (Part II). Finally, I address how exactly we should understand the phenomenology of self-experience in its most basic form so as to present the plausible basis for the being of selfhood. In this context, I argue that self-experience, if it is to present the mode of being of selfhood, should also include a sense of being. I suggest that this is indeed the case and that self-experience involves what I call a *feeling of being oneself*, thereby attributing a more central role to feeling for self-consciousness than is usually acknowledged (Part III).

Investigating self-experience thus from an epistemological, metaphysical, and phenomenological perspective, this dissertation, I suggest, offers a novel approach in at least three regards. First, carving out the challenges that adhere to each of the three skeptical views I propose a new conceptual and methodological framework for the assessment of approaches to selfhood. For, although the three forms of skepticism about self-experience are positions that have repeatedly been defended in the different debates about selfhood, it has not yet been worked out that variants of the three forms are recurring attempts to avoid any philosophical position that comes too close to Cartesianism. Second, by shedding light on metaphysical aspects of self-experience, I also connect phenomenological approaches with ontological discussions in the philosophy of mind, laying bare the ontological implications of phenomenological theorems. Moreover, as it will turn out, it also seems impossible to make phenomenological claims about self-experience without by the same token subscribing to a certain metaphysical understanding of what selfhood is. Third, in arguing in favor of the notion of a feeling of being oneself, I offer a new understanding of how feeling and self-experience might be seen as intrinsically entwined.

The investigation proceeds as follows. Chapter 1 begins with an examination of the first form of skepticism about self-experience. Different contemporary accounts of selfhood that emphasize the brain as the neurological underpinning of self-experience and selfhood are distinguished: the hallucination model, the illusion model, illusionism, and virtual phenomenalism. Identifying the hallucination model with the first form of skepticism about self-experience, the focus of the analysis is on Thomas Metzinger's self-model theory of subjectivity, which is arguably the most elaborated and discussed of its sort. It can easily be seen as paradigmatic case of the first skeptical view that denies the reality of selfhood. Discussing key assumptions and arguments of Metzinger's self-nihilism, the first form of skepticism is rejected.

Chapter 2 addresses the second form of skepticism about self-experience. It focusses on an analysis of what I introduce under the label of the No-Self-Picture that is often attributed to David Hume's empiricist investigation of experience and self-experience. Applying a critique of Evander Bradley McGilvary of William James's notion of the thought as the thinker to the No-Self-Picture, it will be argued that the No-Self-Picture is untenable, as all object-experience involves self-experience. This line of argument will be supported by an examination of whether the stream of consciousness shows up in experience, which will be defended by a discussion of Dan Zahavi's interpretation of Edmund Husserl's notion of inner time-consciousness as self-consciousness. Rejecting recent criticisms of the concept of the stream of consciousness formulated by proponents of the camp of illusionism, the chapter concludes by claiming that the No-Self-Picture, according to which our experience is bereft of any self-experience,

is phenomenologically inaccurate—and that therefore the second skeptical view is unconvincing.

Chapter 3 turns to the third form of skepticism about self-experience, according to which selfhood is elusive to self-experience. It identifies Immanuel Kant's theory of self-knowledge as the paradigmatic case for the elusiveness view and discusses different problems that seem to accrue for Kant's view. The main objection is that the elusiveness view posits an unknowable  $x$  that ought to play the role of the self in a way that still resembles classical soul-theory. It will be argued that these problems are not restricted to the specifics of the Kantian system of philosophy but adhere to any way of cashing out the elusiveness view, notably, also so-called illusion model recently described by Tom McClelland.

After having rejected all three forms of skepticism about self-experience, in Chapter 4, I turn to an alternative way of conceiving the relationship between self-experience and selfhood. On this alternative understanding, self-experience defines what selfhood *is*, and is accordingly also informative about the nature of selfhood. I introduce such a position with the label of self-experientialism. One motivation to defend self-experientialism is that its alternatives would result in the acceptance—explicitly or implicitly—of one of the three skeptical views, which, according to the analyses in the first three chapters, should be avoided. Apart from this argumentation *ex negativo*, I present and discuss further arguments that may corroborate self-experientialism. To do this, first, I identify Dan Zahavi's notion of the minimal self—the thesis that pre-reflective self-awareness is the most basic form of self-experience—as the contemporary paradigmatic proponent of self-experientialism. Then I discuss and defend it in the light of recent criticisms formulated, notably, by Marie Guillot. Although Zahavi has provided short responses to Guillot's objections, I argue that a more extensive and explicit answer is needed to dispel the doubt Guillot has casted on the case for the minimal self. Such an answer, I suggest, can be given by unravelling the ontological commitments and arguments intrinsic to the minimal self-view: an ontological interpretation of self-experience in terms of self-experientialism.

In Chapter 5, I discuss alternative contemporary options for an experiential approach to selfhood that, similarly to self-experientialism, emphasize experience or self-experience as that what defines selfhood. Many of these, however, end up with slightly different proposals because they also take the issue of personal identity to constrain the debate on the notion of selfhood, and, in fact, sometimes seem to prioritize discussions of identity over the discussion of self. First, it will be argued that phenomenological self-experientialism is not to be identified with Cartesianism, despite some evident affinities. Then I work out in which sense self-experientialism bears some similarities with Locke and in which not. Doing so, I aim to disentangle the topic of self and personal identity. On the view I defend, the question of the nature of selfhood is primordial to the issue of

identity, and, in fact, if anything, the concept of selfhood constrains personal identity—and not the other way around. After this prelude, I address the accounts by Martine Nida-Rümelin, Barry Dainton, Galen Strawson, and Tim Bayne. Characteristic of these, I propose, is that they somehow prioritize the issue of personal identity in a way that confronts their accounts with ensuing problems. I conclude the chapter by emphasizing that self-experientialism does not face these challenges.

In Chapter 6, the investigation proceeds on another level, focusing on the phenomenology of self-experience. If the minimal self ought to present the most basic form of self-experience, some have argued, then it would have to be much less formal than Zahavi's account is willing to grant. The minimal self is too thin, the accusation goes accordingly. I defend a formal notion of the minimal self against this objection but suggest that what, in the eyes of its critics, seems to be missing in the minimal self-account is a convincing description of the experience of being. For, without providing a proposal of how pre-reflective self-awareness ought to convey a sense of one's own existence, it remains an open question just how pre-reflective self-awareness is a case of self-experience and not just the mere self-familiarity of experience with itself. I explore and suggest the possibility of the integration of feeling into the minimal self-account in order to address that issue, given that feeling has often been taken to present the primary consciousness of existence. Drawing from Rudolf Hermann Lotze, Edmund Husserl, and Matthew Ratcliffe, I attempt to provide a novel approach to an integration of feeling into the minimal self-account. On this view, compatible with Zahavi's minimal self-view, the most basic form of self-experience incorporates a *feeling of being oneself*.

Chapter 7 further explores the conception of the feeling of being oneself in the light of other recent notions of self-feeling in order to specify its details. I then go on to discuss the psychopathological phenomenon of depersonalization. Depersonalization consists in severe peculiarities in the way a person experiences herself, resulting in first-person reports of experience oneself as not being there. Sometimes this condition is considered and conceptualized as a case against the minimal self-view. Moreover, in a similar vein, some might potentially construe depersonalization also as a case against the notion of the feeling of being oneself. I discuss the accounts of Alexandre Billon and Marie Guillot who indeed formulate the claim that depersonalization should be considered as a condition in which experience lacks subjective character in a way that undermines the minimal self-view. I argue against these accounts by describing in which sense first-person reports of persons so-afflicted, by contrast, should be understood as a manifestation of an altered but not a lacking sense of self. Next, I discuss a classical account of depersonalization by Konstantin Traugott Oesterreich who argues that, while something like the minimal self remains intact, the individual feeling of life (*individuelles Lebensgefühl*) is eroded in depersonalization. That suggests that depersonalization is a condition in which the feeling of being oneself is

lacking. However, I argue that Oesterreich's proposal understood in this way is inaccurate. Although building on his account, I emphasize that to describe depersonalization in terms of a lack of self-feeling would be a misconstrual. For, as I suggest, people with depersonalization do care about themselves and feel great pain about how their experiences unfold. Rather than a lack of feeling of being oneself, therefore, the condition of depersonalization involves what is a very peculiar feeling of being oneself. Depersonalization is among the most extreme forms of self-experience and one of the best candidates to undermine the minimal self-view and my account of the feeling of being oneself. However, given that depersonalization, as it turns out, cannot be described as a case of lacking sense of self and self-feeling, I conclude the chapter by arguing that this speaks in favor of both the minimal self-view and, as an integral aspect of it, the feeling of being oneself.





# Part I



# Chapter 1

## The Hallucination Model of Self and the Deception of Self-Experience

Among all the aberrations of the human mind the following strikes me as the most bizarre: that it could have reached the point at which it casts doubt on its own being and essence, which only can and is immediately lived through by itself and so is known to itself; or that it would take its own being and essence as a product bestowed by some external nature that we know only by second hand and through the mediating knowing of the very mind that is cast doubt on.<sup>12</sup>

[I]f people would only realize that they are confused about the true nature of their experiences, that consciousness doesn't really exist, suffering would vanish *tout court* from the world! ... Such a denial of the authentic nature of experience is a metaphysical counterpart to Cotard's syndrome, a psychiatric condition in which patients deny being alive.<sup>13</sup>

In the past decades, a new wave of skeptical accounts with regard to what self-experiences may tell us about the reality and metaphysical nature of selfhood has emerged. Drawing from recent neuroscientific insights into the cognitive architecture of our experience and their philosophical conceptualization in the cognitive sciences, proponents of such a skeptical view argue that, while we may experience ourselves as existing and as being directly aware of ourselves, self-experiences need to be understood in terms of representational processes brought forth by our brain. Some take these deliberations as the legitimate basis for skeptical views of extreme kind, namely of the first of three forms as described in the introduction. Self-experiences, they claim, fulfill certain cognitive functions or are a by-product of the human organism's way of representing the world, but self-experiences do not refer to any real self. Accordingly, defending a variant of so-called no-self doctrine, proponents use terms such as "hallucination", "illusion" or "virtuality" to describe the ontological status of both self-experience and what self-experience is supposed to be an experience of.

In this chapter, my aim is to present some of the key arguments that underlie such accounts that place emphasis on the deceiving character of self-experience, to examine to what extent they may be persuasive, and whether it is worth adopting what has sometimes

---

<sup>12</sup> Lotze 1884/1856, 295, my translation (slightly grammatically adapted for better readability in English): "Unter allen Verirrungen des menschlichen Geistes ist diese mir immer als die seltsamste erschienen, daß er dahin kommen konnte, sein eignes Wesen, welches er allein unmittelbar erlebt, zu bezweifeln oder es sich als Erzeugniß einer äußeren Natur wieder schenken zu lassen, die wir nur aus zweiter Hand, nur durch das vermittelnde Wissen eben des Geistes kennen, den wir leugneten."

<sup>13</sup> Koch 2019, 3.

been labelled “*the hallucination model of self*”<sup>14</sup>. The hallucination model of self is a specific line of argument, the paradigmatic case of the first form of skepticism about self-experience, according to which we do indeed experience ourselves as existing selves although selves do not exist. I will first specify the idea of the hallucination model of self by distinguishing it from similar approaches. Then I present and discuss its core arguments and conclude the chapter by arguing that the hallucination model of self faces severe and ultimately unsurmountable problems that render it unattractive as an approach to selfhood.

## 1.1 Focus on self-experience and the no-self doctrine

The hallucination model of self is not the only theory family according to which self-representational processes are fallacious. In current—mostly analytic—philosophy of mind approaches that center around such ideas are widespread.<sup>15</sup> Yet, despite some agreement on the deceiving character of self-representational processes, these accounts sometimes vary heavily in how they conceive of the deception they take inherent to self-representation, how they understand the nature of self-representation, and in how they assess the ontological status of self. It is therefore important to keep different lines of thinking well apart from each other, especially where they draw different conclusions with regard to the ontological nature of selfhood. My interest in this chapter is with the specific no-self-doctrine that the hallucination model of self presents and is therefore highly focused. Let me, therefore, first explain what precisely the hallucination model of self consists in and how it differs from similar approaches.

### 1.1.1 Virtual self theory (VST) and the hallucination model of self

First, a clarification regarding the relationship between so-called virtual self theory (VST) and the hallucination model of self is needed. In a recent paper, McClelland describes VST in the following way:

According to the virtual self theory (VST), selves are merely virtual entities. On this view, the self is an elaborate fiction generated by the brain. You are *represented* to exist, but these representations do not correspond to any real

---

<sup>14</sup> McClelland 2019, 28, my emphasis.

<sup>15</sup> To give only a few examples: Blanke & Metzinger 2009, Dennett 1992, Frankish 2017, Hood 2011, Limanowski 2014, Limanowski & Blankenburg 2013, Metzinger 2003, 2009a, b, 2013, Northoff 2004a, b, Seth & Tsakiris 2018.

entity. As such, selves are *merely intentional* entities. VST is a contemporary version of the ‘no-self’ theory.<sup>16</sup>

Prima facie, it seems that conceptualizing selves as virtual entities necessarily amounts to the assumptions of the hallucination model of self: we possess self-experience, yet self-experience fails to refer to anything real external to consciousness and must therefore be characterized as a hallucinatory experience. However, things are more complicated. The following reasons suggest that the relationship between VST and the hallucination model of self are better understood in terms of a partial overlap rather than congruency: Some VST-accounts may be considered a hallucination model, while others not. Bruce Hood and Thomas Metzinger, claiming that the self is only a hallucination or illusion, explicitly deny the reality of self, although they admit that, given the way our experience of the world is phenomenologically structured, it appears as if we had or were selves.<sup>17</sup> But not all defenders of VST draw such drastic conclusions from the fact that, beyond our experience, there is no external object or entity to which our self-experience could legitimately be said to refer to. Tim Bayne’s “virtual phenomenalism”, for instance, despite stressing “that we should think of the self as a merely virtual center of ‘phenomenal gravity’”<sup>18</sup>, thereby “[a]ppropriating Dennett’s (1992) notion of the self as a center of narrative gravity”<sup>19</sup>, does *not* endorse the no-self-doctrine. He makes it perfectly clear that the latter, in his eyes, is untenable a position: “[N]ot only is the denial of the self of dubious coherence—it is certainly not how virtual phenomenalism was advertised. I said I was going to provide an account of the self, not an account that explained the self *away*.”<sup>20</sup> Most importantly, Bayne also highlights that while he believes that no referential relationship between any appearance of being a self and an object external to phenomenal consciousness exists, this does not mean that representations of self do not refer at all:

We can now see where other approaches to the self go wrong: they assume that there must be some ‘real’ entity that plays the role of the self. The only thing that plays the self role—indeed, perhaps the only thing that *could* play the role of the self—is a merely intentional entity.<sup>21</sup>

As the quotation shows, Bayne’s virtual phenomenalism argues that selfhood just consists in being a merely intentional entity, and that for this reason to expect a “real” object—“real” in the sense of a physical object—is wrong-headed. VST, thus, unlike McClelland

---

<sup>16</sup> McClelland 2019, 21f.

<sup>17</sup> Hood 2011, Metzinger 2003.

<sup>18</sup> Bayne 2010, 289.

<sup>19</sup> Bayne 2010, 289.

<sup>20</sup> Bayne 2010, 292f.

<sup>21</sup> Bayne 2010, 290.

seems to suggest, not in all cases amounts to a no-self-view, but is sometimes rather construed as an attempt to understanding the specific ontological character the self or selfhood has as compared to worldly objects.

### 1.1.2 Hallucination vs. illusion model of self

There is another important distinction to consider in order to adequately grasp the hallucination model. As McClelland emphasizes, casting doubt on the epistemological value of self-representation, different interpretations of self-representation remain possible. To be more precise, there are different ways in which one could say self-representations may deceive us: One might also interpret the fallacious self-representations as *misrepresentations* of the self and its nature rather than as representations that simply falsely represent something that does not exist. Hence, from the fact that our self-representations, the way we experience ourselves, doesn't refer to an alleged *immaterial* self-substance or mental self doesn't follow that self-representations don't refer at all to something, for instance, a material self. On what McClelland dubs the *illusion model of self*, self-representation deceives us when it comes to our being, but it doesn't deceive us in suggesting that we exist. According to the view of the illusion model, McClelland emphasizes, self-representations “systematically and dramatically misrepresent what features the self has”<sup>22</sup> and takes the self to be “a real material entity that presents illusory appearance of being immaterial”<sup>23</sup>. Most importantly, thus, while the hallucination and illusion model of self share the notion that self-representation does not indicate the true metaphysical nature of self, the former explains this by emphasizing that the self is not real and therefore has no metaphysical nature other than its merely intentional or virtual character. By contrast, the illusion model of self simply holds that we cannot conclude from self-representations to the metaphysical nature of self and that this does not render the self unreal.

It is therefore important to conflate the hallucination model neither with Bayne's virtual phenomenalism nor with the illusion model of self.

### 1.1.3 Self-representation and self-experience

Third, a further clarification with regard to the relationship between self-representation and self-experience is required. Most importantly, it is necessary to bear in mind that, in—mostly analytic—philosophy of mind, the notion of “self-representation” is understood in

---

<sup>22</sup> McClelland 2019, 28.

<sup>23</sup> McClelland 2019, 28.

at least two ways: first, as the self-representation of a mental act by virtue of which the very act becomes a *conscious* act;<sup>24</sup> second, as the self-representation of a cognitive system that underlies the activity of representational processes, i.e., the representation through which the cognitive system that does the representing represents itself.<sup>25</sup> These ways of understanding “self-representation” are not conceptually congruent. Precisely because this is so, speaking of “self-representation” without further specification is misleading and gives rise to an important question: Does “self-representation” in the second sense, the representation of a cognitive system of itself, necessarily involve phenomenal consciousness or experience? Put differently, when speaking of the self-representation of a cognitive system, does this necessarily involve a form of self-*experience*, i.e., a certain ‘*what-it-is-like*’<sup>26</sup> to represent oneself as the cognitive system doing the representing? Given how “representation” is generally understood in the—mostly analytic—philosophy of mind, there seems to exist no good reason to assume that when authors speak of “self-representation” they also mean by the same token a form of “self-experience”. For, “representation” as a general notion of the—mostly analytic—philosophy of mind is usually understood in terms of *aboutness* or *intentional directedness*, which need not necessarily involve experience or phenomenal consciousness.<sup>27</sup> Accordingly, to speak of “self-representation” in the second sense without further specification only refers to the fact that a cognitive system that self-represents is intentionally directed at itself. For instance, an intelligent machine may be capable of processing information of its environment and of itself such as location, temperature, battery status and so on. Consider a robot vacuum cleaner. Scanning its environment for objects such as tables, chairs or the walls to determine the surface to be cleaned, it represents its environment. At the same time, it also represents itself by keeping track of its own location so it can stop and change direction in the right moment before touching, for instance, an expensive Rococo chair. Does this kind of self-representation include self-experience in that there is something it is like for the robot to represent its location, i.e., itself? For most of us, it will seem impossible to claim that it does. Hence, self-representation and self-experience need to be kept apart conceptually: Not all self-representation amounts to a form of self-experience. Since my interest is with self-experience, more precisely, whether self-experience could be a mere hallucination, in discussing the hallucination model of self, I am only discussing those accounts which refer to self-experience when they speak of self-representation. Moreover, it strikes me as confusing, if not wrong outright, to speak of a *hallucinatory* self-

---

<sup>24</sup> Carruthers 2000, 2005, 2016, Caston 2002, Kriegel & Williford 2006, Kriegel 2006, 2009, 2013, McClelland forthcoming a, Van Gulick 2001, 2004.

<sup>25</sup> Hohwy & Michael 2017, Metzinger 2003, McClelland 2019.

<sup>26</sup> Cf. Nagel 1974.

<sup>27</sup> Pitt 2018.

representation that doesn't involve phenomenal consciousness, i.e., experience. In my view, thus, any proponent of the hallucination model of self when talking about a hallucinatory self-representation must mean "self-experience".

#### 1.1.4 The hallucination model of self vs. illusionism

A fourth clarification is needed. It concerns the relationship between the hallucination model of self and what has been labelled *illusionism*<sup>28</sup>, which is not to be conflated with the illusion model of self I mentioned in Section 1.1.2. Illusionism, far from being a homogeneous approach, is the general view that our experiences seem to have phenomenal properties when really there are not such things as phenomenal properties at all.<sup>29</sup> According to Frankish, the illusion problem is to substitute the "hard problem of consciousness"<sup>30</sup>, which is well-established in the philosophy of mind. Rather than asking how it is that phenomenal consciousness may arise out of what are purely physical processes, illusionism emphasizes the question how it is that there seems to be something like phenomenal consciousness to begin with.<sup>31</sup> Clearly, illusionism and its denial of phenomenal consciousness has repercussions for any account of self-experience.

For instance, Blackmore maintains that there is no stream of consciousness, no contents of consciousness, no continuity of consciousness, and accordingly no unity of consciousness. From this, she infers that no self exists that would underlie our experience as its subject.<sup>32</sup> Where there is no continual stream-like consciousness, there is no self-experience and it does not make sense to speak of an abiding subject of experience. Garfield speaks it out more directly when he says "[t]hat there is nothing that it is like to be me"<sup>33</sup> and that thinking otherwise is nothing but a "cognitive illusion"<sup>34</sup>. Unlike the hallucination model of self (and the illusion model of self), illusionism, thus, does not claim that our self-experience falsely represents the reality of self or misrepresents the nature of self, rather, illusionism highlights that the very idea that we enjoy self-experience is illusory. Hence, illusionism denies that there is self-experience that could turn out to be right or wrong to

---

<sup>28</sup> Frankish 2016, 2017

<sup>29</sup> Frankish 2016, 12, 15.

<sup>30</sup> Chalmers 1995, Shear 1999.

<sup>31</sup> Frankish 2016, 37. This move relaxes the pressure on physicalism about phenomenal consciousness to explain how consciousness can emerge from neurophysiological processes. Put differently, it ought to avoid what is typically considered the *hard problem of consciousness*. For, if phenomenal consciousness doesn't exist it is superfluous to ask how it could arise out of neurophysiological processes. The alternative challenge for the illusionist, then, is to demonstrate how it is that such an illusion about phenomenality develops in the context of neurophysiological information processes.

<sup>32</sup> Blackmore 2016, 59ff. See also Blackmore 1999.

<sup>33</sup> Garfield 2016, 74.

<sup>34</sup> Garfield 2016, 82.



begin with. This, of course, is a completely different version of the no-self view. I address such a line of thinking in Chapter 2. Here, my interest is with the following set of questions: Provided we *do* have self-experience, i.e., we experience ourselves as existing, is it possible to doubt the reality of self? Could it be that I do not exist although I experience myself as being someone, as the hallucination model of self suggests?

### 1.1.5 Skepticism about self-experience and the hallucination model

As the last point has shown, the hallucination model presents a case of the first kind of skepticism about self-experience that I have described in the Introduction of the thesis. On this first skeptical view, we enjoy some form of self-experience, but self-experience does neither reveal the existence nor the being of selfhood. In fact, the first skeptical view denies the reality of selfhood, although it accepts that our experience of the world involves self-experience, though one that is fallacious and misleading. Accordingly, it is a skepticism not about the phenomenology of self-experience, but about the—allegedly lacking—relevance of the phenomenology of self-experience for our knowledge of the *metaphysics* of selfhood.

Above, I distinguished the *hallucination model of self* from similar accounts, such as *illusionism*, the *illusion model of self*, and *virtual phenomenalism*. Crucially, the differences between these accounts also translate into differences concerning the skeptical (or non-skeptical) attitude that is involved in each of them. As it turns out, each of them presents an example of one of the three forms of skepticism about self-experience or, as in the case of virtual phenomenalism, of an experiential account of selfhood that makes no skeptical statement about the phenomenology of self-experience. The following table lists the different accounts and shows how the core idea of each of them corresponds to a certain form of skepticism about self-experience—with the exception of virtual phenomenalism.

<p><b>Hallucination model of self</b> (1<sup>st</sup> skepticism: Chapter 1)</p>	<ul style="list-style-type: none"> <li>- accepts self-experience</li> <li>- denies reality of selfhood (existence)</li> <li>- denies that self-experience reveals being</li> </ul>	<p>All phenomenal consciousness is hallucinatory. Our experience involves self-experience but is hallucinatory and so doesn't refer to any self. The self is not real.</p>
<p><b>Illusionism</b> (2<sup>nd</sup> skepticism: Chapter 2)</p>	<ul style="list-style-type: none"> <li>- denies self-experience</li> </ul>	<p>There is no phenomenal consciousness or there is at least no self-experience.</p>
<p><b>Illusion model of self</b> (3<sup>rd</sup> skepticism: Chapter 3)</p>	<ul style="list-style-type: none"> <li>- accepts self-experience</li> <li>- accepts reality of selfhood (existence)</li> <li>- denies that self-experience reveals being</li> </ul>	<p>We do have self-experience and selfhood is real. But self-experience gives a misleading impression of what selfhood really is.</p>
<p><b>Virtual phenomenalism</b> (No skepticism: Chapter 5)</p>	<ul style="list-style-type: none"> <li>- accepts self-experience</li> <li>- accepts reality of selfhood (existence)</li> <li>- no statement on revelatory character of self-experience</li> </ul>	<p>Our experience of the world involves self-experience. Selfhood is real in experience but it is nothing over and above our experience. Its being and reality consist in being a merely virtual or intentional entity only.</p>

## 1.2 The neuroscientific study of self-experience

Since in this chapter my interest is with the hallucination model of self, the question arises as to which accounts are most suitable in order to investigate key tenets of the hallucination model. Typically, it is philosophers interested in neuroscience and the cognitive sciences who develop a disposition for endorsing a variant of the hallucination model. Just like experiences involved in processes of perception, emotions, and cognition, they suggest, self-experiences can equally become the subject of neurocognitive and neurobiological study.<sup>35</sup> Among the questions they address are:

- What are the neurobiological underpinnings of self-experience?
- What are the brain mechanisms that underlie the development of a sense of self?
- What is the role of processing information revolving around the interaction with the world for experiencing oneself as an agent?

As these few examples demonstrate, self-experience can be and has been made a subject of *empirical* study. Rather than taking self-experience as some kind of epistemic principle that is to serve to enable metaphysical claims about the reality and nature of self, as the tradition had it, the empirical study of self-experience declares self-experience itself an explanandum. Such a shift in attention, intrinsic to the methodological setting of the empirical study of self-experience, has often been accompanied by a shift in assessing the ontological status of self. If self-experience is based on neurological processes, notably brain mechanisms, the rationale goes, doesn't this mean that theories about the neurological processes underlying self-experiences are a better source for assessing the ontological status of self than the self-experiences? Accordingly, aren't thus neuroscientific theories of self-experience the best candidates in explaining what we are, what our metaphysical nature is?

Naturally, not all empirical scientists addressing self-experience have engaged in this kind of philosophical deliberations. Yet, those more prone to or stemming from philosophy have sometimes not shied away from drawing drastic conclusions from neuroscientific studies and—most importantly—from speculating what they *expect* neuroscientific studies will reveal in the future. Prominent examples in that regard are neurophilosophical frameworks such as Metzinger's *self-model theory of subjectivity*<sup>36</sup> or, more recently, the *predictive coding framework*<sup>37</sup>. Surveying many neuroscientific studies,

---

<sup>35</sup> To name just a few examples of the massive and still growing body of literature: Apps & Tsakiris 2014, Blanke & Metzinger 2009, Hohwy 2013, Letheby 2020, Lutz & Thompson 2003, Metzinger 2003, 2013, Millière 2017, Seth & Tsakiris 2018, Trehub 2007, Wiese 2018, 2019.

<sup>36</sup> Metzinger 2003, 2005, 2009a, 2011.

<sup>37</sup> For instance, Clark 2002, 2015, Friston 2010, Frith 2007, Hohwy 2013, Hohwy 2016, Hohwy & Michael 2017, Limanowski & Blankenburg 2013.

proponents of these accounts examine how best to understand and reconstruct the processes through which self-experiences emerge. They also develop philosophical conceptions regarding the reality and metaphysical nature of self. Metzinger, without a doubt, offers the most extreme view of self. Arguably presenting the most elaborate case for the hallucination model of self, he denies the reality of self based on the assumption that our experiential self-representation of ourselves doesn't refer to anything, given that all experience, phenomenal consciousness, is hallucinatory. While Metzinger explicitly and unambiguously denies the reality of self and reduces selfhood to *phenomenal* selfhood, a self-model built-up on representational processes, proponents of predictive coding are more ambivalent in their statement on the ontological status of self. Consider, for instance, Jakob Hohwy's *The Predictive Mind* (2013). In his final remarks of the book, Hohwy highlights that Metzinger's idea of the self as a mere phenomenal self-model that otherwise does not exist and bears no reality is "a reasonable way to go with the notion of the self"<sup>38</sup>. Crucially, he takes it to be compatible with the central hypothesis of the predictive coding framework, *prediction error minimization*, which consists in the view that the brain has the function to continuously reduce the prediction error with regard to information that the organism processes in interaction with its environment. On the view of the predictive coding framework, "perception, cognition, and action are manifestations of a single adaptive regime geared to the reduction of organism-salient prediction error"<sup>39</sup>. Self-experience can thus be seen as an aspect of an experiential process that ultimately serves the purpose for an organism to deal with its environment and to minimize the prediction error, thereby guiding the organism's information-processing and behavior that ensues from it. This, indeed, seems to be compatible with the no-self view akin to Metzinger.

However, although it is certainly possible to construe an account of predictive processing in terms of a no-self view similar to Metzinger's self-nihilism, neither does self-nihilism necessarily ensue from the prediction error minimization view nor do its proponents seem to converge in their views on the ontological status of self.<sup>40</sup>

In fact, even in single thinkers the view of the ontological status of self remains often blurry. As, for instance, Hohwy and Michael highlight, they "agree with much of this [Metzinger's denial of self, P.S.]", but yet intend to "offer a seemingly more metaphysically robust account of self-representation in terms of inferred hidden causes"<sup>41</sup>. What they agree with is that, in interacting and representing the world, a self-model is construed. What is unclear though is whether they ultimately also conclude from this that the self is unreal. It does not seem so. For, they describe the self as "the set of endogenous causes being referred

---

<sup>38</sup> Hohwy 2013, 255.

<sup>39</sup> Clark 2016, 138.

<sup>40</sup> Cf., for instance, Beni 2019, Kiverstein 2020.

<sup>41</sup> Hohwy and Michael 2017, 374.

to by self-models”<sup>42</sup>. On their view, self-models emerge where “hidden patterns of endogenous causes”<sup>43</sup> of an agent’s experiences, sensory inputs or other psychological properties are represented. Yet, they do not maintain that the self only *appears* to be the set of hidden causes that best explains how the experiential life of the agent unfolds. They say it *is* the set of endogenous causes. The self features in the world-model as a cause on a par with other, though, external causes.<sup>44</sup> Moreover, Hohwy and Michael speak of a “realist interpretation of intentional states and of self”<sup>45</sup> and conclude: “In summary, the predictive processing approach to brain function has a central role for both body and self, as *inferred entities* in the world that are represented in the agent’s internal model of the world.”<sup>46</sup>

While the predictive coding framework shares with Metzinger’s self-model theory of subjectivity the idea that a cognitive system, in representing the world, develops a useful self-model, it is not sufficiently clear whether the predictive coding framework amounts to a no-self doctrine. Given that this is so, my following analysis of the hallucination model of self will focus on Metzinger’s account and the key assumptions it relies. Even though his neurophilosophical proposal is not the newest on the market, when it comes to the philosophical assessment of self, it is, in its radicality and explicitness, unrivaled. Moreover, philosophical criticisms of his theory—abundant in the philosophical literature of the past two decades—have a broader application and extend to those proposals of predictive processing that are intended as being in line with Metzinger.<sup>47</sup> In this sense, Metzinger is a or *the* paradigmatic example of the hallucination model of self.

### 1.3 Metzinger’s *self-model theory of subjectivity* and hallucination

Metzinger’s self-model theory of subjectivity is by far the most detailed attempt at trying to explain how it comes that our experiential life includes forms of self-experience and how we come to believe that we are someone when in reality “no such things as selves exist in the world”<sup>48</sup>, as Metzinger lets us know on the first pages of his seminal book *Being No*

---

<sup>42</sup> Hohwy and Michael 2017, 385.

<sup>43</sup> Hohwy and Michael 2017, 385.

<sup>44</sup> “The self is just one set of causes among other causes of sensory input. By minimizing prediction error, the self can be modeled just as other causes can be modeled.” (Hohwy and Michael 2017, 374)

<sup>45</sup> Hohwy and Michael 2017, 386.

<sup>46</sup> Hohwy and Michael 2017, 386 (my emphasis).

<sup>47</sup> Those proposals of predictive processing that do not subscribe to the no-self doctrine in that they suggest the self is an endogenous cause that features in our model of the world take the self to be real but distinct to what our self-experience suggests. They will be covered by my discussion of the third skeptical view about self-experience in Chapter 3.

<sup>48</sup> Metzinger 2003, 1.

*One* (2003). Bearing in mind the three claims Descartes had made, as described in the Introduction, it is immediately evident that Metzinger offers an alternative story about our self-experience. He agrees with Descartes that we have self-experience (claim 1). But he vehemently denies that we could infer from self-experience that we exist (claim 2), let alone what our metaphysical nature is (claim 3). What is his alternative story?

The upshot is this: “[T]he subjective experience of being someone emerges if a conscious information-processing system operates under a transparent self-model.”<sup>49</sup> Regardless of how we experience ourselves and how these experiences may be phenomenologically described, in each case we are underlying a significant mistake: We “constantly confuse” ourselves “with the content of the self-model” that is being run by our brain.<sup>50</sup> What exactly does Metzinger mean by that confusion? How can I confuse myself with something else? The confusion lies in the idea that I take *my* experiences to be my experiences, i.e., as experiences had by a self, had by someone, had by me. According to Metzinger, my experiences only appear as if they were had by myself, that is, someone. I mistakenly take the content that constitutes my conscious life to be myself. On Metzinger’s view, however, what really is happening is that there is nobody else than just a biological organism with a brain that, fulfilling certain criteria,<sup>51</sup> simulates the phenomenal experience of being a self: “Biological organisms exist, but an organism is not a self. Some organisms possess conscious self-models, but such self-models certainly are not selves—they are only complex brain states.”<sup>52</sup>

The alternative picture, thus, Metzinger offers, consists in the idea that only biological organisms can be said to exist and that the “phenomenal property of selfhood as such is a representational construct; it truly is phenomenal property in terms of being an appearance only”<sup>53</sup>. The representational construct, Metzinger argues, comes about when a biological organism self-represents its own physical states in a certain way—in a way that suggests that there was a self that had experiences. Really what is represented in self-representation are “properties of the physical organism”<sup>54</sup>: the organism’s “internal chemical profile”, “states of its internal organs, its spatial and kinesthetic properties, or relational properties with regard to other objects and agents in its environment”<sup>55</sup>. And, accordingly, *that* is what self-reference ultimately refers to: the physical organism and its

---

<sup>49</sup> Metzinger 2003, 1.

<sup>50</sup> Metzinger 2003, 1.

<sup>51</sup> Metzinger 2003, 305ff.

<sup>52</sup> Metzinger 2003, 563. This statement distinguishes the hallucination model from the illusion model that I will discuss in Chapter 3. Unlike the hallucination model, the illusion model argues that the organism or the brain *is* the self.

<sup>53</sup> Metzinger 2003, 563.

<sup>54</sup> Metzinger 2003, 272.

<sup>55</sup> Metzinger 2003, 272.

properties. It is not the alleged self that is carrying out the cognitive act of self-reference, but the physical system itself by producing the phenomenal self-representational states: “[A]ny conscious system is justified in thinking the I\*-thought that at least some physical representational system exists, which right now carries out the act of cognitive self-reference.”<sup>56</sup>

Accordingly, to say that *I* am confused about my existence is saying already too much. Speaking of a confusion, “[we] now know that this was only an introductory metaphor”<sup>57</sup>: “There is no one *whose* illusion the conscious self could be, no one *who* is confusing herself with anything.”<sup>58</sup> Really, there is no one at all: “What exists are information-processing systems engaged in the transparent process of phenomenal self-modeling.”<sup>59</sup>

One might wonder why it is that biological organisms come to represent the world in a way that comes with a self-experience of the kind we enjoy. Metzinger has a clear, naturalistic answer: Our brain, as a control center of the organism, operates with a phenomenal self-model to deal with the comprehensive information and challenges the organism is confronted with in its natural environment: “Generating a single and coherent self-model is a strategy for reducing ambiguity.”<sup>60</sup> Operating with a phenomenal self-model, then, serves a function, it allows the organism to reduce and select information according to its needs. The “virtual organ”<sup>61</sup>, as Metzinger describes the phenomenal self-model, turns out to be an “informational-computational strategy”<sup>62</sup> that presents a significant advantage from an evolutionary perspective.<sup>63</sup>

To conclude, Metzinger’s view is that self-experiences are generated by the biological organism that self-represents its physical states. In his picture, self-experiences ultimately refer to the biological organism which should not be understood in terms of “self” conceived along the lines our self-experience seems to suggest, i.e., (as Metzinger believes) as an immaterial or experiential entity that has experiences. Since self-experiences have no selves as referents and selves do not exist, they are hallucinatory experiences. Whatever our self-experience might suggest, it is fallacious and deceptive. You, I, we do not exist. Obviously, these are tough claims. What reasons does Metzinger give us to make his case?

---

<sup>56</sup> Metzinger 2003, 404f.

<sup>57</sup> Metzinger 2003, 634.

<sup>58</sup> Metzinger 2003, 634.

<sup>59</sup> Metzinger 2003, 337.

<sup>60</sup> Metzinger 2003, 316.

<sup>61</sup> Metzinger 2003, 200.

<sup>62</sup> Metzinger 2003, 312, 316, 337, 343.

<sup>63</sup> Metzinger 2003, 345, 346, 348, 352, 420, 563.

## 1.4 *Being no one*: Key assumptions and arguments

Reading Metzinger's "gargantuan work"<sup>64</sup> of over 600 pages poses a challenge and the complexity of his position is immense. Yet, it is possible to pin down the key assumptions and arguments that underlie his general storyline: that we do not exist. To understand his rationale, it is important to bear in mind that Metzinger's no-self doctrine operates at least on two levels: the phenomenological level of our experience and the subpersonal level of the representational processes that he says underlie our experience and its phenomenology. Let me start my reconstruction of the line of his main argument by describing what Metzinger has to say on the phenomenology of self-experience.

### 1.4.1 The "phenomenological fallacy"<sup>65</sup>

One important element of Metzinger's refusal of the reality of self concerns the critique of self-experience—or better, the way self-experience is usually interpreted. As I have already highlighted, Metzinger, although defending the no-self view and unlike Hume's version thereof, does acknowledge that our human consciousness, the way we represent the world, includes self-experience:

There seems to be a primitive and prereflexive form of phenomenal self-consciousness underlying all higher-order and conceptually mediated forms of self-consciousness, and this nonconceptual form of selfhood constitutes the origin of the first-person perspective. Phenomenal selfhood is what makes us an experiential subject.<sup>66</sup>

It is not only that a world is present; it is that I am a *present self* within this world.<sup>67</sup>

Phenomenal experience not only consists in 'being present'; it also consists in 'being present *as a self*.'<sup>68</sup>

Phenomenologically speaking, if the specious present of a psychological moment is integrated with a self-representation, then it will ... lead to the conscious experience of *someone's existence*.<sup>69</sup>

Autoepistemic closure, as represented by the phenomenal transparency constraint being maximally satisfied for large partitions of the human self-

---

<sup>64</sup> Harman 2011, 32.

<sup>65</sup> Metzinger 2003, 22, 23, 94, 272, 325, 446.

<sup>66</sup> Metzinger 2003, 158.

<sup>67</sup> Metzinger 2003, 310.

<sup>68</sup> Metzinger 2003, 311.

<sup>69</sup> Metzinger 2003, 311.

model, must by necessity lead to the prereflexive phenomenal property of ‘selfhood,’ which simply is the certainty about one’s own existence ...<sup>70</sup>

As this selection of quotations demonstrates, Metzinger is happy to acknowledge that there is something like “phenomenal selfhood”<sup>71</sup> or the “experience of being someone”<sup>72</sup>. Yet, what he criticizes is that this phenomenal property is taken to indicate the real existence of an entity, a nonphysical-kind of object, the subject or self. According to Metzinger, what we do find in introspection is that our experiences are characterized by a form of “subjectively experienced *immediacy* of subjective, experiential content”<sup>73</sup>. On Metzinger’s view, all we can find in introspection is some phenomenal content given in a phenomenally immediate or direct way. Yet, the fact that experiential content is given in an immediate way leads to what Metzinger calls a “necessary self-reification”<sup>74</sup>:

If naively we speak of a ‘content of self-consciousness’ or *the* ‘content of a phenomenal self-representation,’ we reify the experiential content of a continuous representational process. This process is now frozen into an object, which this time is our ‘own’ self, that is, a *subject*-object. We automatically generate a phenomenal individual and thereby run the risk of repeating the phenomenological fallacy. The core of the fallacy consists in the unjustified use of an existential quantifier within the scope of a psychological operator, this time when referring to self-consciousness.<sup>75</sup>

It is the general phenomenal feature of ‘immediacy’ of our experience that makes us believe that we are in direct contact with *ourselves*, for which reason we come to believe to *be someone*: “My own existence possesses temporal immediacy: a sense of being in touch with myself in an absolutely direct and nonmediated way, which cannot be bracketed.”<sup>76</sup> That gives rise to what Metzinger calls a “naïve-realistic self-misunderstanding”<sup>77</sup>. Because our experiences seem to be immediately given to us, we feel that we are somehow directly aware of ourselves as being entities that represent themselves to themselves: “You simply seem to know that all this is an ongoing process of self-representation, while at the same time you are certain that you are the initiator of this process *yourself*.”<sup>78</sup> But the problem is that, according to Metzinger, such immediacy and direct self-reference is fallacious.

---

<sup>70</sup> Metzinger 2003, 398.

<sup>71</sup> Metzinger 2003, 302.

<sup>72</sup> Metzinger 2003, 1, see also 2, 5, 311, 337, 626.

<sup>73</sup> Metzinger 2003, 125, see also 99, 170, 284.

<sup>74</sup> Metzinger 2003, 338.

<sup>75</sup> Metzinger 2003, 272.

<sup>76</sup> Metzinger 2003, 310.

<sup>77</sup> Metzinger 2003, 333.

<sup>78</sup> Metzinger 2003, 333.



The first reason is phenomenological: “[D]escriptions of phenomenal content do not refer to a privileged phenomenal individual—for example, ‘the self’—but only to an introspectively accessible time slice of the actual representation process—that is, to the content of this process *at t*.”<sup>79</sup> Put differently, if we describe experiential content phenomenologically, we will only be able to detect some specific content given in an immediate way and not an object “self”. As Metzinger emphasizes: “The conceptual reification of what actually is a very unstable and episodic process ... is even phenomenologically false: we are not things, but processes.”<sup>80</sup>

The second reason concerns the notion of “direct reference”. We may experience ourselves as immediately present when experiential content is given to us in an immediate way. We may feel that, in experiencing content, we are at the same time enjoying a feeling of being someone to whom said experiential content is given. Yet, as Metzinger is adamant to highlight, none of the many cases of self-reference are ever direct: “There simply is no such thing as epistemically immediate contact to reality. What there is is an efficient, cost-effective, and evolutionary advantageous way of phenomenally modeling reliable representational contents as *immediately given*.”<sup>81</sup> As this is a further crucial assumption and element in Metzinger’s overall argument, let me explain it in more detail.

#### 1.4.2 The impossibility of direct (self-)reference

Metzinger concedes that we appear to have some direct awareness of ourselves and that corresponding experiences—or better to say—the immediacy of the givenness of all our experiences seem to imply that we are directly referring to ourselves. But for Metzinger, to draw from that experience to the belief that we really are directly referring to ourselves is based on the lack of a distinction between phenomenal immediacy and epistemic immediacy:

*Phenomenal* immediacy does not entail epistemic immediacy.<sup>82</sup>

[M]any bad philosophical arguments concerning direct acquaintance, infallible first-person knowledge, and direct access are clearly based on an equivocation between epistemic and phenomenal immediacy ...<sup>83</sup>

---

<sup>79</sup> Metzinger 2003, 272.

<sup>80</sup> Metzinger 2003, 325.

<sup>81</sup> Metzinger 2003, 599.

<sup>82</sup> Metzinger 2003, 599.

<sup>83</sup> Metzinger 2003, 170.

But why should one draw such a distinction according to Metzinger? Why shouldn't we infer from the seemingly direct awareness of ourselves to the belief that we really possess direct access to ourselves? The problem with this kind of experience-based belief is, Metzinger claims, that we oversee that these seemingly *immediately* given experiences are only given to us mediated by representational processes through which a self-model is construed. Put differently, self-experience triggered by the immediacy of our experiences only appears to be direct, when in fact it is indirect and only possible by representational processes that, as it just so happens, do not manifest themselves experientially:

Of course, it is true that only the experiencing subject can refer to its own phenomenal states as phenomenal states. But as we now know, no kind of self-reference is ever truly direct, because it is inevitably mediated through the self-model, because it crucially depends on the subpersonal self-organization of the relevant construction mechanisms, which are introspectively invisible, that is, transparent to us.<sup>84</sup>

What phenomenologically might seem to be direct or immediate givenness, really is indirect and mediated through the representational processes carried out by the physical system, the real bearer of experiences.

### 1.4.3 The transparency of representation and the sense of self

Any experience, since ultimately based on physical processes, is mediated by representational processes. How then, one might wonder, does the experience of direct acquaintance come about? Why do we experience ourselves as having direct access to ourselves? Metzinger has a reductionistic explanation for this: It is precisely because the representational processes, which are really carried out by the physical system, are not phenomenally manifest, i.e., not represented, that experiences gain a direct and immediate character:

The *instruments* of representation themselves cannot be represented as such anymore, and hence the experiencing system, by necessity, is entangled in a naive realism. This happens because, necessarily, it now has to experience itself as being in direct contact with the current contents of its own consciousness.<sup>85</sup>

Since the mediation or production of the content of consciousness through the brain is not phenomenally given, the content of consciousness appears to be not-mediated, that is,

---

<sup>84</sup> Metzinger 2003, 578.

<sup>85</sup> Metzinger 2003, 169.

immediate. The representational processes are something we look right through, in Metzinger's words, they are 'transparent':

The transparency of the self-model is a special form of inner darkness. It consists in the fact that the representational character of the contents of self-consciousness is not accessible to subjective experience.<sup>86</sup>

Obviously, to speak of 'inner darkness' is a metaphor, and perhaps a bad one. But what Metzinger is saying is that when trying to pin down the representational processes through which phenomenal content is given to us, we end up just with the phenomenal content directly. The production of the phenomenal content *as production* does not come into phenomenological sight. And, therefore, the "phenomenology of transparency is the phenomenology of apparently direct perception", which ultimately amounts to "the subjective impression of immediacy"<sup>87</sup>. And, as mentioned before, it is the immediacy of experiences that gives rise to the "nonepistemic feeling"<sup>88</sup> of being someone. Therefore, Metzinger, can say:

The prereflexive, preattentive experience of being someone results directly from the contents of the currently active self-model being transparent.<sup>89</sup>

[T]he subjective experience of being someone emerges if a conscious information-processing system operates under a transparent self-model ...<sup>90</sup>

The phenomenal property of selfhood is constituted by transparent, nonepistemic self-representation ...<sup>91</sup>

Hence, whatever our conscious experience might suggest concerning our own existence as selves or the direct awareness thereof, it must be understood as some form of trick played to *us* by our brain, whereby the *us* still gets the reductionistic picture wrong. Really, the brain, the physical organism that is, is tricking itself:

It is this system, for example, the central nervous system of a biological organism, which really is the thinking thing. It generates *cogitationes* in the form of what I have called in Chapter 3 phenomenal mental models. However, as it is not able to internally represent those models *as models*..., it is not able to recognize its phenomenal ego—that is, the mental model of a *res cogitans*—

---

<sup>86</sup> Metzinger 2003, 331.

<sup>87</sup> Metzinger 2003, 169.

<sup>88</sup> Metzinger 2003, 287.

<sup>89</sup> Metzinger 2003, 337.

<sup>90</sup> Metzinger 2003, 1.

<sup>91</sup> Metzinger 2003, 337.

as a product of its own, ongoing internal representational dynamics, but ‘confuses itself’ with the content of this model.<sup>92</sup>

Hence, according to Metzinger, the physical system, instead of being aware of itself as an experience-producing system, represents itself by constructing a phenomenal self that the physical system then (re)presents as the subjective point of view or the bearer of experience. But really, the phenomenal self is just an invention of the physical system, it is made up.<sup>93</sup>

#### 1.4.4 Self-experience, indirect reference, and the lack of a “self“-referent

As Metzinger points out, self-experience can only be considered to present a form of indirect reference. Even so, one might then want to ask, doesn't it still refer, and don't we, thus, have to infer that there is a self, self-experience indirectly refers to? As I have already mentioned above, Metzinger admits that self-experience has a referent, namely the physical system, the organism. Self-experience is the very—indirect—way through which the organism represents itself. But another question ensues from that: Why does the organism represent itself in a distorting way by construing a self-model? Why doesn't the organism represent its own representational processes and provide a more truthful way of being aware of itself and its own states? The problem, as Metzinger emphasizes, is that this just does not seem to be possible. If the organism were to represent its own representational processes, Metzinger argues, then this would lead into an infinite regress of representations. For, if the representational processes that gives rise to the representation of, for instance, a tree were to be represented and as such experientially manifest as well, then this would require a further representational process. Now, if that latter representational process were to be again itself experientially manifest, then a further representational process would be required—et ad infinitum. The following longer quotation formulates that notion in an intrusive way:

It is important to note that self-modeling, in terms of its logical structure, is an infinite process: A system that would model itself as currently modeling itself would thereby start generating a chain of nested system-related mental content, an endless progression of ‘self-containing,’ of conscious self-modeling, which would quickly devour all its computational resources and paralyze it for all practical purposes.

...

One simple and efficient way to interrupt a circular structure is by introducing an untranscendable *object*. My hypothesis is that the phenomenon of transparent

---

<sup>92</sup> Metzinger 2003, 279.

<sup>93</sup> “The phenomenal self can now be regarded as a weapon, developed in a cognitive arms race. Conscious selves are like instruments or abstract organs, invented and constantly optimized by biological systems.” (Metzinger 2003, 273)

self-modeling developed as an evolutionary viable strategy because it constituted a reliable way of making system-related information available without entangling the system in endless internal loops of higher-order self-modeling.<sup>94</sup>

Metzinger suggests that the phenomenal self is an invention of the physical organism that serves to reduce the information to be processed by the organism. If the organism had to represent all representational processes, such process would consume a high amount of ‘computational resources’, which would be an evolutionary disadvantage. Put differently, on Metzinger’s view, the organism is evolutionarily better off creating a phenomenal self that misrepresents its ontological basis, the physical organism, than trying to represent itself in a more accurate and truthful way.

#### 1.4.5 All phenomenal content is hallucination and the virtual self is unreal

Provided Metzinger were right in conceiving of the self as a useful invention and product of the physical organism, would follow from that per se that the self cannot be real? The case of Bayne’s virtual phenomenalism, which bears high resemblance with Metzinger’s self-model theory of subjectivity, just seems to suggest that it does *not*. Like Metzinger, Bayne highlights that the phenomenal self is produced by the physical organism, notably, the brain:

In generating a virtual self the cognitive architecture underlying the stream of consciousness also ensures that the self is represented as the owner or bearer of those experiences that are responsible for its very existence. The intentional structure of the phenomenal field leads us to experience ourselves as entities that stand over and against our experiences—as inhabiting the ‘centre’ of a phenomenal field.<sup>95</sup>

However, very much unlike Metzinger, according to Bayne, the fact that the experience of self and selfhood is produced by the brain does not render the self unreal. Without naming Metzinger, Bayne seems to address Metzinger’s position when he writes:

Although selves are merely intentional there is nonetheless a sense in which self talk is perfectly legitimate. It needn’t be rejected (indeed, I doubt whether it could be rejected), nor need it be legitimized by finding something in consciousness or its underlying substrate that might qualify as its referent, for it needs no such referent.<sup>96</sup>

---

<sup>94</sup> Metzinger 2003, 338.

<sup>95</sup> Bayne 2010, 292.

<sup>96</sup> Bayne 2010, 293.

As this quotation shows, there is a further assumption operative in Metzinger's refusal of the self: the presupposition that if self or selfhood were to be real, then it would need to show up as an intentional object, given to consciousness as something that stems from within consciousness or as something to be found in the physical world. For Bayne, by contrast, it suffices that experiences are lined up around a virtual center from the perspective of which phenomenal content related to the world is given. According to him, it would be hopeless and naïve to try to find a more real 'self' (or better kind of self) than the one we are familiar with through our experiences:

The kinds of selves that we possess are as real as selves get. This kind of reality might not be enough for some, but I think it provides all the reality that we might have reasonably hoped for here. Perhaps more importantly, it provides all the reality that we *need*.<sup>97</sup>

Bayne seems to see no problem in accepting the specific reality of selfhood as something that presents the *mode* through which we experience the world—or a phenomenological component of the experience of the world—rather than an object to be found in the world. The contrast between Bayne and Metzinger in these points reveals Metzinger's presupposition that for something to be real it must exist as an object. Accordingly, it reveals that the main target of Metzinger's denial of self is the classical understanding of the self or soul as a substance-like object. Metzinger's rationale seems to be: Although we have to concede a “nonepistemic feeling”<sup>98</sup> of being someone, given that we neither in perception nor introspection experience an object 'self', we have to conclude that “there is no single entity in or outside the system that directly corresponds to the primitive, prereflexive feeling of conscious selfhood.”<sup>99</sup>

Accordingly, it might appear that all it takes to riposte Metzinger's attack on the self is to describe selfhood in different, non-objective terms, e.g., like Bayne as a “virtual centre of 'phenomenal gravity'”<sup>100</sup> from which and against the background of which experiences are given. Although discussing what exactly the constraints are for something to be real, and whether the reality of selfhood should and could possibly meet the same criteria as worldly objects, such as tables, mountains, and trees, is important, it will not be enough to reduce Metzinger's argument to the notion that only physical objects are real—although this indeed is an assumption that he makes.

Yet, crucially, the scope of his argument is much broader and even more radical. It spreads out to any understanding of selfhood as something pertaining to the phenomenal,

---

<sup>97</sup> Bayne 2010, 293.

<sup>98</sup> Metzinger 2003, 287.

<sup>99</sup> Metzinger 2003, 564f.

<sup>100</sup> Bayne 2010, 289.

i.e., experiential sphere, be that as an object for experience, a structural component of the phenomenology of experience, or the very mode of experiencing the world. His refusal of the reality of selfhood understood along such lines is more radical because what Metzinger highlights is that any phenomenal content as such is ultimately to be considered as hallucinatory content. Thus, it is not only the fact that our self-experience does not correspond to any object that could possibly be given in experience (as an object) that renders, in his view, self-experience hallucinatory. Self-experience is hallucinatory, because it is experience to begin with:

Phenomenal experience during the wake state is an *online* hallucination. This hallucination is *online* because the autonomous activity of the system is permanently being modulated by the information flow from the sensory organs; it is a hallucination because it depicts a possible reality as an actual reality. Phenomenal experience during the dream state, however, is just a complex *offline* hallucination.<sup>101</sup>

[A]ll phenomenal content is hallucinatory content, because what I have termed its *de nunc* character ... is nothing but a simulational fiction that proved to be functionally adequate. From an *epistemological* perspective it is not a form of knowledge.<sup>102</sup>

The point Metzinger attempts to make is that any phenomenal content, anything we experience, ought to be understood in terms of its phenomenological character which per se is to be judged as fictitious. For, anything that is given to us is given as now, in a temporally immediate way. Yet, as Metzinger demurs at taking that at face value, there is, in a strict ontological sense, no “now”:

The first point to note when shifting back into the third-person perspective is that the physical world is ‘nowless,’ as well as futureless and pastless. The conscious experience of time is a simulational type of mental content that proved to be a useful tool for a certain kind of biological organism on a certain planet.<sup>103</sup>

A physical theory about the universe will never tell us what time is ‘now.’ ‘Nowness’ is a kind of mental content that only appears under phenomenal representations of reality. As extensively discussed above, our experiential present always is a simulated present, because in its extended character (the phenomenal Now being a dilated or ‘smeared’ form of time as opposed to the time of physics) it is a pure fiction.<sup>104</sup>

---

<sup>101</sup> Metzinger 2003, 51.

<sup>102</sup> Metzinger 2003, 250f.

<sup>103</sup> Metzinger 2003, 127.

<sup>104</sup> Metzinger 2003, 250.

Although Metzinger sometimes seems only to maintain that all phenomenal content *could* turn out to be a hallucination,<sup>105</sup> these quotations demonstrate that no matter how one phenomenologically describes—what ultimately is merely phenomenal—selfhood, no possible description or conceptualization of self can lend support to the reality of self once selfhood is restricted to phenomenal selfhood and all phenomenal content is deemed hallucinatory. It is precisely this that presents the final coffin nail for any attempt to salvage the self by building on *self-experience*. The whole point of Metzinger’s maneuver is to argue that selfhood is something that pertains to phenomenal consciousness only, “it truly is phenomenal property in terms of being an appearance only”<sup>106</sup>, and that therefore, given the problematic ontological status of phenomenal consciousness, selfhood must be considered unreal. As a merely phenomenal property, just like phenomenal consciousness per se, selfhood can and must be explained in reductionistic terms. Over and above its status as a phenomenal explanandum it has no further relevance to the project of the explanation of consciousness and the mind, let alone of physical nature. The reductionistic formula can be summed as follows:

**Step 1: *Self-Representation and Experience-Production.***

The physical organism represents physical properties of itself and the world by generating “a phenomenal window of presence”<sup>107</sup> by which intentional content is “transformed into phenomenal content”<sup>108</sup>.

**Step 2: *Transparency.***

The representational processes underlying the generation of phenomenal content themselves are not represented.

**Step 3: *Immediacy.***

The lack of representing of the representational process renders the givenness of phenomenal content immediate givenness.

**Step 4: *Sense of self.***

Phenomenal immediacy gives rise to a sense of self: phenomenal selfhood.

---

<sup>105</sup> “All objects (or subjects) seen through this representational window could at any point turn out to be hallucinations. Therefore any claims to knowledge—that is, to an additional *epistemic* or intentional content going along with the phenomenal content in question—are in need of independent justification. Phenomenal experience is how the world *appears* to you and as such it is nothing more than that.” (Metzinger 2003, 586)

<sup>106</sup> Metzinger 2003, 563.

<sup>107</sup> Metzinger 2003, 98.

<sup>108</sup> Metzinger 2003, 30.



**Step 5: Reduction.**

Phenomenal selfhood is a by-product of the specific way in which the world is represented in the human case. The human-specific phenomenal representation of reality is only a phenomenal fiction and so is its concomitant phenomenal selfhood. Beyond hallucinatory self-experience, no sense or meaning of self remains.

## 1.5 Argumentum ad absurdum. Fending off the hallucination model

Metzinger's version of the no-self doctrine is not only a highly detailed theory of the self and its alleged non-existence. It is a provocative approach on many levels, having a bearing on many issues and scientific disciplines. And, as if his ideas on phenomenal consciousness, representation, selfhood, and many other important subjects of philosophy and related scientific disciplines were not irritant enough, Metzinger combines them with sometimes seemingly aggressive comments on the history of philosophy and other philosophical traditions that do not subscribe to naturalism and materialism like he does. To give a few examples:

[T]heoretical approaches to the mental, still intuitively rooted in folk psychology, have generated very little growth of knowledge in the last twenty-five centuries.<sup>109</sup>

One of the greatest dangers in philosophy of mind and in traditional theories of subjectivity in particular has been to underestimate the richness and complexity of the target phenomenon, for instance, by making absurd claims about necessary features a priori.<sup>110</sup>

And, of course, there is the general rejection of any “dangerous tendency toward arrogant armchair theorizing”, something he ascribes to what he calls “analytical scholasticism”<sup>111</sup>. But it should be clear that the scope of this verdict is broader. Not surprisingly, Metzinger's proposal has had made an impact and tickled out some not unequally toxic reactions in some of his readers. Graham Harman, for instance, cannot resist highlighting that it is “nihilistic younger males who enjoy breaking things into pieces”<sup>112</sup> to whom Metzinger's work seems so appealing. It is also Harman who places Metzinger “in the camp of what David Chalmers calls *Don't-have-a-clue-materialism*”<sup>113</sup>. He describes it in the following

---

<sup>109</sup> Metzinger 2003, 16.

<sup>110</sup> Metzinger 2003, 326.

<sup>111</sup> Metzinger 2003, 3.

<sup>112</sup> Harman 2011, 29.

<sup>113</sup> Harman 2011, 24.

words: “I don't have a clue about consciousness. It seems utterly mysterious to me. But it must be physical, as materialism must be true.”<sup>114</sup>

Perhaps this kind of reaction to Metzinger is not completely uncalled for. At least, if it is true what Harman further states: “Respect for his opponents is not among the chief virtues of Thomas Metzinger”<sup>115</sup>, pointing to an interview of the latter with the journal *Collapse*<sup>116</sup> in which Metzinger “bluntly dismisses a host of near and distant foes”<sup>117</sup>, comparing philosophers with the folk who are both equally guilty of “philosophical conservatism”<sup>118</sup> or accusing other neuroscientists, who are more prone to buy into the views of Metzinger, of “‘corniness’ of seeking ‘media attention’”<sup>119</sup>.

Although the polemic skirmish seems to play a not unessential role in the debate revolving around Metzinger’s theses, more philosophical responses to his self-model theory of subjectivity and its many different facets abound—and Harman himself provides valuable ones.<sup>120</sup> My discussion of these focusses on the assumptions that underlie the idea that our self-experience is hallucinatory.

### 1.5.1 Global phenomenal hallucination?

Clearly, the strongest claim Metzinger makes to undermine the self consists in denouncing all experience as mere fiction. It is not only strong because it is radical but also because it seems global in scope and requires a strong argument. Is Metzinger’s argument strong? Certainly, his highly-detailed neurophilosophical theory wrests deference from us and one shouldn't simply dismiss the idea that all what we experience is merely an appearance because it seems to run against our everyday intuition. This, after all, the thorn of skepticism, has been a pressing philosophical *topos* and motivation for philosophical thinking since its very beginning. Accordingly, as a well worked-out theory of the deceiving character of our experience, it deserves a critical assessment just like any other elaborate theory of knowledge. Hence, the question is whether his approach is capable of launching a theory of knowledge?

The motivated reader will find much in Metzinger’s book, rudiments of a theory of knowledge, however, are not among it. Rather, she will encounter a clear case of presupposed affirmation of naturalistic perspectives on empirical sciences and a disdain of

---

<sup>114</sup> Harman 2011, 24, Chalmers 1996, 162.

<sup>115</sup> Harman 2011, 7.

<sup>116</sup> Metzinger 2009b.

<sup>117</sup> Harman 2011, 7.

<sup>118</sup> Metzinger 2009b, 190.

<sup>119</sup> Harman 2011, 7.

<sup>120</sup> For instance, Beni 2019, Fuchs 2017a, Duus 2017, Harman 2011, Gallagher 2005b, Himma 2005, Rinofner 2005, Thompson 2009, Zahavi 2005.

a priori philosophy. It is the latter, the disdain of a priori reasoning, that seems to be the reason for Metzinger's neglect of discussing the implications of his very own philosophical approach for any theory of knowledge. Such neglect, however, is, as it turns out, highly problematic. For, it is precisely the implications of his hallucination model for the theory of knowledge that renders the very same model impossible. Let me explain why I believe this is the case.

To anticipate, Metzinger degrades all phenomenal consciousness, our experience, to mere hallucinations, on the one hand. On the other hand, he wants to stay clear from any a priori argumentation to the effect that mere thinking will not do for him to buttress a philosophical argument. But then one wants to retort: With the two classical roots of knowledge, experience and thinking, cut off, it remains obscure how his own approach ought to stand on secured epistemological grounds. The point is, I submit, it doesn't. Metzinger's trust in empirical, mostly neuroscientific studies betrays not only that his position has never transcended the belief in phenomenal consciousness as a source of knowledge but also that it relies heavily on assumptions that would require a discussion of a priori reasons. This together renders the whole self-model theory of subjectivity problematic, "paradoxical"<sup>121</sup> even.

Consider the role of phenomenal consciousness in Metzinger's theory. He has made it very clear that phenomenal consciousness must be seen as hallucinatory fiction, a simulation of the brain. But on the other hand, he highlights the many neuroscientific studies of brain processes that are to tell us how the representational processes that underlie our self-model need to be understood. That, however, comes with a bitter taste of contradiction. Harman correctly writes that "objects for Metzinger have exactly the same shadowy status as the self"<sup>122</sup>. If phenomenal consciousness ought to be intrinsically fallacious, by the mere fact of being a form of consciousness through which content is present in a virtual window of presence, then this spreads out to any object given to us in experience. The *empirical* studies Metzinger so highly appreciates, they all rely on the belief in what is given to us in experience. They precisely rest on the assumption that knowledge with regard to natural processes is only to be acquired by rigorous experience-based natural science. But if experience is deemed essentially and generally hallucinatory, one wonders, how can such a drastic denunciation of experience leave our understanding of empirical science unaffected, let alone, intact? Harman phrases his surprise in the following way:

If you insist like Metzinger that the book in your hands is merely a hallucination due to the autoepistemic closure of the nervous system, and then *also* claim that paying attention to the book is like a nonverbal arrow going outside and

---

<sup>121</sup> Harman 2011, 29.

<sup>122</sup> Harman 2011, 26.

pointing at the real book, then this is an almost comical analogue of eating and preserving the cake simultaneously.<sup>123</sup>

The problem is that once phenomenal consciousness, experience, is considered hallucinatory, anything that *seems* to be turning out to be true or false in experience is intrinsically fallacious at the same time. The problem is not that any perception could turn out to be wrong itself again. Nor is the experience that a prior perception was non-veridical problematic. The fact that our perception is vulnerable to error and often needs to be corrected is something intrinsic to the process of perception. Take the example of the mannequin.: I seem to be seeing a person at the horizon and after having come closer to the alleged person I ‘find out’ that it was a mannequin. The experience ‘it was a mannequin rather than a person’ is not infallible either. Approaching the now alleged mannequin, I may find out that it was my friend standing still like a mannequin to fool me. Possibility of falsification is an in-built feature of our perception.<sup>124</sup> But Metzinger’s hallucination theory claims more than that our perception is uncertain. His doctrine implies that our perception is *certainly wrong*. What use is it if in my hallucinated perceptual experience something turns out to be rather *p* than *q*? Whatever phenomenal experience makes me believe that rather *p* than *q*, it is only valid within the “virtual space”<sup>125</sup> of my experience. And as such it is intrinsically detached from reality: “[T]he content of their [of human beings’, P.S.] subjective experiences always is counterfactual content, because it rests on a temporal fiction”<sup>126</sup>.

But if that is so, all we can learn from empirical studies is similarly based on a fiction. The whole setting of these studies is built on our lifeworld, which is an experiential world. Even the belief that we possess brains and that brain mechanisms underlie our experiential life, even such belief is based on and dependent on our phenomenal consciousness. Fuchs describes this point in the following way:

Of course, even neuroscientists or neurophilosophers continue to live with this insight in the everyday world of ‘naïve realism.’ And they are well advised to do so; for if the world of our experience were in fact only a virtual product of our brains, how could we ever find out anything about the actual world ‘out there?’ If we were only in contact with a *reality simulation*, a world of mere appearances, how could neuroscientists make any statements about ‘real brains?’ Already in terms of knowledge theory, such a position is untenable.<sup>127</sup>

---

<sup>123</sup> Harman 2011, 32.

<sup>124</sup> Husserl 1966b (Hua 11), 25ff., Melle 1983, Reis Piedade 2002.

<sup>125</sup> Metzinger 2003, 25.

<sup>126</sup> Metzinger 2003, 57.

<sup>127</sup> Fuchs 2017a, 4.

The problem with declaring all phenomenal consciousness as hallucinatory content bereft of any epistemological value is that it leads to a peculiar form of epistemological nihilism: Although there must be something like reality, I cannot know anything about it; everything that appears to me is a mere hallucination that no experience could correct; I, myself, not even am real; only reality is real, producing the experience of a virtual, non-real phenomenal world that does not exist as it appears for equally non-existing selves that only appear to inhabit it. “The title *Being No One* could easily have been expanded to *Being Nothing*, covering objects of experience no less than the experiencing self”<sup>128</sup>, writes Harman. There is certainly something to that, but it doesn't get the picture right. It is true that, on account of the hallucination view, nothing we experience is real but only hallucinatory content. Yet, for Metzinger, this doesn't mean nothing exists.

The question, therefore, arises as to what precisely of our knowledge is it that survives the excruciating discovery that all that is phenomenally given to us, including ourselves, is merely hallucinatory?<sup>129</sup> To formulate it differently, and to commemorate Descartes: What is it that the *deus malignus* cannot deceive us about? The Cartesian answer is present to all of us. It is the doubting or thinking as such, i.e. that what for Descartes amounts to the same as the ego, that we cannot cast doubt on. This, of course, is unacceptable for Metzinger. For, it is his explicit goal to show that our “Cartesian intuitions ... are ultimately rooted in the deeper representational structure of our conscious minds”<sup>130</sup>. And, obviously, his book is devoted to unfolding that very thought. What survives the hyperbolic doubt for Metzinger, then, is the representational process itself—including what is doing the representing. In other words, what remains as existing in the picture of Metzinger is the *deus malignus* and his act of deception, as it were. What in Descartes was a thought experiment, a methodological, artificial, and, most importantly, temporary hypothesis to demonstrate the reality of self and what it consists of, in Metzinger becomes what is ought to be *really* happening. Admittedly, Metzinger does not understand the representational process as being done by an evil god. On his view, neither is it god who does the representing; nor is the deceiving an act of malignity, for no one gets deceived, as we, as experiencing subjects, don't exist. Rather, it ought to be “Mother Nature” who is undertaking some form of “self-conversation”<sup>131</sup>, as Metzinger circumscribes the picture. Metzinger commits to a “realist ontology”<sup>132</sup>, and his hallucination model, although it comes across as a strong form of skepticism, is, as he constantly highlights, throughout naturalistic.

---

<sup>128</sup> Harman 2011, 25.

<sup>129</sup> Cf. Zahavi 2005, 11.

<sup>130</sup> Metzinger 2003, 2.

<sup>131</sup> Metzinger 2003, 634.

<sup>132</sup> Metzinger 2003, 462.

Yet, the specific version of his naturalism undermines—*ex hypothesi*—any empirical support for the very same version. If it were true, no empirical science could prove it. The myriads of studies Metzinger consults and examines in an otherwise very interesting way become worthless. If Metzinger’s version were true, the only thing that could be said would be this: There is the experience of a phenomenal world including a phenomenal subject, produced by nature as the only reality which is the bearer of the representational processes through which the experience of a phenomenal world including a phenomenal subject emerges. But the problem is how can Metzinger then show that his specific version of naturalism is true in the first place?

That he would need to show that his version of naturalism is true independently of the support of empirical studies is something Metzinger seems to acknowledge:

Trivially, if an internal representation of the system itself exists, according to the fundamental assumptions of any naturalist theory of mind there also has to exist a physical system which has generated it. I call this the ‘naturalist variant of the Cartesian *cogito*.’

...

[E]ven an only weakly self-conscious machine would be justified in assuming that it possesses *some* kind of hardware. All the details of its current conscious self-representation may be false, but the underlying existence assumption is always justified. Of course, naturalism itself would have to be argued for by this machine, on independent grounds.<sup>133</sup>

The crucial point lies in the last sentence. The naturalist variant of the Cartesian cogito is reliant on independent reasoning in favor of naturalism. But what exactly are those “independent grounds”? It seems there cannot be such grounds altogether. For, given the terms of the hallucination model, any reference to experience-based sources of knowledge will not do. Moreover, the representational processes are not something we could possibly be aware of from a first-personal perspective. But even if they were, as experiences they would *ex hypothesi* lack the required validity to support the naturalism the hallucination model targets. But even if one were willing to grant that empirical studies despite being based on experiences may tell us something about the nature of representational processes and the dependence of experience on brain mechanisms, “it is not entirely made clear how the ‘empirical’ data Metzinger offers either falsifies all competing theories of mind, or verifies his specific conclusions about consciousness and the self”<sup>134</sup>. Although Metzinger makes a fascinating job in making it look like it, but the many studies he refers to are far from pointing unambiguously to his interpretation of self-nihilism. Crucially, it is even highly doubtful that such strong metaphysical conclusions could possibly be derived from

---

<sup>133</sup> Metzinger 2003, 279.

<sup>134</sup> Thompson 2005, 3.

empirical data directly.<sup>135</sup> Yet, neither could any rationalistic approach—a certain interpretation of the pure ‘I think’ or thinking—yield an argument for the self-nihilism Metzinger propagates. Quite the contrary, such rationalistic inquiries traditionally rather paved the path that lead into the assumption of the reality of self and examinations of its nature. Moreover, it is also Metzinger who precisely declines to make use of a priori arguments. Hence, affirmatively and explicitly building on some a priori argument that is built upon the mere activity of thinking is not an option for the hallucination model.

That said, and undermining Metzinger’s stance, clearly his own approach is not bereft of a priori assumptions. As Harman, correctly in my view, highlights, “even if the self is causally generated by the physical-organism-as-a-whole, this does not mean that it is *nothing but* a group of disconnected nerves and cells.”<sup>136</sup> The fact that self-experience or experience more generally might be bound to brain activity as such does neither render phenomenal selfhood nor experience in general hallucinatory nor does it entail that the self does not exist. To claim so, rests “on a simple *a priori* dogma that if something has causal antecedents, then only those antecedents can have independent reality”<sup>137</sup>. Similarly, the related idea that only mind-independent objects and processes can be said to be real is something that Metzinger does not argue in favor of, but rather takes to be the guiding interpretative framework and background for interpreting the empirical data.

In conclusion, it must be stated that Metzinger’s hallucination model faces severe problems that render its position highly unattractive. First, the model problematically undermines and devaluates the two traditional sources of knowledge—experience and thinking—without providing alternative ways to conceive of knowledge and its acquisition. Second, although the hallucination model dismisses both experience and thinking as sources of knowledge, it is built on assumptions that are reliant on the validity of at least some experiences and a priori reasons. Hence, the hallucination model is contradictory, because it rests on a fundament that, on its own terms, ought to lack any validity and falls generally short in lending support to any possible justification.

### 1.5.2 Partial phenomenal hallucination? The incoherence of self-hallucination

Some might acknowledge the criticism that I just presented, but insist that the hallucination model of Metzinger’s kind might still tell us something interesting about how selfhood emerges out of physical processes, and that one might *mutatis mutandis* construe varied forms of arguments against the reality of self that avoid the critical points against the

---

<sup>135</sup> Zahavi 2005, 5.

<sup>136</sup> Harman 2011, 29.

<sup>137</sup> Harman 2011, 29.

hallucination model mentioned earlier. Let me survey and discuss the most important of such possible variations of the hallucination model.

**Option 1:** Phenomenal consciousness is fictitious in the sense that it requires further epistemic justification, which is in general possible.

As I have mentioned earlier, Metzinger's work is sometimes ambivalent and occasionally seems to suggest that the only problem with phenomenal consciousness is that the contents that are given in it *could* turn out to be wrong: Seeing the table may be a hallucination, but it need not either.<sup>138</sup> The idea then would be that phenomenal consciousness simply lacks the feature of functioning as a verification principle rather than saying that phenomenal consciousness per se is fallacious. Accordingly, self-experience as such does not warrant the reality of self. Further epistemic justification would be needed. Of course, from the fact that further epistemic justification is needed alone it is impossible to infer that self does not exist. It just means that self-experience alone is not sufficient to justify the assumption of the reality of self and that further proof is pending. Hence, the modifications in Option 1 wouldn't suffice to make the hallucination model more convincing.

**Option 2:** Phenomenal consciousness is fictitious in the sense that it requires further epistemic justification, which is in general possible. Yet, only in cases of some particular phenomenal content it is impossible.

One could further modify Option 1 and claim that in some cases phenomenal content is indeed hallucinatory, namely when it—unlike in fantasy and imagination—does suggest that it refers to something while in fact it couldn't possibly. One could then argue with Metzinger that this precisely is the case with phenomenal selfhood: We have self-experience but it is generally impossible that it turns out as right, for we could not possibly gather any further epistemic reasons to lend support to the idea of self other than mere experiences of being someone. On that view, because we will always *only* have self-*experience* and experience is generally in lack of further epistemic proof, we must conclude selfhood is hallucinatory phenomenal content. Now, such a route leading to the denial of self could only possibly make sense if in other cases of phenomenal content, say the experience of the table in front of me, it would be possible to find experience-independent epistemic proof. But here the problem arises again: If phenomenal consciousness and its content is deemed in need of further epistemic justification, it remains unclear what exactly could provide such epistemic proof. What—other than the experience of table or related

---

<sup>138</sup> See, for instance, Metzinger 2003, 57, 586.



experiences—could possibly give us the epistemic justification to believe that there is in fact a real table? It is hard for the defender of such a version of the hallucination model to demonstrate that some phenomenal content (the table) can be epistemically justified in an independent, i.e. experience-independent, way, while other phenomenal content (phenomenal selfhood) cannot. For, it is unclear how such an experience-independent epistemic justification of the belief in the reality of what is concretely phenomenally given should run altogether. But, as long as it remains obscure how any experience-independent epistemic justification could be provided, it doesn't make sense to say in some cases of phenomenal content it is possible, while in others it is not. Option 2 doesn't help either.

**Option 3:** Certain types of phenomenal consciousness or content are fictitious. Whether a given type of phenomenal consciousness or content is fictitious depends on whether the object it refers to is a physical object, something pertaining to the physical world.

Another alternative of formulating the hallucination model would consist in the idea that not all phenomenal content is fictitious in the sense that it requires further independent epistemic justification. One might claim that those things that appear to us as physical objects such as the table are not in need of independent epistemic justification. Only types of experiences such as self-experiences are in need of further epistemic justification, because what they are an experience of is nothing we encounter *in* the world like the table, but rather something that is intimate, private and only experiencable of the self in question: we might experience ourselves, but we do not see, hear, touch, smell nor taste ourselves—at least not as *experiencing* subjects in that we do not see our seeing, hear not our hearing and so on. The feeling of being someone might arise when we see, hear, touch, smell or taste the world and its objects, but it is not a feeling of something that could itself be seen, heard, touched, smelled or tasted. Our self-experience, our sense of self, thus, is hallucinatory, because it seems to refer to something that cannot possibly be given in *perceptual* experience, i.e., perception as we usually understand it when saying 'I see a table', 'I heard you screaming' and the like. Because we experience something that cannot be given through our senses in perception altogether, we must conceive of it as a hallucination. Now, formulating the hallucination model along such lines seems to fit better with the idea that our empirical sciences may indeed produce knowledge as long as they painstakingly focus on what is given to us through our senses (or through what is physically measurable by reliable and valid instruments). Unfortunately, if selfhood is not among that which can be perceived in a sensory way nor physically measured otherwise, then our best scientific explanations of the world seem to imply that we must conclude the self is only a hallucination.

Yet, there are several issues by which such a version of the hallucination model is troubled. Accordingly, Option 3 neither seems to render the hallucination model a viable position. Let me focus on three main points that, in my view, undermine Option 3:

*1) Physicalism needs further justification.*

It clearly relies on the assumption of physicalism according to which only that what is physically measurable is real, which is an assumption that is in need of further justification, usually lacking in accounts that are based on it. What does it mean that further justification is needed? Earlier, I said that the fact that we have self-experience might be taken that it does not warrant the reality of self but that, from that, it doesn't follow that the self is unreal. Now, to say that physicalism requires further justification similarly doesn't necessarily mean that physicalism is false. Yet, I wish to keep those two cases well apart from each other. For, in the case of self-experience, we *do* have a motivation for believing in the reality of self. Our self-experience, our feeling of being someone, as Metzinger has it, does lend support to the belief in the reality of self. The burden of proof to show that the self is unreal, therefore, is on the side of the one who denies the reality of self, whereas it is perfectly reasonable to believe in the reality of self when feeling oneself being someone. But the case of lacking epistemic justification for physicalism is different. We do *not* have an experiential motivation for physicalism as we have an experiential motivation for the belief in the reality of self. Accordingly, in the case of physicalism, a lack of further epistemic justification renders physicalism a merely theoretically possible position. No less, but not more. Hence, Option 3 still doesn't avoid the problems Metzinger's hallucination model is facing, as it is unclear what possible justification could be given. For, it still remains unclear why we should only accept physical objects to be real.

*2) Self-experience is a non-sensory phenomenon and the notion of hallucination ill-chosen.*

If self-experience is a type of experience to be essentially distinguished from sensory perception, then it is problematic to make use of the concept of 'hallucination'. Hallucinations, as usually conceived, are non-veridical perceptions<sup>139</sup>, something that could at least possibly turn out as true perception. Hallucinations are aberrations of perceptual experience and refer to physical phenomena: *hearing* voices that are really not there although experience suggests there are; *seeing* things that are not really there although experience suggests there are etc. But self-experience, if considered as distinct from sensory perception, does not suggest there is something that could be perceived when really there is nothing to be perceived. Self-experience, the experience of being someone who is

---

<sup>139</sup> In the literature, hallucinations are usually taken as perceptual states that are, unlike proper perceptions, not evoked by external stimulation of the corresponding sensory organs. See, for instance, Frith 1992, 68, Halligan & Marshall 1996, 242, Hayward, Berry & Ashton 2011, 1314, Ratcliffe 2017, 6.

undergoing experiences, as the argument itself propagates, is precisely different to sensory perception. But, if that is so, then it is problematic to suggest the same interpretation criteria as in the case of sensory perception. A perception-like experience, when wrongly suggesting there is something perceivable, may be called a hallucination, but why should an experience like self-experience that doesn't even refer to a sensory phenomenon be judged a hallucination? It is paradoxical to emphasize self-experience is different from perceptual experience and then treat it like a non-veridical perception. There are not two kinds of feelings of being someone: one veridical, the other non-veridical. If Metzinger is right and selfhood is a merely phenomenal property, then there simply are no criteria external to experience that would need to be fulfilled to speak of real selfhood. Seeing the table has also external criteria to meet in order to be a veridical perception, criteria that transcend my perception of the table: There must be a table perceivable also for others. But the whole point of Metzinger's restriction of selfhood to *phenomenal* selfhood was that there are precisely no such external criteria in the case of selfhood. Selfhood completely pertains to the experiential realm and could not possibly be given in a way akin to physical objects. But why then assess its ontological status in the way we evaluate claims about the reality of physical entities?<sup>140</sup> Clearly, the only motivation for doing so is that one has already presupposed all along that only physical entities, i.e., anything measurable by the science of physical nature, are real. But again, that is a strong metaphysical claim that would require further arguments and justification. And it is hard to see how the hallucination model of self could provide such in a non-question-begging way. In any case, applying a notion that intrinsically refers to sensory phenomena such as 'hallucination' to self-experience while emphasizing that self-experience is not a sensory phenomenon is a non-starter.

### 3) *Moderate physicalism doesn't necessarily entail Metzinger's self-nihilism.*

If one were to buy into the physicalist assumption that only what is part of the physical world is real, it is unclear how explaining self-experience in terms of brain mechanisms would render selfhood unreal. If it were true that we could give a perfect description of how the feeling of being someone emerges out of brain mechanisms, i.e., physical processes, we are still not in need to infer from such fact that selfhood is unreal. For, one could equally argue that the fact that phenomenal selfhood is a product of physical processes is precisely what makes it an aspect of the physical organism. In other words, if phenomenal selfhood ought to be a product of physical processes, and all real there is ought to be physical, why should we deem phenomenal selfhood that emerges out of physical

---

<sup>140</sup> For a similar line of thinking, cf. Bayne 2010, 293.

processes unreal?<sup>141</sup> Why not declare phenomenal selfhood as a feature of certain physical objects, for instance, certain organisms and creatures that possess consciousness or self-consciousness?<sup>142</sup> Or, why not say that for phenomenal selfhood to be possible, information-processing physical structures such as brains are required? I do not wish to make a case here for any possible interpretation of the fact that phenomenal selfhood seems to have some physical underpinning, but it needs to be stated that even under the premise of physicalism there are several alternative interpretations possible that are compatible with the view that selfhood is real and the arguments delivered by Metzinger fail to rule them out.<sup>143</sup> Option 3 may make the hallucination model seem less radical, as it interprets the physicalism in a way that allows for some—*physical*—things to be real. But once acknowledged the reality of physical objects that can be given through our senses, a further explanation for why we should deny the reality of self is needed and it seems one can be a physicalist without thereby denying the self in a way Metzinger does. Option 3 would only be an interesting modification of Metzinger’s proposal if it retained the denial of the reality of self, but it doesn’t. The fact that it doesn’t further illustrates that Metzinger’s self-nihilism is not intrinsic to physicalist perspectives.

## 1.6 Conclusion: The hallucination model of self is inapt

The first kind of skepticism about self-experience that I presented consisted in the idea that we might admittedly have self-experience in terms of a feeling of being someone, of being the subject undergoing experiences, but that such self-experience is deceptive in that it suggests the existence of a subject that really does not exist. Accordingly, proponents of such a view like Metzinger’s have characterized self-experience as hallucinatory. Is such a line of thinking convincing?

In this chapter, I have presented the core arguments on which the notion of hallucinatory self-experience rests and discussed several variants thereof. First, I focused on Metzinger’s strong version of the hallucination model of self. Its main argument can be summed up in the following way:

---

<sup>141</sup> Gallagher, in a commentary on Metzinger, highlights that the fact that it is the real body that generates self-experiences undermines the idea of self-experience as hallucinatory: “I am not a fiction or ghost generated by the body. What I pre-reflectively experience of this existence is not an illusion, since my pre-reflective experience is itself generated by my real embodiment—by a brain that is part of a real bodily system, which is part of a real environmental system. . . . That there are real bodies means that there are real selves who experience themselves both as bodies and as more than bodies.” (Gallagher 2005, 8)

<sup>142</sup> See, for instance, Panksepp 1998, Ghin 2005, Hobson 2005.

<sup>143</sup> That is precisely the point of McClelland 2019.

- (1) We have self-experience: phenomenal selfhood.
- (2) We *only* have self-experience: There is no further reason to assume the reality of self.
- (3) *All experience*, all phenomenal consciousness and its contents are merely a *hallucination*, a virtual and essentially fictitious depiction of reality, a simulation.
- (4) Since selfhood is restricted to phenomenal selfhood and phenomenality per se is hallucinatory, we must conclude that selfhood is hallucinatory and there really is no one.

As this summary shows, the key to the denial of self lies in the reduction of selfhood to *phenomenal* selfhood and the degradation of phenomenality as such. Yet, such account, I argued, faces severe problems. For, denying any epistemological value of experience ultimately lets any plausibility of Metzinger's specific representational narrative that is to buttress his no-self-doctrine dissolve. We wouldn't even be able to claim there are brains that simulate a world and subjects in it in the first place, all empirical knowledge would crumble away and with it the representational narrative, too. Moreover, it turned out that the mere fact of restricting 'selfhood' to 'phenomenal selfhood' does not render selfhood unreal per se.<sup>144</sup> To draw such a conclusion, it would be necessary to consider all phenomenality, all experience, as fallacious, which amounts to an incoherent, because self-defeating, epistemological standpoint.

I then considered modified versions of the hallucination model of self, particularly one that seems to be able to refrain from questioning all experience, while devaluing only experiences of a certain kind: experiences that do not refer to something perceivable through our senses, i.e., experiences that do not refer to a physical object. As it turned out, such maneuver is based on the general assumption that physicalism is true, i.e., the view that only physical objects and processes are real. However, it remains unclear whether the hallucination model of self has the means to launch the thesis of physicalism at all. Furthermore, even if physicalism were true, it is an open question, to say the least, whether the specific interpretation of it implied in the hallucination model of self is the best interpretation of physicalism or that physicalism entails the denial of the reality of self. For, there are not only other interpretations of physicalism possible that do not render selfhood unreal. The interpretation of physicalism implied in the hallucination model of self also seems to rely on what is a category error: to apply norms of reference akin to perceptual experience to self-experience. Since self-experience, at least according to the hallucination model, does not intend a physical object nor corresponds to any physical object, the

---

<sup>144</sup> I have already mentioned one alternative: Bayne's virtual phenomenalism. In Part II, I will discuss more alternatives.

argument goes, it must be considered hallucinatory. But the crucial difference to a hallucination is that, unlike self-experience, a hallucination is a perception-like experience that propagates to present a perceptual object. In the case of a perceptual experience that presents a physical object although the object does not exist, it might make sense to speak of a hallucinatory experience. But in the case of self-experience there is no mismatch between appearance and reality if there is no physical object that corresponds to self-experience. Self-experience, at least as it is described in the hallucination model, does not purport to refer to a physical object. Accordingly, it is not even clear whether a distinction between appearance and reality applies in the case of self-experience.<sup>145</sup>

To address that issue, however, will require a more detailed analysis of the phenomenology of self-experience and of what exactly is meant by the feeling of being someone—or the feeling of being oneself as I will later describe it. That will be my aim in Part III. In this first chapter, my aim was mostly negative: to show that casting doubt on self-experience while having it comes at high conceptual costs. Whether it is still worth to adopt the hallucination model of self heavily depends on alternative ways to treat the issue of self-experience. Before yielding a final verdict on the hallucination model of self, therefore, a discussion of the alternatives is required. As I will later argue, there *are* alternative ways to respond to the Cartesian proposal that are more convincing than the hallucination model.

---

<sup>145</sup> Cf. Bayne 2010, 293, Zahavi 2014a, 43.

## Chapter 2

# What Self-Experience? No Self, Hume, and the Grinning of the Sceptics

Against those who would maintain that ... an empirical unity of consciousness is not sufficient for philosophical purposes, I would urge that what is good enough as a correct statement of facts as they actually are, ought to be good enough for philosophy. Of course, philosophy must not stop with this single fact, but must bring it into intelligible relation with other facts which this statement does not include. When this is done, I venture to think that *the continuity of consciousness will prove to be near the center of gravity of the philosophical system.*<sup>146</sup>

In the previous chapter, I discussed a first form of skepticism about self-experience. It consisted in the notion that self-experience, although an admitted phenomenological ingredient of our experience of the world, is hallucinatory and so doesn't warrant the reality of self. In this chapter, I turn to the second skeptical view about self-experience, which differs significantly from the first: It doesn't accept the view that there is self-experience to begin with. On that view, our experience of the world is bereft of any sense of self, for which reason proponents of such a view deny that there are any experiential grounds for any proposition about selfhood.

Clearly, the most prominent example associated with this line of thinking is Hume's empiricist treatment of self and his introspective analysis of our experience.<sup>147</sup> Hume's main point is that when we reflect on our experience we are unable to witness any positive perceptual phenomenon that could be able to make evident or at least indicate the existence of what in his times had been referred to as the 'soul'. And, he drew from his findings that we are, at least on empirical grounds, not warranted to make any claims about an alleged substantial self. Hume being an empiricist, moreover, believed there is no other fundament than our perceptual experience to build knowledge upon, wherefore he predicted little hope to find any secure clue that could lend support to the belief in the reality of self (as traditionally conceived).

Despite the major influence of such an approach, other prominent figures in the history of philosophy have invested much effort to show that we cannot dismiss selfhood

---

<sup>146</sup> McGilvary 1907, 231 (my emphasis).

<sup>147</sup> As I will show in the chapter, it is not fully clear whether Hume was indeed a proponent of the so-called no-self doctrine, if by the latter it is meant that selfhood is not real. However, he is typically associated with such a view. See, for instance, Albahari 2011, 90, Dreyfus 2011, 115, Foster 1991, Giles 1993, among others.

on the mere ground that it is not an experiential phenomenon in the same way as perceptual objects present themselves to us in experience. One important tradition along such lines runs from Kant via Fichte to Husserl. Building on Husserl and the phenomenological tradition, Dan Zahavi, more recently and quite prominently, has promoted the view that our experience of the world involves intrinsically a sense of self.<sup>148</sup> Most importantly, he suggests that the sense of self should not be understood as if it presented a specific conscious content but rather must be taken as the *form* in which any conscious content may be present to us.<sup>149</sup> On that view, experience is essentially imbued with self-experience. The phenomenological argument to countenance Humean introspective worries about self-experience consists in the notion that for something to be experientially present, i.e., given as now, it is required that the flow of experience as such manifests itself experientially too. The phenomenological argument that Zahavi extracts from the Husserlian account of inner time-consciousness, thus, suggests it is only against the background of the temporally structured stream of consciousness that something can be present in consciousness to begin with. Accordingly, one might formulate the phenomenological response voiced by Zahavi to Humean skepticism about self-experience in the following manner: While Hume may have correctly seen that we cannot describe our sense of self as a certain content passing through our conscious life, unfortunately, he has overseen and failed to consider the *way* contents pass through our conscious life. Although Hume was right in directing his attention towards the phenomenology of experience, he fell short in providing an accurate description thereof, the criticism runs. On account of the phenomenological tradition, if Hume had offered an adequate phenomenological description of our experience he would have found that the very streaming of experience is always experientially co-present together with what is given in experience. Such a line of argument has encountered broad reception.

Yet, in the past years, some have attacked the view that our stream of consciousness and its self-manifestation may give rise to a sense of self. Supporting the no-self doctrine, proponents of such a line of thinking have emphasized that talk about the stream of consciousness as such is phenomenologically unwarranted, and that therefore arguments in favor of some alleged self-experience based on the flow of experience rest on an experiential illusion of sorts. The reasons why some believe there is no stream of consciousness, at least not in the way it is usually conceived of, vary in the individual cases.

The aim of this chapter is to discuss these modern variants of Humean skepticism about self-experience—together with the traditional version. First, I will sketch the Humean, empiricist case against self-experience and examine whether the denial of self-experience necessarily entails the no-self doctrine. Second, I carve out which assumptions

---

<sup>148</sup> Zahavi 1999, 2005, 2014.

<sup>149</sup> Zahavi 2014, 22f., Zahavi & Kriegel 2016, 49.



underlie Hume's argument and reconstruct recent considerations that question these assumptions. Third, I present Zahavi's account of minimal self and his phenomenological descriptions of the structure of the stream of consciousness as an attempt to overcome the skepticism about self-experience of the Humean kind. Fourth, I then go on discussing recent no-self-views that launch their denial of self through a critical examination of the phenomenological structure of the stream of consciousness.

## 2.1 The Humean challenge: Is self-experience possible?

Hume's famous treatment of personal identity begins with a brief sketch of the kind of hypothesis about self-experience he aims to shed light on: the idea that "we are every moment intimately conscious of what we call our SELF; that we feel its existence and its continuance in existence; and are certain, beyond the evidence of a demonstration, both of its perfect identity and simplicity"<sup>150</sup>. It is that idea which he examines from his empiricist perspective. Hume's empiricist investigation of selfhood involves two steps:

- 1) Analysis of whether self-experience, given the way self is commonly understood, could be possible at all. Such an analysis consists in the theoretical examination of the compatibility of the notion of self (as substance) with Hume's empiricist framework that makes assumptions about the way experience is structured and how we gain knowledge through experience.
- 2) Second, an introspective analysis of our experience studying whether our experience features a sense of self in a way that could possibly give rise to the belief in a substantial self.

In both cases, Hume makes it very explicit: Neither does he hold that self-experience is *possible*, given the empiricist framework he is convinced of and the notion of self at stake; nor does he believe that self-experience *in fact* features in our experience of the world. Let me reconstruct both these analyses in more detail, before assessing whether the argument they give rise to is convincing.

### 1) Given the concept 'self' of the soul-theory, is self-experience even possible?

First, it is important to note that, discussing the possibility of self-experience, Hume presupposes both a certain notion of 'self' and of 'experience'. As the brief quotation above demonstrates, Hume is primarily concerned with the self understood in terms of classic

---

<sup>150</sup> Hume 2007/1739, 164 (1.4.6.1).

soul-theory. Soul-theory conceives of the self as substantial, single, and persisting, the metaphysical bearer of our experiences. Moreover, Hume promotes a certain understanding of ‘experience’: According to Hume’s empiricist doctrine, all occurrences of our mind, i.e., of our stream of consciousness, are either “impressions”<sup>151</sup> or “ideas”<sup>152</sup>, whereby impressions are considered primary compared to ideas. Ideas, if they are to be judged as real, need to be grounded on some impression. For, on Hume’s view, “[i]t must be some one impression, that gives rise to every real idea”<sup>153</sup>. So, asking the question whether self-experience is possible, we must find out whether we have an idea of self that is grounded on impressions. It is here where Hume formulates severe worries about the coherence of the view that we possess a sense of self—‘self’ as the soul-theory has it. The problem Hume sees is that no impression could possibly play the role of grounding the idea of self as an abiding substance. The self as conceived in the soul-theory is precisely not a specific impression among others but rather something “that to which our several impressions and ideas are suppos’d have a reference”<sup>154</sup>. Accordingly, Hume argues, we do not even have a “clear and intelligible”<sup>155</sup> idea of self in said manner. Since the self in question ought to be the bearer of all experiences—of all impressions and ideas—the self cannot be said to be an impression in the first place and is consequently unable to provide the experiential ground for any real idea of self.

But even if one were to grant that there could be an impression giving rise to the idea of self, Hume continues, it would have to possess a character that makes its existence unlikely: “that impression must continue invariably the same, thro’ the whole course of our lives”<sup>156</sup>. But our impressions are changing all the time and there seems not to be one single impression that persists and accompanies all other impressions that follow each other. If there were such impression that never changes and is prevalent in every moment of our stream of consciousness, then it would be a salient impression and we would know about it. But we don’t. Hume concludes that from the impressions that pass our mind we may not possibly yield the idea of self as the soul-theory has it: “It cannot, therefore, be from any of these impressions, or from any other, that the idea of self is deriv’d; and consequently there is no such idea.”<sup>157</sup> The way Hume and his empiricist doctrine conceives of the

---

<sup>151</sup> “Those perceptions, which enter with most force and violence, we may name *impressions*, and under this name I comprehend all our sensations, passions and emotions, as they make their first appearance in the soul.” (Hume 2007/1739, 1 (1.1.1.1/1))

<sup>152</sup> “By *ideas* I mean the faint images of these in thinking and reasoning, such as, for instance, are all the perceptions excited by the present discourse, excepting only, those which arise from the sight and touch, and excepting the immediate pleasure or uneasiness it may occasion.” (Hume 2007/1739, 1 (1.1.1.1))

<sup>153</sup> Hume 2007/1739, 164 (1.4.6.2).

<sup>154</sup> Hume 2007/1739, 164 (1.4.6.2).

<sup>155</sup> Hume 2007/1739, 164 (1.4.6.2).

<sup>156</sup> Hume 2007/1739, 164 (1.4.6.2).

<sup>157</sup> Hume 2007/1739, 164 (1.4.6.2).

content of experience—impressions and ideas—is, thus, not compatible with the idea of self that the soul-theory promotes. Accordingly, Hume’s verdict on the *possibility* of self-experience is negative.

2) *Do we have any sense of self that could give rise to the idea of the self as substance?*

Second, to further buttress that verdict, Hume also provides an introspective analysis of experience, which aims at showing that we *really* do not encounter one bit or component in our experience that would give rise to the idea of self:

For my part, when I enter most intimately into what I call *myself*, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never can catch *myself* at any time without a perception, and never can observe any thing but the perception.<sup>158</sup>

According to Hume, introspection shows that his empiricist verdict regarding the self is true. Whenever we reflect upon our experience, we do not spot anything else than our perceptions. Moreover, he emphasizes there is no other way to direct our attention to ourselves unless by turning our introspective gaze towards our perceptions. Our experience is exhausted by perceptions, for which reason Hume maintains that we “are nothing but a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement.”<sup>159</sup>

## 2.2 Strawson and the limits of Humean skepticism

Hume explicitly denies that there is any proof for a sense of self, that self-experience is incompatible with what he takes to be the right understanding of our experience, and holds that our experience, accordingly, as a matter of fact, does not feature anything that could count as self-experience. What is the scope of such a line of reasoning and is it convincing? To assess the Humean line of thinking and its force, I address four interrelated questions. Answering these questions will reveal that, despite its alleged radical nature, Hume’s verdict is best understood as quite restricted in scope:

---

<sup>158</sup> Hume 2007/1739, 165 (1.4.6.3).

<sup>159</sup> Hume 2007 /1739, 165 (1.4.6.4).

- (1) Does Hume’s skepticism entail the no-self doctrine?
- (2) Does Hume’s argument rule out all alternative notions of self or is another form of self-experience (provided a different understanding of selfhood) compatible with Hume’s argument?
- (3) Does Hume’s empiricist conception of experience provide the right background to understand self-experience which might turn out to be an experience of genuine kind?
- (4) Finally, in which sense do the specific notions of ‘self’ and ‘experience’ operative in Hume’s argument constrain his introspective analysis and is it phenomenologically accurate at all?

Let me start by addressing the question whether Hume’s argument amounts to the no-self doctrine. Clearly, Hume is one of the typical figures in Western philosophy—if not the most prominent one—to be associated with the no-self view. And certainly, if one wishes to make a case against the reality of self, denying self-experience seems to be a good point of departure, given that the notion of a sense of self is usually taken to be the primary reason to form the belief about the reality of self. Furthermore, *prima facie*, it seems out of doubt that Hume explicitly proposes and defends the no-self doctrine when he exclaims that “there is no such idea [of self, P.S.]”<sup>160</sup> or that he is “certain there is no such principle in me” that would give rise to the perception of “something simple and continu’d”<sup>161</sup> and that he could call himself. Moreover, saying that “the rest of mankind ... are nothing but a bundle or collection of different perception”<sup>162</sup> seems to match best with a no-self view. A widespread, perhaps the mainstream interpretation of Hume, thus, quite understandable, attributes to him the view that we are only—or, better, the mind is—a bundle of perception, something that does not allow for the notion of self or soul to be applied, a notion of which we do not even have a clear and intelligible idea.<sup>163</sup>

Yet, not all agree on the mainstream interpretation of Hume’s argument. Take, for instance, Strawson’s discussion of Hume in chapter 13 of his *The Subject of Experience* (2017). In this chapter, Strawson makes three fairly controversial claims regarding Hume’s treatment of self:

---

<sup>160</sup> Hume 2007/1739, 165 (1.4.6.2).

<sup>161</sup> Hume 2007/1739, 165 (1.4.6.3).

<sup>162</sup> Hume 2007/1739, 165 (1.4.6.4).

<sup>163</sup> However, that doesn’t necessarily mean that Hume could not defend a sort of “soft self-realism”, as, for instance, Kristjánsson (2010, 46ff.), emphasizes. On this view, Hume rejects the self understood as a thinking soul-substance, but not the moral self we are concerned about in our everyday life.

- (1) “that Hume doesn't think that the mind is just a ‘bundle’ of perceptions”;
- (2) “that the bundle account of the mind that he expounds in 1.4.6 [in his *Treatise*, P.S.] doesn't involve any sort of denial of the existence of subject of experience”
- (3) “that he never claims that the subject of experience isn't encountered in experience”<sup>164</sup>.

Put differently, Strawson holds that Hume neither denied the reality of selfhood nor that we have self-experience. Unfolding his argument is highly informative for understanding better the scope of Hume's skepticism about self-experience. First of all, Strawson emphasizes that there are two possible readings of Hume's argument: one *epistemological*, the other *ontological*. According to the latter, which seems to fit the mainstream interpretation, Hume defined the being or nature of the mind as that of a mere bundle or collection of perceptions. On the ontological view, thus, we *are* bundles of perceptions. By contrast, on the epistemological view, Hume's verdict is only concerned with what we can possibly know about the nature of the mind and of ourselves. Hence, Hume's reference to bundles of perceptions is only meant to clarify what kind of epistemological resources we possess to build judgments about the self. On that reading, all we got to determine the nature of self are bundles of perception. We simply cannot say more than that there are perceptions—impressions and ideas, more specifically—and that we cannot detect anything else when it comes to ourselves. Strawson admits that in section 1.4.6 of the *Treatise* what we find are rather “epistemologically unqualified ontological formulations of the bundle theory of mind”<sup>165</sup>. However, he holds against this that in other sections of the *Treatise*, notably in the Abstract and the Appendix, we find “explicitly epistemologically qualified statements” which are “claims to the effect that this is all we can know or clearly conceive of the mind”<sup>166</sup>:

The mind ‘*as far as we can conceive it*, is nothing but a system or train of different perceptions’ (Abs§28/657, my emphasis). We have no ‘*notion of...self..., when conceiv'd distinct from particular perceptions...we have no notion of...the mind..., distinct from the particular perceptions*’ (App§§18,19/635, my emphasis).<sup>167</sup>

---

<sup>164</sup> Strawson 2017, 253.

<sup>165</sup> Strawson 2017, 255.

<sup>166</sup> Strawson 2017, 255.

<sup>167</sup> Strawson 2017, 255.

Moreover, convincingly in my view, Strawson points to the fact that Hume as a self-declared skeptic and empiricist “doesn't claim to know the ultimate nature of things”<sup>168</sup> other than in the way they are experientially given to us. On the epistemological view, which Strawson defends, Hume only states that we do not really know what the nature of ourselves is and that, epistemologically speaking, all we can ever yield to know about ourselves is confined by perceptions. On such a reading, the result of Hume's analysis is mostly negative: It simply states that we cannot gain insight into an alleged substantial nature of self as per the soul-theory; it simply states that to conceive of the self as an abiding substance is experientially not warranted, because all we can possibly find in introspection are bundles of perceptions.

One important repercussion of highlighting the strong epistemological aspect of Hume's analysis is that the latter ought to be mostly considered as an attack on the soul-theory of self and as nothing much more. Such an attack, generally, could even leave open the possibility that we do possess or are souls. It is merely an attack on the claim uttered in favor of soul-theory that we have a sense of self that gives rise to the idea of a substantial self. Accordingly, denying that we have self-experience, understood in terms of experience of the self considered as a substance, does not necessarily rule out that we have a different form of self-experience whereby the ‘self’ in self-experience is not understood in terms of a substantial or object-like self.

In fact, that is precisely what Strawson maintains: that Hume's account does not deny the existence of the subject of experience and that it doesn't deny that the subject of experience appears in experience. To understand how Strawson comes to such a belief, it is necessary to consider another aspect of his interpretation of Hume, namely his understanding of ‘perception’. As will become evident in a moment, Strawson's discussion of ‘perception’ also illuminates how the notion of ‘experience’ works in Hume's argument and how a different notion of ‘experience’ may pave the way for an alternative conception of self-experience. This will further buttress the view that Hume's argument not necessarily entails a no-self view.

According to Strawson, we need to be aware of the “troublesome ambiguity”<sup>169</sup> of the notion of ‘perception’ in Hume's argument. On the one hand, namely, ‘perception’ is understood as a *type* of experience: perception in contrast to thinking, desiring, or imagining. On the other hand, however, with ‘perception’ Hume also denotes—more generally—any experiential occurrence whatsoever, i.e., any *token* of experience, be it an episode of perceiving *p*, desiring *p*, imagining *p* or thinking that *p*.<sup>170</sup> As Strawson remarks, if ‘perception’ is understood in terms of a certain type of experience, then certainly one

---

<sup>168</sup> Strawson 2017, 255.

<sup>169</sup> Strawson 2017, 260.

<sup>170</sup> Strawson 2017, 260.

might get the impression that Hume is saying that all we ever encounter when introspecting is some form of experience of perceptual kind, something like “an actual occurrence of redness-experience”<sup>171</sup> or any other form of sensory perception. If one understands ‘perception’ like that, then clearly, Strawson argues, Hume’s argument is interpreted as saying that whenever introspecting we only encounter sensory-like perceptions “without in any sense coming across a subject of experience”<sup>172</sup>. Yet, Strawson retorts:

It seems, however, enough to read and understand the words ‘when I enter most intimately into what I call myself...’ to be clear on the point that Hume believes in the existence of a subject of experience of some sort—as H. H. Price remarks.<sup>173</sup>

Moreover, Strawson highlights that “[o]f the fifty-nine occurrences of the word ‘perception’ in the passages of Hume’s writing that primarily concerns us now ... only two are even *prima facie* candidates for being type uses ...”<sup>174</sup>. Hence, Strawson suggests that, at least in the section on personal identity, one should read ‘perception’ in terms of the token-use. On that interpretation, when Hume says that humans are nothing but bundles of perceptions he only means that all we can possibly reveal about ourselves is the experiencing, i.e., the different experiential episodes be they of perceptual, thinking, desiring or imagining kind; and that the experiencing does not include anything that would transcend experience, i.e., experiencing does not involve the presence of an immortal soul that exists over and above the experiences which, as per the soul-theory, ought to inhere in the soul.<sup>175</sup>

Crucially, for Strawson, to say that all we find in introspection is the experiencing does not mean we could not encounter the subject of experience as something present *in* the experiencing: “But to say this is in no way to say that you don’t come upon a thinking being, a thinking being considered in its mental being, when you come upon the experience—the experiencing.”<sup>176</sup>

To see this, it is important to consider what Strawson calls the “*Experience/Experiencer Thesis*”<sup>177</sup>, which he formulates, referring to Frege and Shoemaker, in the following way: Any experience, any experiential episode *conceptually* requires an experiencer, someone undergoing the experiential episode. Accordingly, Strawson argues, when reflecting on the experiencing one is by the same token reflecting

---

<sup>171</sup> Strawson 2017, 261.

<sup>172</sup> Strawson 2017, 261.

<sup>173</sup> Strawson 2017, 261. See also Price 1940, 96–97.

<sup>174</sup> Strawson 2017, 260.

<sup>175</sup> Strawson 2017, 265.

<sup>176</sup> Strawson 2017, 265.

<sup>177</sup> Strawson 2017, 253.

upon on the experiencer who is conceptually implied in the experiencing. Yet, it would be a mistake, Strawson warns us, to believe that the subject of experience must therefore make an appearance in the experience in the same way the experiential content of a given experiential episode—for instance, a tree or the rising of the sun—is present. In this sense, “Hume’s phenomenological report ... is correct: there is indeed no thetic presentation of the subject as such”<sup>178</sup>. That, however, does not make the subject of experience unreal. Rather, as Strawson goes on, to draw such a conclusion stems from a wrong expectation:

The natural idea is that any such presentation would *ipso facto* be presentation of the subject as something in some way distinct from the experience as a whole; that this is what a thetic focusing of attention on the subject would inevitably amount to, a singling out of the subject; and this, as Hume rightly says, is precisely what does not happen. And yet there is, nevertheless, awareness of the subject.<sup>179</sup>

According to Strawson, thus, experiencing involves a non-thetic sense of self, a form of “presence”<sup>180</sup> of the subject of experience distinct to the presentation of experiential content, which might be “marginal” but “none the less real, a real concrete occurrence, part of the actual course of one’s experience”<sup>181</sup>. To sum up, Strawson believes that Hume never denied such a sense of self as an intrinsic aspect of any experiential episode. Although Hume himself might not have had a clear understanding of how to conceive of such a non-thetic form of sense of self, his analysis in the *Treatise*, according to Strawson, does not rule out a notion of self-experience of this kind.

Strawson calls John Foster as a witness for his claim, who precisely formulates a response to Hume along such lines. According to Foster,

a natural response to Hume would be to say that, even if we cannot detect ourselves apart from our perceptions (our conscious experiences), we can at least detect ourselves in them—that when I introspectively detect an experience, what is revealed is the complex of myself-experiencing-something or myself experiencing-in-a-certain-manner. Indeed, it is not clear in what sense an experience *could* be introspectively detected without the detection of its subject.<sup>182</sup>

However, unlike Foster, Strawson does not conceive of the notion of experience as a complex that includes a sense of self as a *response* to Hume. Rather, he believes that the line of thinking the quote by Foster circumscribes just *is* Hume’s position: John Foster’s

---

<sup>178</sup> Strawson 2017, 269.

<sup>179</sup> Strawson 2017, 269.

<sup>180</sup> Strawson 2017, 267.

<sup>181</sup> Strawson 2017, 271.

<sup>182</sup> Foster 1991, 215.



“only error, in my view, is to think that Hume ever thought otherwise. Hume’s target was the simple unchanging persisting subject of the philosophers and the church, nothing less and nothing more.”<sup>183</sup> Hence, Strawson believes Hume neither denies the possibility of self-experience nor the reality of self.

Is Strawson’s fairly controversial and in the ears of some Hume scholars perhaps even provocative interpretation of Hume’s treatment of self in the *Treatise* convincing? Since my interest is not with scholarship or Hume particularly, I remain agnostic about whether the depiction by Strawson best matches with what Hume really had in mind. The focus of my interest rather lies with the systematic aspects of Hume’s position, its scope and force. Thus, here, the more important question is what exactly we can learn from Strawson’s discussion with regard to the four questions I mentioned at the beginning of this section: Does denying self-experience per se amount to a no-self doctrine? Does Hume’s argument rule out alternative understandings of self-experience and self? Is Hume’s introspective analysis phenomenologically accurate?

Much I believe can be learned from Strawson’s discussion of Hume’s position. First of all, arguing in favor of an epistemological rather than ontological reading of Hume, Strawson shows that the denial of self-experience as such need *not* amount to a denial of self. For, one might simply hold that as far as we know from our experience we have no reason to *assume* the reality of self as justified, while one may still *believe* in the latter, for instance based on religious or other eschatological convictions. To say we do not have self-experience or self-knowledge is not tantamount to saying we know there is no self. Accordingly, one might argue that for the being of selfhood, self-experience is inessential. On such a line of thinking, selfhood might be real even though I do not undergo self-experience. Moreover, with Strawson, one could hold that Hume only denied the reality of a certain idea of self, the incoherent notion of self as an abiding substance proposed by soul-theory, and believe that Hume’s account would be compatible with an alternative notion of self in which case it might be possible to say selfhood is real.

Strawson, as mentioned, further argues that Hume indeed had an alternative notion of self in mind, namely the subject of experience as a necessary structural component of the experience itself. He believes that Hume would have or has in fact accepted the idea of a subject of experience that is intrinsic to experience and that is not conceived as a soul-like substance existing independently of experiences. Foster might not believe that Hume had this kind of idea of self in mind. But he believed that, with such a notion of self, an objection against a Humean denial of self-experience is possible. Now, regardless of whether Strawson or Foster is right, it seems that an alternative notion of self remains untouched by the strong skepticism formulated by Hume against an alleged self-experience (understood as the experience of a substantial self). At least, and that was a further point

---

<sup>183</sup> Strawson 2017, 272.

by Strawson, as long as one understands ‘perception’ as the general notion for experience and not as a certain *type* of experience: perception in contrast to thinking, desiring or imagining. If one were to understand ‘perception’ in terms of sensory perception only, i.e., as experience confined by our senses, only then is Hume’s dictum that all we find in introspection are ‘perceptions’ a clear rejection of self-experience *in toto*. For indeed impossible it seems that self-experience could consist in a sensory-like datum on a par with experiences such as the redness of a tomato, its taste, its smells, the way it feels when I squeeze it or the sound it makes when I drop it on the floor. Crucially, if Hume were indeed to mean *that*, unlike Strawson maintains, then one could still formulate the criticism that he must though fail to provide an argument why we should only allow *sensory* experience to count as experience. Again, it is Foster who makes a similar point like Strawson with the only difference that he construes it as an argument *against* Hume. The three following quotes illuminate his line of reasoning well:

One weakness in this argument is that the radically empiricist view on which it relies is open to question. The underlying assumption is that any genuine idea (coherent concept) must draw its whole content, either directly or by analysis, from the data sensory and introspective experience. And it is far from clear why this principle should be accepted.<sup>184</sup>

A further weakness in the argument is that, even if the empiricist principle is accepted, it is not clear why we need to have an impression of the self in order for our concept to pass the empiricist test. ... We might still be able to achieve an empirically respectable conception of the self in some more complex way—presumably in terms of its role as a basic subject with respect to the introspectible items of mentality.<sup>185</sup>

Hume seems to be supposing that introspection works like sense-perception: a mental item is presented as the object of introspective awareness in the same way as a color-pattern is presented as the object of visual awareness. On this presentational model, it is hardly surprising that he finds no room for self-awareness.<sup>186</sup>

Foster criticizes that Hume’s empiricist doctrine of experience is not only in need of an independent argument, but, even if such a doctrine were true, he says, there would still be the possibility of a positive description of self-experience “empirically respectable” and despite the fact that there is no one impression that corresponds to the self. The mistake of

---

<sup>184</sup> Foster 1991, 213.

<sup>185</sup> Foster 1991, 214.

<sup>186</sup> Foster 1991, 215.

Hume is, Foster states, that he wrongly conceived of “introspection”<sup>187</sup> in terms of sensory-perception. Put differently, Foster argues that it is problematic to apply the standards of the experience of a perceptual world-object to the way we might be said to experience ourselves. The alleged denial of self-experience, thus, is an artefact of a wrong conception of experience. For, only once self-awareness is evaluated against the criteria of sensory perception, then self-awareness becomes impossible.

Strawson claims Hume never understood ‘perception’ in such a restricted way when he said that the mind is constituted by nothing but a ‘bundle of perception’. Foster claims Hume did. Who is right? Again, for my analysis, it matters not. *Either* Hume had indeed the narrow understanding in mind when he made his statement. In this case, one has to remark that his rejection of self-experience is limited in scope insofar as it presupposes a notion of experience that reduces experience to sensory experience, which requires a further argument. *Or* Hume, like Strawson thinks, never intended to reject all forms of self-experience but only such self-experience that is conceived of in terms of the experience of a substantial self.

In both cases, we must conclude that the Humean denial of self-experience is very limited. For, either it is not meant as a denial of all kinds of self-experience or a further discussion of the reasons why we should believe self-experience if to exist at all must— notably, *per impossibile*—be a kind of sensory experience. Put differently, if one wants to follow a line of argument according to which there is no self-experience, one has to offer an argument as to why self-experience or any other kind of experience must be an experience that is characterized by the phenomenological quality of sensory perception— regardless of whether one attributes such a line of argument to Hume or not.

## 2.3 The No-Self-Picture is phenomenologically inaccurate

Whether all experience can be reduced to a form of sensory experience or sensory phenomenology is a question that is far from evoking agreement among philosophers and thinkers. Recently, the question has been much debated under the label of the “cognitive phenomenology debate”<sup>188</sup>. Roughly, two camps can be distinguished that give different answers to that question. Proponents of so-called “cognitive phenomenology”<sup>189</sup> hold that our experience is much richer than the experiences that can be described by properties of

---

<sup>187</sup> In the quotes “introspection” seems to refer to both introspection as a kind of awareness of one’s experiences and introspection as the experiential act of reflection in which one reflectively attends one own experiential process.

<sup>188</sup> Bayne & Montague 2011, 3.

<sup>189</sup> Chudnoff 2015, Kriegel 2015, Montague 2016, Strawson 2010, 2011, Pitt 2004, Siewert 2011.

our sensory perception. Particularly, they argue that thinking has a *sui generis* kind of phenomenology that cannot be reduced to the phenomenology of sense perception: cognitive phenomenology. Opponents of cognitive phenomenology, by contrast, defend the view that any experience whatsoever can be described by referring to qualitative characteristics ascribed to sense perception: sensory phenomenology. On the latter, “restrictive”<sup>190</sup> view, if there is any phenomenology of thinking at all, it must be considered as some form of sensory phenomenology rather than a *sui generis* kind of phenomenology.

Although, here, I cannot discuss the specifics of this wide-reaching and complex debate, the only point I wish to make is that a Hume-inspired denial of self-experience would need to specify why we should conceive of experience in terms of sensory phenomenology. Since it is also not entirely clear how we are to understand Hume’s position on these matters—Strawson seems to take Hume as proponent of cognitive phenomenology, while Foster ascribes to Hume the conservative view that reduces experience to sensory phenomenology—it is also hard to find a conclusive answer to the question whether his introspective analysis is phenomenologically adequate. My focus, to repeat, lies not on what the historical Hume had in mind but rather how far the no-self doctrine ascribed to him reaches and to what extent it might be convincing. The phenomenological question then is whether the picture of experience tied to the no-self doctrine, according to which our mind is constituted by *bundles of sensory-like perceptions*, is phenomenologically adequate.

Such a picture of experience, let me call it the *No-Self-Picture*, is characterized by three interrelated phenomenological claims:

- (I) *Sensory phenomenology*: Any experiential content has the phenomenological quality of sensory phenomenology. The phenomenology of experience can be exhaustively described in terms of sensory phenomenology.
- (II) *Bundles*: Our experience is constituted by bundles of sensory-like experiential episodes.
- (III) *No sense of self*: Our experience does not include any sense of self.

Setting aside the issue of (I) sensory phenomenology and (III) sense of self or the lack thereof for a moment, I now want to address the second phenomenological claim: (II) that our mind is constituted by *bundles* of perceptions. Since my interest now is with the aspect of bundles, I will consider a permissive or broad understanding of ‘perception’ and will allow the latter to include any kind of experience, also a tentative non-sensory self-experience. Doing so allows me to apply a criticism formulated by Evander B. McGilvary

---

<sup>190</sup> Prinz 2011, 175.

(1907)—today rather unknown—against William James’s theory of succeeding *distinct subjects of experience* to Hume’s bundle theory.

### 2.3.1 McGilvary on James

In his *Principles of Psychology*, James equated the different experiential episodes that follow each other in our mind each with a distinct ego. The “thought is itself the thinker”<sup>191</sup>, he concluded his chapter on ‘The Consciousness of Self’. With every single thought, therefore, he postulated a corresponding single subject of experience that vanished together with the end of the thought. Just like Hume, thus, James proposed a bundle theory. However, he conceived of the bundles that are to constitute our mind not of mere ‘perceptions’ but of ‘thoughts’ that he says amount each to an isolated and distinct ‘thinker’. Our mind, according to the Jamesian doctrine, thus, embraces a bundle of “little egos”<sup>192</sup>.

Such an idea of a *bundle*, however, seems problematic to McGilvary who attacks the Jamesian view of “temporal psychological atomism”<sup>193</sup>. According to McGilvary, the latter stands in sheer contradiction to another doctrine James himself proposes and to which McGilvary ascribes much more plausibility: the idea of the *stream* of consciousness. He points to the following quotation by James to indicate what he believes is the true and real empirical fact, the way our experience, from a phenomenological perspective, really is:

Consciousness, then, does not appear to itself chopped up in bits. Such words as ‘chain’ or ‘train’ do not describe it fitly as it presents itself in the first instance. It is nothing jointed; it flows. A ‘river’ or ‘stream’ are the metaphors by which it is most naturally described.<sup>194</sup>

Although these are his words, James nonetheless maintained that thoughts came in “pulses”<sup>195</sup> and with the thought the “thinker”<sup>196</sup>. What made him yield a contradictory verdict of this sort? According to McGilvary, it is due to a confusion James had—pertinently on account of McGilvary—warned against himself: “the confusion ... between the *thoughts themselves*, taken as subjective facts, and the *things of which they are aware*”<sup>197</sup>. On McGilvary’s view, we may find something like ‘pulses’ within our

---

<sup>191</sup> James 1950/1890, 401.

<sup>192</sup> McGilvary 1907, 230.

<sup>193</sup> McGilvary 1907, 232.

<sup>194</sup> James 1950/1890, 239.

<sup>195</sup> James 1950/1890, 371.

<sup>196</sup> James 1950/1890, 401.

<sup>197</sup> James 1950/1890, 240 (italics by McGilvary), McGilvary 1907, 226.

experiences when it comes to the objective content of our experience, say, now I perceive a tomato and suddenly, instead of looking at the tomato, I fully attend an imagined lemon. In the transition of the tomato-perception to the lemon-imagination there seems to be a clear cut. The tomato-perception is one pulse, the lemon-imagination another. But these different pulses, McGilvary emphasizes, are only on the side of the object of thought, not on the subjective side of thinking:

[M]y contention in this paper is that *the doctrine of 'perishing pulses of thought' within the empirically continuous stream of thought is the result of just this confusion. The stream of thought is the fact, the pulses of thought are not facts; the only pulses are in the objects of thought.*<sup>198</sup>

The idea is that, yes, content-wise our stream of consciousness may sometimes seem jumpy, but these transitions that can happen quickly are something, according to McGilvary, that occurs *within* our experience. Again, he refers to a quote by James himself who seems to agree:

A silence may be broken by a thunder-clap, and we may be so stunned and confused for a moment by the shock as to give no instant account to ourselves of what has happened. But that very confusion is a mental state, and a state that passes us straight from the silence to the sound. The transition between the thought of one object and the thought of another is no more a break in the *thought* than a joint in a bamboo is a break in the wood. It is a part of the *consciousness* as much as the joint is a part of the *bamboo*.<sup>199</sup>

Hence, we should not believe that only because the content of our experience, the object of thought, may change and often changes quickly, that thinking or experiencing as such was interrupted. The transition itself is experienced and demonstrates the continuity of the stream of consciousness. Moreover, the continuity or unity of the stream of consciousness, according to McGilvary, is also manifest in the present moment of consciousness:

[B]ut rather must we say that in our knowledge of objects, the objects known as past or future are objects of the *one* knowledge that also knows the present.”

...

The speciously present consciousness continues flowing down the stream of time, always comprehending some past object as past, but, at every moment, itself so much of a unity that at no time in its steady flow can it be, except by *a violent abstraction*, spoken of as a new ego coming to birth immediately on the decease of its predecessor.<sup>200</sup>

---

<sup>198</sup> McGilvary 1907, 226.

<sup>199</sup> McGilvary 1907, 226.

<sup>200</sup> McGilvary 1907, 229 (my italics).

According to McGilvary, phenomenologically speaking, I experience myself—or more neutrally, experience experiences itself—as being the very same one who experienced the tomato-perception in the past and the lemon-imagination now. The tomato is an ‘object for the one knowledge that also knows the present’ object, the lemon. Only by ‘violent abstraction’ is it possible to postulate two distinct subjects of experience, one of the perceptual experience of the tomato, another of the imagining experience of the lemon. Moreover, such an abstraction is only possible by the mentioned confusion between the *object of thought* with the *process of thinking*. There are clearly two distinct objects, a tomato and a lemon, but to infer from that that there are two distinct subjects, according to McGilvary, is a mistake. For, the shift in the objects given in experience does not imply an interruption of the stream of consciousness: both objects are ‘known’ by the same ‘knowledge’, they are both given to the same consciousness, one ‘as past’, the other ‘as present’.

However, McGilvary, most certainly referring to dreamless sleep, does admit that there are cases in which the stream of consciousness gets interrupted, “as periodically it does”<sup>201</sup>. “When this happens we have a real breach in the continuity of consciousness and at the same time an ‘appropriation’ by a later consciousness of the selfhood involved in a prior one.” In these cases, McGilvary argues, the stream might be interrupted, the past consciousness is ‘appropriated’ by later consciousness. McGilvary takes such an appropriation as a phenomenological datum but anticipates that one might demand further explanation. Here he points to the Jamesian dictum of the “‘quality of warmth and intimacy and immediacy’ which attaches to the content of the prior consciousness”<sup>202</sup>, which “seems to be the reason, when the quality appears in the content of a subsequent consciousness, for the latter’s identification of itself with the former”<sup>203</sup>. Put differently, the identification with a remembered past experience occurs because the latter possesses a certain phenomenological quality of warmth, intimacy, and immediacy. But, crucially, McGilvary holds that for such identificatory process to be possibly successful at all, a prior identity inherent to any moment of consciousness is necessary:

But identity here has a meaning, furnished by the previously felt continuity of the unbroken consciousness with itself; and the same ‘warmth and intimacy’ which characterized all the contents of the former consciousness can now be considered, when appearing as marks of the contents of the second consciousness, as pointing to an identity of the two consciousnesses—an

---

<sup>201</sup> McGilvary 1907, 229.

<sup>202</sup> McGilvary 1907, 230.

<sup>203</sup> McGilvary 1907, 230.

identity similar to the actual identity which the first consciousness felt in itself.<sup>204</sup>

What McGilvary is trying to say is that if we feel, for instance when waking up after a long night, continuous with the conscious process that we remember as having taken place before we fell asleep, then this is only because our on-going experiential process as such is continuous. The continuity is the primary phenomenon. And, it is only because consciousness feels structurally continuous that remembered episodes of consciousness that are present as remembered with a similar warmth, intimacy or immediacy, are experienced as belonging to the same stream of consciousness, even though the latter seems to have been interrupted during phases of dreamless sleep.

The argument, I suggest, thus, can be formulated as follows: Those who believe that our experience is constituted by bundles of perceptions, thoughts, or experiences *either* need to say that we experience bundles, i.e., different perceptions, thoughts or experiences as separated in a strict sense; *or*, if they accept the *phenomenology* of continuity of experience but maintain that really there are *ontologically* separated bundles of experience that constitute experience, they need to come up with an explanation how it is that we experience continuity where there really ought to be just separate bundles. James's explanation is that of an appropriation through the quality of 'warmth, intimacy, and immediacy'. But such quality is essentially a feature of the on-going, i.e., present, experiential process. It can only function for appropriation because it denotes the continuity of the on-going experiential process. Hence, its function for appropriation is derived or lent from its undoubted function in the on-going present experiential process. Put differently, one has to accept the phenomenology of continuity as a valid characterization of the current on-going experiential process if one wants to claim that it is such quality that produces the feeling of continuity with past experiences that seem to have occurred before a gap of consciousness (e.g., during dreamless sleep).

What Professor James, therefore, seems to have done is to have taken a psychological explanation of the mediated identity of two successive, but actually separated, consciousnesses and to have used it in the explanation of a continuity which is not interrupted. But before he can so use the explanation he must break the actual continuity of consciousness into successive 'thoughts'. These separate 'thoughts', thus obtained in an alleged isolation, may then be reunited by the same principle that his 'Psychology' proposes with great plausibility to employ in uniting what is actually separated.<sup>205</sup>

According to McGilvary, James wrongly took an explanation of how consciousness despite being separated—for instance, through dreamless sleep—may still feel continuous as the

---

<sup>204</sup> McGilvary 1907, 230.

<sup>205</sup> McGilvary 1907, 230.



principle for how consciousness generally feels continuous. In doing so, he wrongly took the separation of consciousness to be the general case and used an explanation that only works if continuity is the general case and that is meant to explain the continuity between really separated episodes of consciousness (as in dreamless sleep) as an explanation of why we feel continuity to begin with. That way, James first artificially separates the continuous flow into little pieces and then applies an explanatory principle derived from the phenomenology of the continuous flow. And that, McGilvary argues, is not possible:

This looks like making a problem where none exists, in order to solve it. It comes in the last resort to exactly the same sort of procedure that Professor James finds fault with in Hume. Hume, according to Professor James, tried to glue up the *disjecta membra* of consciousness by using the principle of association. Professor James uses two kinds of glue, ‘feelings of transition’ and the ‘quality of warmth and intimacy.’ But neither Hume nor Professor James succeeds in making the glued aggregate anything more than a counterfeit presentment of the really continuous psychic life.<sup>206</sup>

To conclude, McGilvary holds that the basic phenomenon of our experience is that of a ‘continuous psychic life’ and that any atomistic view about experience is only possible through a, phenomenologically speaking, secondary abstraction from the continuous life: one considers the different *contents* or *objects of experience* and divides *the process of experiencing* into little pieces corresponding to the distinct objects of experience; then one tries to explain how it comes that the distinct bundles of experiences, despite allegedly being separated, may make their appearance in the same consciousness and points to transition markers (‘warmth, intimacy, and immediacy’), which only can so function if continuity of the psychic life is already presupposed.

### 2.3.2 Applying McGilvary’s critique to the No-Self-Picture

While McGilvary directs his line of argument against James, who indeed equated the thinker (experiencer) with the thought (experience), now the question arises as to where his argument leaves us with regard to our assessment of Hume’s phenomenological claim that the mind is constituted by bundles of perception? Again, we have to decide how we read ‘perception’ in order to not mix things up. If we read ‘perception’ in the way Strawson suggests, namely as something that simply refers to the general notion of experience which always is to include an experiencer, Hume’s theory of bundles amounts to something quite similar to the Jamesian view of succeeding egos (and Strawson’s James-inspired “pearl view”<sup>207</sup>). Yet, here, my interest is with the plausibility of the no-self-doctrine ascribed to

---

<sup>206</sup> McGilvary 1907, 230.

<sup>207</sup> Strawson 1997, 424.

Hume and not with a theory of a plurality of selves constituting the mind,<sup>208</sup> so I want to know whether McGilvary's criticism of James can be applied to the no-self version of the bundle view. McGilvary accused James of having confounded the *object of experience* and the *process of experiencing*. Does the accusation apply to the "sensationalists", who claim that the mind is constituted by nothing else than "sensations and their copies and derivatives, juxtaposed like dominoes in a game, but really separate, everything else verbal illusion"<sup>209</sup>? How can there be a confusion between *object of experience* and *process of thinking*, if one holds that the mind exhaustively consists of the sensations or impressions or sensory-like perceptual experiences? The defender of the no-self-view who takes the mind to be really nothing else than succeeding impressions and its compositions seems to be able to avoid the verdict of McGilvary by pointing to the claim that the *process of experiencing* as such is not a phenomenological datum characterizing the phenomenology of experience. What Hume had in mind, on this line of thinking, was the view that, when introspecting, we really do *not* get *anything else* in sight other than some specific impression or perception built upon on a complex of impressions. Thus, not even the "flux"<sup>210</sup> Hume speaks about, is a phenomenological datum, but rather 'verbal illusion', a proponent of the no-self doctrine might bring on.<sup>211</sup> Hence, on this line of thinking, since there is just some single or group of impressions to be found in our experience and since there is nothing else to experience (no *subjective aspect* of the *process of experiencing*), we can separate experience on the basis of its objects. Since there are only the objects of experience, and of distinct objects of experience there are many, we are completely justified in claiming that experience is constituted by *bundles*.

Yet, although, *prima facie*, such a line of thinking might seem potentially successful, it has a few problems the weight of which is too heavy.

### 1) *How to account for the phenomenology of memory?*

First, and that is a problem Hume completely acknowledges,<sup>212</sup> if only bundles of experience and nothing else are to constitute our minds, one wonders how it is possible that we possess a memory of past experiences. Aren't objects that are remembered as experienced in the past not also *experienced* as past? Remembering an object that I experienced in the past does not mean that it exhaustively fills my current experience again. Rather my remembering of a past object is itself a different kind of experience, more specifically, it is an experience united with a different experience of some present object.

---

<sup>208</sup> I will discuss Strawson's 'pearl-view' in Section 5.5.

<sup>209</sup> James 1950/1890, 245.

<sup>210</sup> Hume 2007/1739, 165 (1.4.6.4).

<sup>211</sup> I will discuss more recent versions of such a view in Section 2.5.

<sup>212</sup> Hume 2007/1739, 400 (App. 20).

Sitting at the table and remembering what happened last night, I have a present experience of the table and other objects currently around me *and* some form of experience of the objects of experience of last night or the whole experiential episode considered as a present object of experience. How is this unity possible if my mind ought to consist of nothing but ‘some one perception’ after another? Hume found that question “too hard”<sup>213</sup>. But the problem is not so much that the version of the no-self-doctrine ascribed to him cannot give a plausible explanation of how it comes that we experience this sort of unity. Rather, if the experienced unity is indeed a fact, which I believe it is, then this already renders the account of mere bundles of succeeding unrelated impressions implausible.

2) *The justification of the claim of the bundle view is not compatible with its own terms*

Second, if one were to deny that such a unity is a phenomenological characteristic of our experience and wishes to make the claim that our experience is characterized by a constant shift of separated and unconnected perceptual experiences, one wonders how such a verdict could possibly be justified? If my mind were constituted in the way the hypothesis assumes, then all what can be found when reflecting on my experience is ‘some one perception’. But if that is so, how can I make any judgement about *all* experiences that are to constitute my mind if all I ever am possibly aware of is ‘some one perception’? To form a judgment about my experience based on introspection, I need some way to gain an overview about the different single ‘perceptions’ that ought to be distinct and separated. But if they are intrinsically unrelated and each contains no experiential connection with the others, it seems impossible to understand how I can even know I have different ‘perceptions’. If any experiential episode consists of nothing but some current perceptual content and *nothing more*, then introspection will only reveal the present perceptual content, which would not be sufficient in order to draw any conclusions about some alleged other past or future perceptual content. For introspection to be possible and to be able to reveal anything about experience in general and not only something about the current token experience, it is required, thus, that the *process of experiencing* is somehow manifest *in experience*. Otherwise I could never say that a certain number of ‘perceptions’ constitute my mind. McGilvary puts the point in the following way:

If, now, we are hardy to believe that psychology is possible, we must believe that there is an awareness of awareness whenever there is an awareness of objects.<sup>214</sup>

---

<sup>213</sup> Hume 2007/1739, 400 (App. 21).

<sup>214</sup> McGilvary 1907, 234.

To put it in yet other words: If my experience at any single point of time  $t_1$  were to consist only in an awareness of objects, say  $o1$ ,  $o2$ , and  $o3$ , then at a later point of time in which I have an awareness of  $o4$ ,  $o5$ , and  $o6$ , how could introspection at  $t_2$  in any way reach out to  $t_1$ ? Introspection would ever only encounter  $o4$ ,  $o5$ , and  $o6$ . Of course, one might say that during introspecting one could also remember  $o1$ ,  $o2$ , and  $o3$ . But to remember  $o1$ ,  $o2$ , and  $o3$  as a *past* experience of  $t_1$  that preceded  $t_2$ , one has to put  $t_1$  and  $t_2$  in an according order. But if experience at  $t_1$  and  $t_2$  each only involved its objects and absolutely nothing else, it is just incomprehensible to see how this should be possible at all.

A similar point has been made by a more recent thinker, Dan Zahavi, when he emphasizes:

Ultimately, reflection presupposes the constitution of a temporal horizon. It would be impossible without the ecstatic unity of the flow. When reflection sets in, it initially grasps something that has just elapsed, namely, the motivating prereflective phase of the flow. The reason why this phase can still be thematized by the subsequent reflection is that it does not disappear, i.e., it is not cut off from the Living Present but remains united with it through retention.

...

When an experience occurs, it automatically acquires an unchangeable location in the stream. I can only locate an act in recollection if it already has a position.<sup>215</sup>

Hence, in order to have some sort of Humean introspective knowledge about how one's own mind is constituted, one must already be familiar with one's own experiences and their temporal location. Only then is it possible to distinguish different tokens of experience and their temporal order. And only then would it be possible to make the claim that all these tokens of experience are nothing but 'some one perception'. But once it is clear that in order to be able to make a judgement about a set of constantly changing perceptions one needs to have a consciousness of them as being succeeding, i.e., a consciousness of succession, it is not possible to make the claim any longer that any token of experience exhaustively consists of perceptual content and—strictly—nothing else.

Hume is only able to describe the 'flux' in introspection, because he has some form of knowledge about or consciousness of the fact that *different experiences* constitute the flux, and so he does need to have some form of knowledge about or consciousness of the fact that these different experiences are somehow *continuous*. And, as McGilvary further highlights:

To experience an experience as continuous with a past experience, must one not have an experience of experience? It matters not whether we call this self-

---

<sup>215</sup> Zahavi 1999, 79.

awareness an experience of experience, or a knowledge of knowledge, or a consciousness of consciousness, or an awareness of awareness ...<sup>216</sup>

To say different experiences constitute a flux, presupposes that one takes them to be continuous. But to take them as continuous is only possible if each experiential episode that is said to belong to a singular flux is imbued with self-consciousness. It is only because the different perceptions are experienced *as being experienced by the same consciousness that experiences the present objects* that I can reflect upon my current experience and compare it with some remembered experiences.

Accordingly, it does *not* seem possible to claim that the *process of experiencing* itself is *not* a phenomenological aspect of experience. And therefore, one must conclude that sensationalists who conceive of the stream of consciousness as a mere flux of succeeding and separated perceptual contents equally confound the *subjective process of experiencing* with the *experienced contents*. Instead of distinguishing between the awareness of an object and the awareness of awareness (or self-awareness), they simply hold that awareness of objects is the only kind of awareness there is.

### 3) *The reflection of experience x and x itself become indistinguishable*

There is another related problem with the bundle-view of experience. If awareness of objects ought to be the only kind of awareness there is, how can one distinguish between the *experience* of seeing a red tomato and the *reflection upon the experience* of seeing a red tomato? Or put differently, in words more akin to Hume, how can one distinguish a perception and one's introspection of the very same perception? Are the experience of introspecting while seeing a red tomato and simply undergoing the experience of seeing a red tomato phenomenologically congruent? Or isn't there at least a slight difference?

*Either* there is a difference, then it seems the difference can hardly be explained by a difference in perceptual content. Rather, any such difference seems only to be explainable by an awareness of the very act one is undergoing in each case: perception vs. introspection, i.e., by an awareness of the *subjective process of experiencing*.

*Or* there is no difference. But then introspection amounts to just the same as the perception. And this is problematic because then, as mentioned before, while introspecting I just have the perception and nothing else, which also means that the 'nothing else' does not come into sight *as a fact* about my experience. Rather, my experience is completely absorbed by the perceptual content. It is only in reflection that I can become aware that the experience I am reflecting on was exhaustively concerned with the perceptual content, the red tomato. While being exhaustively concerned with the perceptual content, the red tomato, experience is *not* characterized by some awareness of the fact that nothing else than

---

<sup>216</sup> McGilvary 1907, 233.

the perceptual content is experienced. Such an awareness can only be the result of an examination of the experience of seeing a red tomato as a whole in reflection and introspection. But if it is claimed there is no difference between perception and introspection, then I technically could not come up with any insight into the specific quality of my experience of that sort. Whenever I were asked to comment on *what it is like* to undergo the experience I am currently having, all I could possibly be saying is 'I see a red tomato'. But from this answer we could never draw any conclusion about the general character of experience that introspection is to reveal. All we could say is that we have this or that perception. What we could precisely not say, like Hume suggests, is that we *always* have this or that perception. A single token perceptual content cannot provide comprehensive insight into the general character of experience. If introspection ought to be just like perception, then introspection equally cannot provide such insight any longer. Introspection would always just present some single perception and *nothing else*.

Hence, either must we admit there is a difference between perception and introspection. Then the bundle view of the No-Self-Picture has already been given up. For in that case one already concedes a difference between perception and introspection which cannot possibly be explained by referring to perceptual content only but only by some difference in the kind of awareness at play: Conceding that there are different forms of awareness is tantamount to acknowledging that awareness cannot be reduced to awareness of sensory-like, perceptual content. Distinguishing between perception and introspection amounts to acknowledging that the *subjective process of experiencing* somehow must be a phenomenological aspect of experience the awareness of which is distinct to the awareness of perceptual content.

Or we maintain that there is no difference between perception and introspection. But then we cannot defend the bundle view of the No-Self-Picture to begin with. For, in that case, introspection *ex hypothesi* loses its specific meaning and is reduced to some single token perception. Accordingly, all we could ever do is to state that our *current* experience is that of a perception and report what we see, hear, smell, taste or feel through touch. But we could never make the statement that *all* our experiences are of such a mere perceptual kind. If introspection were to be identified with perception, then we would have to be much more modest regarding our knowledge about any other experiences, for introspection would always only consist in one single perception. In maintaining that our introspection shows that all our experiences are nothing but perceptions, the No-Self-Picture problematically transcends its own presuppositions and proposes a verdict that, on its own terms, could not possibly be justified.

#### 4) *A wrong understanding of introspection and reflection*

McGilvary provides a fourth argument against the notion of our experience as a bundle of succeeding separated perceptions (each with a different subject of experience). He describes the tendency to distinguish between distinct perceptions as the isolated elements of our experience as a fallacy that stems from a wrong interpretation of the reflective or introspective act. Thus, he considers the bundle view to be an artefact of a wrong understanding of reflection and introspection. Again, he argues, such an error occurs as a consequence of a confusion of the *object of experience* with the *subjective process of experiencing*. The idea is that when one reflects upon the experience of seeing a red tomato, one makes the *subjective process* of seeing a red tomato the *object* of one's reflective or introspective experience. And then, since objects are usually considered to be distinct and separable, one wrongly infers that each reflective or introspective act witnesses a distinct and separable object: an isolated and independent 'bit' of the subjective process of experiencing. McGilvary puts the point in the following way:

It looks very much as if the psychologist who denies such awareness of awareness himself commits a fallacy very similar to the psychologist's fallacy. The psychologist examines his experiences *objectively* so far as he can, and then he is prone to think that because the psychosis under examination is known by him as distinct from the psychosis with which he knows, therefore every psychosis in order to be known requires as knower a distinct psychosis.<sup>217</sup>

The accusation, correctly in my view, is that since *self-awareness* is not acknowledged by the No-Self-Picture, while reflecting on *my* experience, I need to take the current reflective act to be distinct from the act under reflective scrutiny and oversee that the act reflected upon is *my* act, i.e., an act of the very same *subjective process of experiencing* that is currently doing the reflection. In doing so, the act reflected upon is considered as an object completely distinct and separated from the subjective process of experiencing through which it is given. For, reflection or introspection is construed as a form of pure object-awareness and not as a form of self-awareness in which the subjective process of experiencing is examined as an object. The mistake, according to McGilvary, thus, consists in confounding 'objective self-awareness' with 'mere object-awareness'. If reflection and introspection are taken as forms of mere object-awareness bereft of any sense of self, then each token reflection and introspection will discover an experiential act or phase of the subjective process of experiencing as a distinct object. But again, if that view were correct, it is throughout unclear how the distinct objects (the alleged isolated parts of the subjective process of experiencing) revealed by allegedly distinct and independent acts of reflection could all be taken as belonging to the one flux that ought to constitute our mind.

---

<sup>217</sup> McGilvary 1907, 233.

To sum up, the bundle view of the No-Self-Picture is incoherent for various interrelated reasons and does not present a phenomenologically accurate depiction of the phenomenology of our experience. The main problem with it is that it not only denies the phenomenology of continuity of our experience, which most take for granted, but that it also requires the temporal unity of different episodes of experience in order to make its own case. That renders the bundle view self-contradictory and thus self-defeating. To make the claim that the flux of experience is made of separate bundles of perception only it needs to posit some form of consciousness of the fact that the alleged separate bundles of perception belong to the same flux of experience. Without any unifying consciousness of the different perceptual contents that are said to run through our mind, all what could be possibly be said on the bundle view's own terms would be one isolated bundle of some concrete though not unified perceptual content. But then no talk of any flux or constantly changing perceptions would be possible.

My focus has been on the phenomenological claim (II) of the No-Self-Picture: the bundle view. However, as the discussion showed, the three phenomenological claims constituting the No-Self-Picture are strongly interrelated. The bundle view, i.e., maintaining that experience is constituted by the succeeding perceptions only, implies the notion that our experience is bereft of any self-experience because experience is reduced to experience of the perceptual content, the *object of experience*. This includes the claim that the *subjective process of experiencing* doesn't feature as a phenomenological aspect in experience. Claiming that the subjective process of experiencing is not phenomenologically manifest just *is* saying that experience consists of bundles of separate perceptions only. Moreover, emphasizing that any experience is characterized by sensory phenomenology *is* saying that something like self-experience which is precisely not the experience of something given through our senses is impossible. Saying that any experience is of the kind of sensory-like perception just *is* excluding self-experience.

Accordingly, if my discussion and rejection of the bundle view is correct, then this has also significant ramifications for the two other phenomenological claims of the No-Self-Picture: (I) the view that all experience can exhaustively be described by pointing to the phenomenological quality of sensory phenomenology and (III) that our experience lacks any sense of self. In fact, if our experience is more than the appearance of succeeding isolated perceptual contents, then this might cast severe doubt on the view that all the phenomenology of experience there is must be sensory phenomenology. Moreover, if it is true that we need to acknowledge the fact that our experience involves some form of consciousness of the continuity with which the different experiential episodes follow each other, we must accept that the *subjective process of experiencing* is manifest in experience



somehow, along with the different *objects of experience*. But if that is so, then we also need to acknowledge that such experiential manifestation of the subjective process of experiencing might present a case of self-experience, i.e., of a sense of self.

Does that mean that those defending the view that we possess a sense of self in terms of a substantial self, a simple, persistent and absolute entity are right? McGilvary's conclusion seems to insinuate that such a conclusion would be a mistake:

This awareness of awareness never exists except in the one indissoluble awareness of something else. We can never be aware of being aware without being aware, in the same awareness, of an 'other' than the awareness. Self-consciousness, in the sense of consciousness of I, is not an independent 'entity.'

...

If we are aware of awareness when we are aware of some 'object', we may not say that our awareness is an 'object' of itself, without using 'object' in two senses. Consciousness is conscious of itself *as consciousness*; it is conscious of its 'object' not as *consciencing*, if I may use this word, but as *conscienced*.<sup>218</sup>

Again, the mistake would consist in a confusion between the *subjective process of experiencing* and the *object of experience*. The former is a phenomenological aspect of our experience, but it is not a manifestation of an object—at least not of an object in the strict sense. No, the self-experience McGilvary has in mind is the experience of the *process* of experiencing and decisively not the experience of an alleged substantial self-object. However, although McGilvary's deliberations have been proven helpful to assess the No-Self-Picture and lay bare its weaknesses by showing that any experience of an object must involve, by the same token, an experience of the subjective process of experiencing itself, he unfortunately does not further develop a more detailed approach of self-experience. That is, it remains an open question just how exactly we are to conceive of the structure of self-experience of the continuous stream of consciousness and in which way we ought to understand the kind of sense of self he suggests. In other words, in his proposal we only find an argument as to why we should not believe our experience is merely constituted by succeeding sensory-like perceptions bereft of any further but different kind of (self-) experience of the subjective process of experiencing. What is lacking is a *positive* account of self-experience. To reject the view that there is no self-experience, it will not suffice, though, to merely provide reasons why denying self-experience is incoherent. An account of *how* self-experience, a sense of self, indeed features in our experience is required.

To offer such a positive conception of self-experience and its phenomenological structure developed along the lines of an argument astonishingly similar to McGilvary's proposal is the merit of a more recent account harboring a whole arsenal of systematically orchestrated arguments and phenomenological descriptions collected from numerous

---

<sup>218</sup> McGilvary 1907, 234.

thinkers from the phenomenological tradition: Zahavi's account of pre-reflective self-awareness or "minimal self"<sup>219</sup>. In the next section, I will present his account and discuss some recent criticisms of the phenomenological claims it propounds.

## 2.4 Zahavi with Husserl: The stream of consciousness shows up

Despite the arguments that speak against the No-Self-Picture ascribed to Hume, many philosophers seem to agree on his introspective verdict about the absence of any form of self-experience and the related phenomenological claims mentioned above. As Brie Gertler thinks: "Most philosophers find Hume's claim phenomenologically plausible."<sup>220</sup> A similar assessment has been formulated by Shoemaker who believes the No-Self-Picture ascribed to Hume has "commanded the assent of the majority of subsequent philosophers who have addressed the issue"<sup>221</sup>. I believe such an assessment to be highly biased, given not only the huge body of literature in continental philosophy, and, particularly the phenomenological tradition, which is critical of Hume's empiricist understanding of experience, but also because there are quite a few recent authors working in the field of the philosophy of mind who have indicated their discontent with Hume.<sup>222</sup>

Moreover, there is also a considerable number of authors who explicitly argue in favor of the existence of self-experience and make attempts to provide phenomenological descriptions of what constitutes our self-experience. Matt Duncan, in a recent article, for instance, emphasizes that "the self normally shows up in experience as the *author* or *agent* of one's thoughts", which he does not take to be the "most *basic*" but "one such way"<sup>223</sup> in which we have self-experience. That this is so, he argues, can be demonstrated by contrasting our experience with reports from patients with 'thought insertions' who feel some of the thoughts they have on their mind have been authored by other people. In contrast to other thoughts or to how healthy individuals normally experience their thinking, they feel forced to execute a certain thought and to be exposed to the unfolding of the thought. What they lack, thus, is a "feeling of doing", the experience that normally gives us "the sense that we are our thoughts' authors/agents"<sup>224</sup>. In a similar vein, Sebastian Watzl argues "that in conscious experience we can become aware of actively changing our focus

---

<sup>219</sup> Zahavi 1999, 2005, 2014.

<sup>220</sup> Gertler 2011, 210.

<sup>221</sup> Shoemaker 1994, 188.

<sup>222</sup> To give just a few examples: Bayne 2008, Billon 2014, Graham 2002, Guillot 2013, Madell 2015, Kriegel 2004, Musholt 2013, Schwerin 2007.

<sup>223</sup> Duncan 2019, 303.

<sup>224</sup> Duncan 2019, 305.

of attention”<sup>225</sup>. Such processes through which we focus on a certain given object we usually do not experience as something to which we are passively extradite. Although we can feel drawn into a certain event or happenings in our environment, usually, we are able to shift focus and concentrate on a given issue voluntarily. In fact, in many cases we have even trouble to resist the distractions that are omnipresent. According to Watzl, it is precisely in the experience of resisting distraction and attending on something that we develop self-experience in form of a sense of mental agency: “When we resist distractions, we are not just aware of external items, but aware of our own mental agency.”<sup>226</sup> Another example of an account of self-experience is offered by Brentyn Ramm who introduces and performs several first-person experiments to carve out the way we experience ourselves: “The goal of the experiments was to bring attention back to what it is to be the looker in your own experience.”<sup>227</sup> The results are summarized as follows:

I seem to be looking out of a gap. This open SPACE apparently lacks sensory qualities and is seemingly single. I also found that this SPACE encompasses other sensory properties; that is, it functions as a bearer of experience in general. Neither does it seem to be a mere gap, as I seem to be looking from here. It is an aware-space. As the experience meets the criteria for being a self-experience, it hence provides *prima facie* justification for the reality of the subject.<sup>228</sup>

Ramm considers self-experience to consist of an experience of a void that precisely because it is not some specific content is capable of functioning as the locus of experience, i.e., the wherein experiential content is given.

#### 2.4.1 Zahavi and the minimal self against the No-Self-Picture

All these descriptions of self-experience are astute, original, insightful, and pertinent to a comprehensive understanding of self-experience. Yet, here, in this section, my focus is on Zahavi’s account of minimal self and pre-reflective self-awareness. There are a number of reasons why his account deserves special attention when discussing positive examples of self-experience.

First, it is not only widely discussed but it is also far-reaching in scope, i.e., in terms of what it aims to explain. As Zahavi points out, there are several problems “which have to be examined if a theory of self-awareness is to prove convincing”<sup>229</sup>. That is, any theory of self-awareness or self-experience has to offer a clear and persuasive statement with regard

---

<sup>225</sup> Watzl 2018, 12.

<sup>226</sup> Watzl 2018, 11.

<sup>227</sup> Ramm 2017, 153.

<sup>228</sup> Ramm 2017, 161.

<sup>229</sup> Zahavi 1999, 38.

to a number of issues in which light any theory of self-awareness is to be evaluated.<sup>230</sup> The following questions give a taste of how well-elaborated a theory of self-experience or self-awareness is required to be: Is there a phenomenology of self-experience, i.e., can one describe self-experience phenomenologically or is it something that we are forced to posit only for theoretical reasons? How do reflection and self-experience relate? How must self-awareness be constituted so that it can provide the possibility of reflection? How is “self-awareness across temporal distance”<sup>231</sup> possible? How is self-awareness to be considered as a form of self-experience to begin with? In which sense is self-awareness egological? How does self-awareness relate to the body? How can we understand self-awareness in a fashion that does not make intersubjectivity impossible? As Zahavi emphasizes, “it will not do to conceive of the manifestation of subjectivity in such exclusive terms that it becomes incomprehensible how I should ever be able to recognize a foreign body as being in fact a foreign-embodied subjectivity”<sup>232</sup>. Moreover, the question arises as to how self-awareness and intentionality, i.e., the different kinds of consciousness of an object, are interrelated. As will turn out later in the second and third part of the dissertation, there are even more issues to consider when it comes to evaluating theories of self-experience. The only point I wish to make now is that the conception of the minimal self aims to explain how self-awareness is to be conceived in light of various philosophical issues, offers concrete analyses to this end and is therefore to be considered as probably the best candidate available for a comprehensive theory of self-awareness and self-experience.

Second, the minimal self is a conception of a very—in fact, the most—basic form of self-awareness and therefore targets any experiential episode whatsoever. As such it serves well to address the Hume-inspired no-self challenge. Although Duncan’s, Watzl’s, and Ramm’s descriptions provide examples of self-experiences that cast doubt on the verdict that our experience can exhaustively be described by sensory-like phenomenal qualities (and without any sense of self), they do not show why self-experience *must* feature in our experience. To recall, McGilvary had put emphasis on the claim that in order to experience something as continuous with a past experience, any experience of an object must include an experience of that experience, i.e., self-experience. But he did not give any further explanation of the structural intertwinement of object-experience and self-experience. The conception of the minimal self offers just that. And in so doing, it also expounds why self-experience is real whenever there is experience—even if the latter were to consist in nothing but succeeding perceptual experiences. That is, the conception of the minimal self proposes an explanation of why even experiences of allegedly merely

---

<sup>230</sup> Zahavi 1999, 39–41.

<sup>231</sup> Zahavi 1999, 39.

<sup>232</sup> Zahavi 1999, 40.

perceptual kind intrinsically involve a kind of experience that cannot be reduced to sensory-perceptual content.

Third, the conception of the minimal self not only submits an offer for how we ought to understand pre-reflective self-awareness, it is also an account of the continuity of experience. That is, it aims to explain pre-reflective self-awareness by pointing to the experience of continuity intrinsic to experience and thereby gives an account of the latter alike.

Fourth, in linking self-experience and the experience of continuity the conception of the minimal self also provides an alternative way of understanding self-experience. Instead of conceiving of self-experience in terms of a sense of a substance-like self-object, it argues that pre-reflective self-awareness, the most basic sense of self, presents a form of non-objectifying awareness.<sup>233</sup> That is, the most basic way in which we experience ourselves, according to the minimal self-account, is *not* a form of *experience of an object*. However, and that is the specific merit of the minimal self-account, it does not describe pre-reflective self-awareness, the most basic self-experience, in negative terms only. Rather, it takes self-experience to be primarily the experience of the *subjective process of experiencing* and proposes a positive description of the phenomenological structure of how *the experiential process* is phenomenologically *manifest* within experience as *self-experience*.

To sum up, thus, Zahavi's conception of the minimal self makes three phenomenological claims that stand in direct opposition to the No-Self-Picture and its three phenomenological claims:

- (I\*) *Not only sensory phenomenology*: Experience, even any period of allegedly merely perceptual kind, is never only characterized by sensory phenomenology.
- (II\*) *Continuity*: Experience unfolds phenomenologically in a continuous fashion, it flows, and is not knitted up out of single points of experience.
- (III\*) *Sense of self*: Experience always and intrinsically involves a minimal form of self-experience, i.e., pre-reflective self-awareness.

All these three claims hinge upon the phenomenological description of the structure of pre-reflective self-awareness. When Zahavi, in his first work on the issue *Self-Awareness and Alterity* (1999), developed his account of pre-reflective self-awareness, his point of departure was a discussion of the *reflection theory of self-awareness*<sup>234</sup> and its criticism

---

<sup>233</sup> Zahavi 1999, 61, Zahavi 2005, Ch.3., Zahavi 2007a, 276.

<sup>234</sup> Zahavi mentions Locke 2008, Armstrong 1993, and Rosenthal 1993, 1997. Cf. Zahavi 1999, 15ff.

through the so-called *Heidelberg School*<sup>235</sup>. The reflection theory of self-awareness argues that self-awareness is only possible if the *awareness* of an object is itself taken as an *object* of awareness. On this view, a given act of experience is not given to itself *per se* but only becomes self-aware through a second act of experience in which the first act of experience is taken as an object.

The Heidelberg School took issue with such a line of reasoning. One—if not the—main argument against the reflection theory consists in the following problem that arises if self-awareness ought to be enabled through a second act of experience (*q*) reflecting on a first act of experience (*p*). The problem is that *q* reflecting on *p* could never give rise to *self-awareness* if it didn't already take itself to be somehow *identical* with *p*. For, if *p* and *q* are considered as separate acts, it remains unclear how the awareness of *p* by *q* should be an experience of *self-awareness*. In which sense could the awareness of a different act constitute *self-awareness*? To conceive of reflection as an act of self-awareness it is required that the consciousness of the fact that *p* and *q* belong to the same experiential process must already be established.<sup>236</sup> It cannot be constituted by reflection.<sup>237</sup> In criticizing the reflection theory of self-awareness in this and other ways, proponents of the Heidelberg School—Zahavi names Henrich—have developed a list of features that original, i.e., the most basic form of self-awareness ought to possess: it ought to be *pre-reflective*, *irrelational*, *egoless*, and *private*.<sup>238</sup>

Zahavi concurs with much of the analysis of the Heidelberg School. However, on his view, the own theory of self-awareness offered by the Heidelberg School remains insufficient insofar it determines self-awareness only with features *ex negativo*. Although he acknowledges its “contribution” to a theory of self-awareness as “of decisive importance”<sup>239</sup>, he ultimately sees it falling short in the following manner:

The account offered by the Heidelberg School is insufficient, since it never explains how a subject essentially characterized by a kind of complete irrelational self-presence can simultaneously be in possession of an inner temporal articulation; how it can simultaneously be directed intentionally toward something different from itself; how it can be capable of recognizing other subjects (being acquainted with subjectivity as it is through a completely unique self-presence); how it can be in possession of a bodily exteriority; and finally how it can give rise to the self-division found in reflection.<sup>240</sup>

---

<sup>235</sup> Tugendhat 1979, 10.

<sup>236</sup> Cramer 1974, 563, cf. Zahavi 1999, 17–21. I will discuss this in more detail in Chapter 4.

<sup>237</sup> A similar line of argument can be found in McGilvary 1907.

<sup>238</sup> Zahavi 1999, 16.

<sup>239</sup> Zahavi 1999, 42.

<sup>240</sup> Zahavi 1999, 41f.

Apart from these rather specific issues, Zahavi emphasizes that the Heidelberg School fails to provide an own account of the “concrete” experiential way in which we are self-aware and only delivers “more formal considerations”<sup>241</sup>. And one of the main aims of the conception of the minimal self is to offer a plausible and concrete description of the phenomenology of the most basic form of self-experience. What then is the offer of Zahavi? What exactly is the phenomenological structure of pre-reflective self-awareness? And in which sense is it itself an experiential aspect of experience? How is *the subjective process of experiencing* itself manifest in experience?

#### 2.4.2 Husserl’s “inner time-consciousness” as the basis of the minimal self

To give an answer, Zahavi turns to Husserlian phenomenology and the latter’s analysis of inner time-consciousness.<sup>242</sup> Examining time-consciousness, Husserl was interested in the question how it is possible that we can have consciousness of an object that is only given insofar as it unfolds over time, such as a melody, which is constituted by different tones that follow each other. If all moments of the melody were only given *after* each other without any givenness of the past moments of the melody, as the No-Self-Picture would have it, perception of a melody as a temporal object would be impossible. In any moment  $t_x$  of the melody, there would always only be the exact number of tones simultaneously given at  $t_x$ , but no consciousness of the tones at  $t_{x-n}$  and  $t_{x+n}$ . But to perceive a melody means to *perceive a succession* and dynamic of tones over the course of time and not merely a series of pulse-like groups of tones. Moreover, to perceive a melody includes experiences of the following kind: being aware of how the melody started off, for instance, with a bright and serene intro, followed by a slow progression towards a dramatic climax, yielding finally a sad and dark taste at the end that is accompanied by a feeling of surprise given how the melody had started. Such experiences would be incomprehensible if there were not some form of consciousness of the passed moments including passed anticipations of how the melody further develops. The surprise in the end reveals the consciousness that the melody took a turn in a way not anticipated, i.e., a consciousness of the past moments of the melody. Zahavi concludes in agreement with Husserl: “Thus, consciousness must in some way or another transcend the punctual now, and be conscious, of that which has just been and is just about to occur.”<sup>243</sup>

From that point of departure, Husserl developed what is a complex, highly technical even pedantic analysis of the different aspects involved in the constitution of objects through time-consciousness, an analysis that had kept him busy over several decades of his

---

<sup>241</sup> Zahavi 1999, 49.

<sup>242</sup> Zahavi 2003.

<sup>243</sup> Zahavi 1999, 63.

life and that brought forth many arguments, perspectives, and more issues.<sup>244</sup> However, a number of insights into the structure of consciousness have remained more or less the same over the course of his engagement with the consciousness of time and have emerged as Husserl's view about inner time-consciousness. These are pertinent for Zahavi's understanding of pre-reflective self-awareness.

*1) Consciousness is a temporal field.*

First, instead of understanding any moment of consciousness as a *pure now-point* isolated from other now-points, Husserl suggests to conceive of consciousness rather as a *temporal field*. This field, he emphasizes, is characterized by three non-independent aspects: *primal impression*, *retention*, and *protention*. Primal impression is the consciousness of the, for instance, tones of a melody given *as now*. But the field is not exhausted by the tones given as now. Together with the primal impression, yet not given as simultaneous, any moment of experience includes a consciousness of the just past tones (or what else had just been given). That is retentional consciousness. Moreover, any moment of experience also involves a form of consciousness of what is to be expected to be given next—based on primal impression and retention. That is protentional consciousness. All three aspects together are constitutive of any given moment of consciousness. They can be distinguished in abstraction, but in concretion they are always structurally intertwined. That is, they are phenomenologically interdependent and always experienced together. It is in this sense that a pure 'now' is impossible (let alone a succession of pure now-points): If my experience were exhausted by what is given 'now', it could never be experienced '*as now*'. It is only within the temporal horizon of retentional and protentional consciousness that something can appear as given 'now'. The phenomenological character of 'now' with its intuitive fullness of impressional kind is constituted precisely against the background of an intuitively empty kind of consciousness. What has just been present to me, e.g., the tone 'c', when retained, is not given in the same manner as when the tone 'c' was present. When retained, the tone 'c' is still somehow there in consciousness but it does not resound any more. Retentional consciousness lacks the intuitive fullness that designates the impression. Protentional consciousness, in a similar fashion like retention, is an empty form of consciousness. Hearing 'c', a protention towards a further tone 'c' is built, but the tone 'c' to come is not yet there (and might not come at all, for instance, when instead of 'c' a different tone, such as 'd', follows). The sense of 'nowness' of the impression, its intuitive fullness (the full presence of the tone 'c') is experientially only there for us in contrast to the emptiness of retention and protention in which the tone 'c' is meant but not present in

---

<sup>244</sup> Here is not the place to discuss the different aspects, developments, insights or problems of Husserl's account of time-consciousness. For overview and discussion, see, for instance, Kortooms 2002, Römer 2010, de Warren 2009, Zahavi 2004, 2010.



an intuitive manner. Awareness of having just heard a ‘c’, doesn't mean to still hear a tone ‘c’, and protending the next phase of the ‘c’, is not yet hearing the next phase of ‘c’.<sup>245</sup>

2) *Consciousness as a temporal field is dynamical.*

The temporal field is not a static constellation. Rather, the structure of impression-retention-protention is dynamical and under continuous change. What is given to me as now (say perceptual content *p*) will become past in the next moment. When *p* becomes retentionally modified, its intuitive fullness is progressively emptied. Though not intuitively filled, the retention of *p* is still a form of consciousness, for its meaning, e.g., the tone ‘c’, is retained. It is only by progressive emptying that something else can be presented as now and that *p* can be retained.<sup>246</sup> Otherwise, we would, for instance, hear a succession of three tones as a triad.<sup>247</sup> Moreover, every new moment of experience does not come at surprise, as a kind of pulse out of nowhere, as it were. When *p* is given as now, it is a fulfillment of a prior protention and what was given before *p* is retained.<sup>248</sup> What is given in impressional, retentional, and protentional consciousness is constantly changing, but not in a completely arbitrary manner, but rather are all changes covarying with each other. Impressional consciousness fulfills retained protentional consciousness, while retentional consciousness empties impressional consciousness and retains it—together with preceding experiential phases. Moreover, impression and retention determine what is protended: current and past experiences motivate what is expected to happen next.<sup>249</sup>

3) *Two forms of ‘intentionalities’ constitute the full-fledged structure of intentionality*

The dynamical and constant temporal process of experiencing includes two original forms of intentionalities which are taken to be constitutive of the full-fledged structure of the intentional directedness of object-experience and which function in form of a kind of double intentionality in retentional consciousness.<sup>250</sup>

a) *The first is transverse intentionality (Querintentionalität).*

It focusses—to remain in the example of the melody—on the tone ‘c’ as given now, i.e., in primal impression and then run from there through the different retentional modifications

---

<sup>245</sup> Cf. Husserl 1966a (Hua 10), §§ 12, 13, Husserl 1966b (Hua 11), §§ 27, 28, Husserl 2001 (Hua 33), §§ 2, 3.

<sup>246</sup> For a description of the interplay of intuitive fullness and emptiness in perception, see Husserl 1966b (Hua 11), § 2.

<sup>247</sup> Husserl 1966a (Hua 10), 11.

<sup>248</sup> Cf. Husserl 2001 (Hua 33), §§ 2, 3.

<sup>249</sup> Lohmar 2002.

<sup>250</sup> Cf. Husserl 1966a (Hua 10), § 39.

of the just past phase of the sounding of 'c' and the preceding retentions of yet prior phases of 'c'. This way the tone 'c' becomes manifest as enduring, as 'there', persisting.

b) *The second is longitudinal intentionality (Längsintentionalität).*

It is directed at that what is new in primal impression as compared to what is given together with the primal impression in retentional consciousness. In doing so, I can direct my awareness on the flowing character of the subjective process of experiencing. For, given that both impressional and retentional consciousness are constantly modified, in reflection, I can become aware that what had just been in impressional consciousness now is retained and becomes even further retentionally modified when yet new phases of experience come about.

Crucially, both forms of intentionalities are possible as reflective attitudes that I can adopt and in which different aspects can become the object of my reflective attention. However, that does not mean that the kind of consciousness I so attend as such and before my reflectively attending it has a *reflective*, nor for that matter, *intentional* character (in terms of an intentional directedness at an object). In fact, both transverse and longitudinal intentionalities, according to Husserl, are the very experiential infrastructure through which a full-fledged intentional experience of an object is possible to begin with. Transverse intentionality enables the presence of an enduring object, while longitudinal intentionality constitutes the "streaming unity"<sup>251</sup> of the flow. While in reflection and abstraction these two intentionalities can be distinguished as intentionalities, pre-reflectively and experientially, they are two aspects of the very same temporal process of experiencing and its phenomenology. The experiential manifestation of an object as given, as persisting in consciousness as long as it is given, by the same token, involves a self-manifestation of the stream of the flow and its *unity in streaming*. Or, to formulate it differently, that what constitutes the duration of a tone, the presence of the object, also constitutes the temporal order of different phases of experiences in form of an experienced continuous flow. The experience of a melody (or of any other object) that stretches over a period of time cannot but register its own process as flow. Husserl puts it like this:

The flow of the consciousness that constitutes immanent time not only *exists* but is so remarkably and yet intelligibly fashioned that a self-appearance of the flow necessarily exists in it, and therefore the flow itself must necessarily be apprehensible in the flowing. The self-appearance of the flow does not require a second flow; on the contrary, it constitutes itself as a phenomenon in itself.<sup>252</sup>

---

<sup>251</sup> Zahavi 1999, 73.

<sup>252</sup> Husserl 1991, 88, Husserl 1966a (Hua 10), 83.

By contrast, putting it the other way around, if the different phases of experience weren't manifest in experience as a proceeding stream of consciousness, the argument goes, then an object like a melody could never be perceived. All what we would be left with are single tones following each other. But the specific sound image of a melody is constituted by the specific dynamical and temporal order. The sound image is a Gestalt that develops over time and can only be constituted if the different phases are *given as* being presented in a temporal order (and not only are following each other in a certain temporal order). Crucially, and to repeat, the constitution of the givenness of the temporal order is phenomenologically not distinct, i.e., from the constitution of the givenness of the temporal object. The flow of consciousness, i.e., the subjective process of experiencing, manifests itself *within* the manifestation of an object. I can focus in reflection on either of these aspects (object or process), but, pre-reflectively, they are phenomenologically interwoven. The self-manifestation of the stream of consciousness is *intrinsic* to the consciousness of objectivity. And it is in this sense that Husserl can speak of '*inner time-consciousness*'.

Now, it is this inner time-consciousness which Zahavi suggests to interpret as pre-reflective self-awareness:

Inner time-consciousness is the prereflective self-awareness of the act, and to say that the act is constituted in inner time-consciousness simply means that it is brought to givenness thanks to itself. It is called inner consciousness because it belongs intrinsically to the very structure of the act itself.<sup>253</sup>

Most importantly, Zahavi not only argues that this is the most basic phenomenological structure of self-awareness, he also suggests to interpret it as the most basic form of *self-experience* in terms of selfhood. Contrary to so-called non-egological theories of self-awareness, he explicitly takes the self-manifestation of the stream of consciousness as a form of *egological* self-awareness. For, as he argues, even though it might be true that only in reflection I become explicitly aware of myself having this or that experience *p*, becoming *explicitly* aware is not a discovery or invention of having been the one who was having *p*. Rather, it is a thematization of an implicit awareness of myself having been the one who was having *p*.<sup>254</sup>

There are more reasons he gives for why we should conceive of the self-givenness of the experiential flow as an experience of selfhood. These are multifarious and complex. They deserve special attention and I will discuss them later in a different context. For the question whether there is self-experience at all, deciding between an egological and a non-egological theory of self-awareness is not decisive. For, denying that the self-appearance

---

<sup>253</sup> Zahavi 1999, 71.

<sup>254</sup> On the notion of thematization, cf. Gurwitsch 1974, Zahavi 2005, 21, 23.

of the flow presents a case of egological self-awareness need not amount to a denial of self-experience *in toto*. One can deny that pre-reflective self-awareness is egological but still acknowledge that there are forms of self-experience, for instance, as an ingredient of higher-order acts of consciousness.<sup>255</sup> The discussion, thus, whether a non-egological or an egological theory of self-awareness is more appealing need not be settled in the present context which is concerned with the denial of self-experience altogether.

My focus here, hence, is on other approaches which have promoted no-self views by casting doubt on the phenomenology of the stream of consciousness. In the next section, I will discuss recent approaches that have been introduced as modified versions of the no-self doctrine attributed to Hume. But before doing so, let me sum up Zahavi's argument against the No-Self-Picture.

The No-Self-Picture claims that all experience is exhausted by an infinite number of impressions, that our experience are nothing over and above of bundles of perceptions, that therefore experience can be reduced to sensory phenomenology, and, that accordingly our experience lacks any sense of self. All what we find in experience are objects. The subjective process of experiencing is not manifest in experience. The minimal self account proposed by Zahavi, by contrast, argues that the subjective process of experiencing *qua* experiential flow is indeed manifest in experience and necessarily so. Only if the stream of consciousness self-appears, i.e., if experience has a streaming character can some object be experienced to begin with. *Object-experience* and the experience of the subjective process of experiencing, i.e., *self-experience*, are intertwined structures. Therefore, it is never correct to say that experience is exhausted by sensory phenomenology. Even if all our acts of experience were only to consist in sensory-like perceptions, they would always include a form of self-experience, an experience of the streaming character of the continuous experiential flow. Put differently, our experience always goes over and beyond the purely impressional givenness. Moreover, since any experience of something goes along with the self-appearance of the continuous stream of consciousness, it is incorrect to start the analysis with an alleged pure now-point isolated from other allegedly separate pulses of consciousness. For something to be given always means that it stands out in a temporal field as *given as now* rather than yet to come or just passed.<sup>256</sup>

---

<sup>255</sup> In fact, that is the position of many, e.g., Bayne 2010, Blanke & Metzinger 2009, McClelland 2019, forthcoming b, Metzinger 2003, 2009a, Peacocke 2014, Sartre 1991/1936, 1956.

<sup>256</sup> A similar point is made by Ratcliffe 2013b, when he argues that experience cannot simply be reduced to (sensory) content, given that any content also involves a sense of the possible.

## 2.5 Illusionism and recent doubts about the stream of consciousness

McGilvary, after rejecting James's idea of separated egos succeeding each other by emphasizing another Jamesian concept, the stream of consciousness, exclaimed:

Against those who would maintain that such an empirical unity of consciousness is not sufficient for philosophical purposes, I would urge that what is good enough as a correct statement of facts as they actually are, ought to be good enough for philosophy. Of course, philosophy must not stop with this single fact, but must bring it into intelligible relation with other facts which this statement does not include. When this is done, I venture to think that *the continuity of consciousness will prove to be near the center of gravity of the philosophical system.*<sup>257</sup>

Around the same time, and, as far as I know, independently, Husserl was racking his brains about how to conceive of the stream of consciousness, how its constitution is to be conceptualized, and how it stands in relation to the philosophical inquiry into consciousness. Most certainly, the stream of consciousness was at the very center of gravity of his phenomenological philosophy, the very structural ground of all objectivity and knowledge. By emphasizing the Husserlian account of the temporal constitution of the stream of consciousness, when discussing pre-reflective self-awareness, Zahavi has put back the continuity of consciousness, its nature as stream, to the center of gravity—at least when it comes to understanding self-experience and selfhood. The concept of pre-reflective self-awareness and his specific understanding of it has found quite wide-spread reception and is acknowledged by many. Those critical of the concept often at least accept the notion that the stream of consciousness presents the structure of pre-reflective self-awareness, although they are skeptical about the latter's egological character, i.e., about pre-reflective self-awareness presenting a case of self-experience. Yet, there are also quite radical criticisms of the idea of the stream of consciousness as such and of the idea of a sense of self related to the stream of consciousness. In this last section of the chapter, I will discuss three recent versions of radical criticism which reject the idea of the stream of consciousness throughout. I will first briefly sketch each of them and then assess their claims as well as their underlying assumptions together.

---

<sup>257</sup> McGilvary 1907, 231 (my italics).

### 2.5.1 Benovsky: The stream of consciousness as an illusion of the brain

Jiri Benovsky, who first seemed to endorse a view of self similar to the kind of James,<sup>258</sup> but then, recently, has moved towards the no-self view.<sup>259</sup> He suggests that while we might experience continuity in form of a stream of consciousness, the latter is only an illusion produced by our brain, created by similar interpretative mechanisms that make the perception of movement possible:

It's like in the cinema: static images are shown in rapid succession, usually twenty-four or twenty-five images per second, which creates an illusion of continuity where there only is a succession of individual static frames. Continuity—or at least the appearance of continuity—can arise from metaphysically discontinuous elements.

...

I suggest that one can think similarly of the way we experience ourselves. A 'metaphysically gappy self' only consisting in *there being successive psychological states is sufficient to account for our phenomenal experiences of continuity and unity* ...<sup>260</sup>

As the quotes demonstrate, Benovsky is not so much worried about the phenomenology of the stream of consciousness. He seems to agree that our experience appears as continuous and unified. Yet, he believes these experiences are based on a trick performed by the brain which interprets individual and really separated psychological states in such a way that they gain the phenomenal character of continuity.

### 2.5.2 Blackmore: The stream of consciousness as a delusion of consciousness

Another skepticism about the stream of consciousness is formulated by Susanne Blackmore. Rather than speaking of an illusion, she describes the idea of a stream of consciousness as a delusion because talk of the stream of consciousness is rather the product of “theorising” that “is largely intellectual”<sup>261</sup> than an illusionistic experience. Yet, she does acknowledge that our experience, at least sometimes, seems to have the structure of a stream of consciousness, namely always when we reflect on our consciousness: “I suggest that consciousness appears as a stream only when we reflect on it as such. The rest of the time, multiple parallel processes carry on, sometimes interacting with each other, often not.”<sup>262</sup> Put differently, she takes the view that our experience is stream-like as an

---

<sup>258</sup> Benovsky 2014, 2017. Benovsky here similarly like James and Strawson endorses the view that a person is constituted by an intra-individual plurality of selves.

<sup>259</sup> Benovsky 2018.

<sup>260</sup> Benovsky 2017, 550f.

<sup>261</sup> Blackmore 2016, 53.

<sup>262</sup> Blackmore 2016, 59.

artefact of introspection: “There is thus something very curious about the nature of consciousness—that looking into consciousness reveals only what it is like when we are looking into it—and most of the time we are not. So introspecting on our own minds is thwarted by the very fact of introspecting.”<sup>263</sup> The problem, she claims, is in believing our experience is stream-like we wrongly infer from how our experience appears to us in introspection to how it is when currently not introspecting. Hence, there is a wrong overgeneralization at play when we make claims about experience in general based on introspection. Equally, she maintains, there is usually no unity in consciousness. Unity of consciousness only occurs when we attend certain objects:

It can certainly seem that way, as if the sensations of sitting here, the pictures on the wall in front of me, the sounds around me, all form some kind of unified field. ... Yet, it takes only a little careful observation to see that as my attention shifts these sensations come and go and that these acts of attention pull some together and let others disappear. This is the only sense in which there is ‘unity of consciousness’.<sup>264</sup>

Blackmore, thus, takes the experience of both continuity and unity to be tied to introspection and attention, something that comes about only when these acts of focus are at play. Really, that is. when we do not introspect nor attend any particular object, our consciousness, she seems to suggest, is a more or less organized chaos of “multiple parallel processes with no clear distinction between conscious and unconscious ones”<sup>265</sup>.

### 2.5.3 Garfield: There is no phenomenal consciousness

The most radical criticism of the idea of the stream of consciousness, however, can be found in the approach of Jay Lazar Garfield, it seems. Instead of denying that our experience has a certain character, a stream-like character, he does not make prisoners in claiming: “There is no phenomenal consciousness; there is nothing ‘that it is like’ to be me.”<sup>266</sup> If phenomenal consciousness *in toto* does not exist, then, clearly, it makes no sense to speak of a stream of consciousness, and consequently it cannot function as a sense of self. But is Garfield really denying phenomenal consciousness?<sup>267</sup> A closer look reveals that it is rather a peculiar notion of phenomenal consciousness that he dismisses:

---

<sup>263</sup> Blackmore 2016, 54.

<sup>264</sup> Blackmore 2016, 60.

<sup>265</sup> Blackmore 2016, 52.

<sup>266</sup> Garfield 2016, 73.

<sup>267</sup> Cf. Niikawa 2021. It could be argued that the disagreement between illusionists and anti-illusionists is merely terminological.

Once again, it is important to be clear about what phenomena we do know in perceptual experience. First and foremost, we know the objects around us and their perceptible properties—a mango and its yellowness and sweetness; middle C on the oboe and its timbre; sandalwood and its smell, etc. When we introspect to find the experiences that correspond to the awareness of the contents of perception, we can discover *that we are aware of a sweet, yellow mango, the sound of middle C, or the scent of sandalwood*, but nothing more is added; we discover not two objects of knowledge—the external and the internal medium that makes the external object known—but only the object and the fact that we are aware of it.

...

Whence the illusion that there is not only something that the world is like but also something that it is like to experience the way the world is? The source, at least at a first pass, is the failure to distinguish the subjective from the objective side of experience.<sup>268</sup>

As the quotation shows, Garfield does not deny that the world as it is like can be given to us. What he denies, however, is that apart from the qualities or properties of the objects of the world there are also some additional features of experience. Similar to the No-Self-Picture, thus, he maintains all there is to experience is object-experience. Accordingly, when he rejects phenomenal consciousness, what he means is the *subjective process of experiencing*. On his view, it does not show up, nor for that matter does he believe it has any feature or structure.

#### 2.5.4 Response to the skeptics

Do these different objections invoke severe worries for the defenders of the idea of the stream of consciousness? I don't think so. While they are creative in formulating strong claims about consciousness that sound provoking, admittedly interesting, and revelatory, on a second look, they do not reach far beyond what the No-Self-Picture offers. Some of their descriptions and assumptions are even evidently self-contradictory or otherwise problematic.

When Benovsky claims that the experience of continuity, the stream-like character, is the product of a mechanism similar to watching TV, then there is something deeply wrong. First, one may ask how he ends up with the assumption of isolated atomistic 'individual psychological states' that are to play the role of the individual frames needed to produce a film. Clearly, it is not the phenomenology of experience that suggests so, for he concedes it has stream-like character. Accordingly, to speak of individual psychological states seems like a secondary abstraction, most certainly a product of an ill-understood application of reflection. As discussed above, in reflection one correctly distinguishes between the reflecting from the experience reflected on and then wrongly separates both

---

<sup>268</sup> Garfield 2016, 78.



from each other as allegedly isolated and independent acts. But the mistake is to oversee that such a reflection in which certain aspects of experience can be singled out as states is only possible because experience has stream-like, i.e., continuous character. The distinction between 'individual psychological states' is secondary *to* and an abstraction *from* the primary phenomenon: the stream of consciousness. It therefore seems wrong to attempt at explaining the continuity by applying a gluing-mechanism of whatever kind to turn allegedly static mental pictures into a moving stream.

In addition, the mechanism or the metaphor for the mechanism of stream-production Benovsky propagates is far from providing anything near a plausible picture of how the stream of consciousness comes about. One must ask how it is possible in the case of film-production that the presentation of single pictures may give rise to movement? Why do we perceive movement in our environment, say when I drop the cup and it falls on the floor? Would this sort of perception be possible without experience having a stream-like character? I don't think so. It is precisely because of retentional consciousness that I can see the cup falling, because its prior positions are still retained in consciousness along with constantly fulfilled protentions of the next positions of the cup. It is rather because of the fact of the continuity of consciousness that I can be tricked into thinking that I perceive movement where in fact is only a succession of static pictures. Equally, it would require some sort of stream in which the alleged mental pictures were to be given so that a stream of consciousness could be generated. But it is clear that such a description cannot possibly explain the stream of consciousness, because it is circular. Moreover, it also suggests a kind of stream-homunculus in the brain that "looks" at mental pictures, the 'individual psychological states'. Yet, it seems hard to deny that such an implication amounts to a *reductio ad absurdum* of the camera-metaphor as an explanation of the stream of consciousness rather than an attractive version of a theory of consciousness.

Blackmore's suggestions and the problems I see in her argumentation also concern the understanding of reflection. She maintains that we come to believe in a stream-like character of experience based on reflection. And she accepts that in reflection experience indeed seems to unfold stream-like. Yet, she denies that experience flows also when we do not reflect. Let me start by granting that there is a phenomenological difference between pre-reflective and reflective experiencing. It makes a difference whether I am simply driving a car and am concerned about traffic as compared to reflecting on the very same experience. It feels different. Let me also grant that experience in some forms of reflection has a particularly strong stream-like character. If I artificially hold still and then slowly turn my head from left to right and focus on my visual perception, I notice how my visual field slowly changes. I can track that the bottle that is directly in front of me, after a while, is at the left of my visual field. The slowly getting from the center to the periphery of my visual field of the bottle unfolds temporally and slowly. Clearly, this reflective experience has a

strong streaming character. Even more so, if I focus either on the transverse or longitudinal aspects of retentional consciousness I directly focus on the streaming character of experience. Most certainly, concerned with everyday life issues, my focus will be different.

Consider, for instance, while slowly turning my head from left to right, my girlfriend calls and asks whether I have done what she asked me to do. I haven't. So, I get stressed, because I know it mattered much to her. And my thinking becomes nearby chaotic. Will she forgive me? Can I still do it? What time is it? How can I do it? What else did I forget? Am I unreliable? I can't think right and I walk through the apartment trying to get dressed while brushing my teeth. I don't think things through. I don't answer the questions that pop up in my mind in a staccato-like manner and linger in and around me, and create an unpleasant tension in my head, my whole body. I hit the table with my knee, the bottle drops. Bam!

Evidently, such an experiential episode is phenomenologically quite different from the calm reflection in which I focus only on my visual field and the bottle. And the stress episode, one may indeed say, does not seem to have the structure of a well-balanced stream as in the reflective episode. But I believe it wrong to infer from that that only the latter has a stream-like character. First of all, it is unclear to me how Blackmore thinks reflection or attention bring about continuity and unity in otherwise chaotic and non-temporally structured impressional data. It might be true that at the periphery of my experience, at the periphery of my experiential field, things are less clear than in the center of my attention. I hear 'something' out there, a noise, but it could be anything. At the same time, I hear the humming of the oven, I know it is the oven while I look at the cake being baked inside it. So yes, there is a difference in clarity. At the periphery of my experiential field things seem wilder, less organized, more degrees of freedom, and hard to predict. But Blackmore seems to imply that at the periphery of reflection and attention, there are clearly given chaotic impressional data, that become tamed only through reflecting and attending them. But I don't think that is what we experience. Consider again the difference between simply driving the car and reflecting on it, but this time, focus on what you think both experiences share in structure. Do you think reflecting on the driving experience alters it from a discontinuous chaotic process to a stream-like experience? Of course, Blackmore could say that, in reflection, it might seem that both pre-reflective and reflective experience of driving are stream-like, and that *that* is precisely the bias and artefact of reflection. But then she must also say that whenever we introspect or reflect on our experiences, we do not really get a grasp of what it is like to have them before we reflected on them. But if that were the case, then how can she know that before reflection and attention experiences were not stream-like and discontinuous? Isn't it through some form of reflection? What special access to experience does she possess that we don't? Hence, I believe Blackmore's approach not only has a wrong picture of the phenomenology of experience, fails to provide

an explanation how reflection ought to inflict a stream-like structure on otherwise unorganized data, but also oversteps her own epistemological boundaries. Finally, her theory also does not—nor for that matter seems to be apt to—give an explanation of how reflection is possible in the first place.

Garfield's attack on the stream of consciousness, as I mentioned above, is mostly due to the belief that the subjective process of experiencing is not manifest in experience. According to Garfield, all what is manifest in experience are the perceptual objects. His position, thus, although *prima facie* very radical and specific, is very similar to the No-Self-Picture. Accordingly, I suggest that the arguments I put forward against the latter also apply, by the same token, to Garfield's account. Let me, therefore, instead of repeating them, bring a different argument against Garfield that also further buttresses the rejection of the phenomenological claims of the No-Self-Picture.

As Zahavi, correctly in my view, points out, Garfield's view is also very similar to recent transparency views of consciousness.<sup>269</sup> On the transparency thesis, phenomenal consciousness is exhausted by properties of the object.<sup>270</sup> The view is exemplified by the following quotation by Tye (and which Zahavi discusses<sup>271</sup>):

Try to focus your attention on some intrinsic feature of the experience that distinguishes it from other experiences, something other than what it is an experience of. The task seems impossible: one's awareness seems always to slip through the experience to blueness and squareness, as instantiated together in an external object. In turning one's mind inward to attend to the experience, one seems to end up concentrating on what is outside again, on external features or properties.<sup>272</sup>

Again, just like in the No-Self-Picture, it is maintained that experience can be reduced to perceptual content. However, as Zahavi emphasizes, our experience is not only concerned with the object and its properties, it also includes an awareness of how the object is given to us:

If you compare a situation where you perceive a green mango with a situation where you remember a green mango or imagine a green mango, there will be an experiential difference between these different situations, and it is hard to

---

<sup>269</sup> Zahavi 2020a.

<sup>270</sup> For a concise description of the transparency thesis, see Montague 2016, 70ff. Note that the so-called transparency thesis is to be distinguished from the transparency of the self-model described by Metzinger and which I discussed in Section 1.4. Metzinger doesn't emphasize that our experience of the world is exhausted by phenomenal properties of objects. After all, he acknowledges phenomenal selfhood. What he means with transparency refers to the representational (i.e., brain) processes that underpin our experience and that do not show up in experience. These processes are transparent. Yet, Metzinger does acknowledge self-awareness as a phenomenon of our experience of the world. Hence, his transparency view is quite another matter.

<sup>271</sup> Zahavi 2020a.

<sup>272</sup> Tye 1995, 30.

see how that could be explained simply by referring to the objectual properties of the green mango.<sup>273</sup>

I believe that argument is correct. Referring to object properties only is insufficient to distinguish between different kinds of experiences of the same object. Yet, it *is* possible to differentiate between different kinds of *acts* of experience. And such differences are phenomenological, they are experientially manifest. The subjective process of experiencing is manifest in experience, since only then can we distinguish between the different kinds of acts in which an object is given to us. That applies equally to the No-Self-Picture. Experience cannot only consist of sensory-like perceptual content and nothing else, for there are quite different ways in which we are intentionally directed at the object that is given to us. And, we are usually aware of the kind of intentional directedness currently in play, an awareness that cannot be explained by objective properties, since these properties are the same across the different kinds of intentionalities. Kant's famous 100 thaler have the same object properties, whether I perceive them in my pocket or whether I fantasize about having them or whether I remember how I had them.<sup>274</sup>

## 2.6 Conclusion: The stream of consciousness against the No-Self-Picture

In this chapter, I presented and discussed a second form of skepticism about self-experience: the view that our experience does not include any sense of self, i.e., the view that there is no self-experience altogether. I started by a discussion of the so-called no-self doctrine attributed to Hume and whether his position indeed was that selfhood is unreal or whether he simply rejected a certain understanding of self: self as a substance-like object as the metaphysical bearer of experience. I then characterized the No-Self-Picture that has been attributed to him by three phenomenological claims:

- first, that our experience is exhausted by sensory phenomenology;
- second, that our experience consists of bundles of sensory-like perceptions;
- third, that our experience is bereft of any sense of self.

I discussed these claims in light of McGilvary's criticism of James's account of separated thinkers and by focusing on the assumption that our experience is constituted of bundles. I argued that the bundle view of experience cannot account for the phenomenology of

---

<sup>273</sup> Zahavi 2020a.

<sup>274</sup> Kant 1996/1781, 584 (B 627).

experience, and that the latter not only includes object-experience but also the subjective process of experiencing. For reflection or Humean introspection to be possible, I emphasized, the subjective process of experiencing must itself be manifest in experience. The continuity of experience is manifest as an experience of continuity. To give a positive description of the phenomenological structure of how the subjective process of experiencing is manifest in experience, I built on Dan Zahavi's theory of pre-reflective self-awareness or minimal self which suggests to interpret Husserl's notion of inner time-consciousness as the phenomenological structure of the way in which the subjective process of experiencing is experientially manifest. I also followed Zahavi in understanding the self-manifestation of the experiential flow, the stream of consciousness, as the most basic form of self-experience. As I further showed, Zahavi's conception of the minimal self promotes three phenomenological claims that are directly contrary to the No-Self-Picture:

- first, that not even allegedly mere perceptual experience is ever exhausted by sensory phenomenology;
- second, that experience is characterized by continuity, i.e., that it has stream-like character;
- third, that experience involves pre-reflective self-awareness understood as the most basic form of self-experience or sense of self.

Finally, I presented and discussed three recent criticisms of the idea of the stream of consciousness, and concluded that their arguments are insufficient to undermine the notion of the stream of consciousness.

The conclusion that I draw from my discussion of the No-Self-Picture and the experiential manifestation of the conscious flow is that the second form of skepticism about self-experience could only possibly be convincing if there was agreement on the fact that self-experience would have to consist in the experience of myself as a substantial self or some other form of self-object. However, as my discussion further illustrated, an agreement of that question is hardly in sight. Quite the contrary, having presented some defenders of self-experience, notably, Zahavi's interpretation of the temporal self-manifestation of the conscious flow as a case of basic self-experience, it became clear that skepticism of the second form is insufficient to dissolve talk of self-experience.

What I have not yet sufficiently addressed is the question for what reasons exactly we are able to understand the self-manifestation of the experiential flow as the most basic form of self-experience. That question, which has recently attracted increasing attention in the current debate, I will address in Chapter 4. As it turns out, a more refined *metaphysical* discussion will be required to fully elucidate the phenomenological claims concerning the most basic form of self-experience. But before I can devote myself to that task, a third kind

of skepticism about self-experience needs to be discussed, which I will do in the chapter to come.

# Chapter 3

## Elusiveness, Illusion, and the Limits of Self-Experience

[W]e are not unaware of what we are except in the inability to characterise it beyond each one's apprehension of himself in being himself.<sup>275</sup>

We call the substrate [i.e., the brain, P.S.] of the conscious appearances 'I' for no other reason than the fact that the substrate is deemed to underlie the immediately experienced 'I', which, in the first instance, exclusively owns the right to carry the name 'I'. Without the immediately experienced 'I' the real 'I'—the substrate—would be nothing more than a nameless something.<sup>276</sup>

In the previous chapters, I discussed two radical forms of skepticism about self-experience. The first consisted in the idea that, whatever self-experience we might have, it is throughout hallucinatory to the effect that selfhood is not real. The second consisted in the idea that there is no self-experience that could motivate the notion of selfhood to begin with. In this chapter, I address a third—less radical—possibility to cast doubt on self-experience. Characteristic of this possibility is that it is compatible with the intuition of most people about themselves: that they exist. Unlike the first two forms of skepticism about self-experience, thus, the third does not present a version of the no-self doctrine. It does not deny selfhood. It accepts both that we have self-experience and that such self-experience indeed indicates the reality of selfhood. However, it emphasizes worries regarding the question as to what extent one may gain insights into the metaphysical nature of selfhood based on self-experience. The third form of skepticism denies that self-experience may provide *self-knowledge*. All self-experience can provide, on such a line of thinking, is knowledge about how we *experience* ourselves. Accordingly, it suggests there is a gulf between how selfhood *appears* to itself and what selfhood really *is*. Hence, while it accepts that self-experience reveals the reality or *existence* of selfhood, it denies that self-experience is revelatory with regard to the *being* of selfhood. In self-experience, on this view, we learn *that we are* but not *what we are*.

---

<sup>275</sup> Lewis 1982, 172.

<sup>276</sup> Lipps 1901, 41, my translation: “So nennen wir auch das Substrat der Bewusstseinserscheinungen ‘Ich’ aus keinem anderen Grunde, als weil es das dem unmittelbar erlebten Ich, dem zunächst allein dieser Name zukommt, zu Grunde Gelegte ist. Ohne dies wäre das reale Ich nichts als ein namensloses Etwas.”

Roughly, there are two views that may give rise to or cash out the skeptical attitude of this third kind. Usually, the second view involves the first, while the second is not necessary for the first:

### 1) *Elusiveness*

On this view, self-experience is our best candidate to investigate selfhood, but, unfortunately, even the best candidate is limited and can only bring into sight selfhood *insofar as it is experienced* in the context of an experience of some object *x*—but not how it is considered *independently of our experience of x*. Given that selfhood, as it is in itself, i.e., without being concerned with some *x*, necessarily evades our experiential gaze, our true metaphysical nature of selfhood is *elusive*. The classical and most relevant approach to date is Kant’s theory of self.

### 2) *Illusiveness*

On this view, self-experience is veridical in the sense that it correctly provides a sense of one’s own reality or existence. Self-experience has a real referent: e.g., the brain.<sup>277</sup> Yet, given that in self-experience we don’t experience ourselves as brains, we must conclude that self-experience “systematically and dramatically misrepresent[s] what features the self has”<sup>278</sup>. Hence, it is not that we couldn’t gain insight into the true metaphysical nature of selfhood. It just so happens that self-experience is not a good guide for such an attempt at all, because the true nature of selfhood is elusive to self-experience. But even worse, self-experience purports to present us with the being of selfhood when it doesn’t. Self-experience is *illusory*. Accordingly, not only does it need to be complemented but also corrected by other means, for instance, the natural sciences. This, at least, is the suggestion by what Tom McClelland describes as the *illusion model*.<sup>279</sup> Like the *hallucination model* discussed in the first chapter, the illusion model understands self-experience (or self-representation) as a product of the brain. Unlike the hallucination model, however, it sees no reason to deny the brain the status of self and so acknowledges the reality of selfhood.

In this chapter, I discuss different approaches that involve one or the other of the two views and the challenges they face. Most importantly, while these views come dressed as skeptical positions, a closer examination reveals that they involve and reproduce essential elements of the classical soul-theory they usually purport to undermine. As it will turn out, thus, behind the alleged skepticism protrudes a metaphysical stance that requires further

---

<sup>277</sup> Howell 2010, McClelland 2019.

<sup>278</sup> McClelland 2019, 28.

<sup>279</sup> McClelland 2019, 28. Cf. Howell 2010.



argumentation. However, given the problems that are generally attached to both views and specifically to those versions I discuss, I will conclude that this third form of skepticism about self-experience doesn't offer an attractive line of thinking.

### 3.1 Elusiveness I: Kant and the indeterminability of the pure 'I think'

Kant's approach is a complex and comprehensive one that involves many specifics and genuine features. Crucially, it is also open to interpretation and has been subject to controversy for more than 200 years. However, here, for the purpose to show what is wrong about elusiveness, it is sufficient to focus on Kant's stance on self-knowledge, i.e., on the question what we can know about the self. Kant's answer to this question is far less ambiguous than other components of his philosophical system and there is—relatively speaking—a consensus at least with regard to the general line of his argument and his view that based on self-experience we cannot determine the metaphysical nature of self as it is in itself.

I will now reconstruct the argument and further detail the metaphysical assumptions that underlie it. I will then describe how these metaphysical assumptions are problematic and examine whether the problems they produce are bound to the specifics of the Kantian approach. That is, whether the problems could be avoided by modifying some of the Kantian specifics. I will conclude the section by denying the latter and arguing that the problems apply to any approach that takes some alleged true metaphysical nature of self to lie behind or to remain unwitnessed beyond our experience, inaccessible as it were.

#### 3.1.1 Kant on self-knowledge

To understand Kant's argument, it is necessary to bear in mind how he conceives of knowledge and the structure that enables knowledge in the first place. Moreover, it is important to consider the context in which Kant addresses and examines the possibilities of self-knowledge. Let me start with the former.

Knowledge, for Kant, always begins with experience, to the effect that any claim of knowledge must be related to our experience.<sup>280</sup> Now there are two general ways of that relationship through which our knowledge might be said to refer to experience. Either knowledge is based on something *given* in experience or it *makes* experience as such

---

<sup>280</sup> Kant 1996/1781, 44 (B 1).

possible.<sup>281</sup> Kant calls the first kind of knowledge empirical, the second kind transcendental knowledge.<sup>282</sup>

For something to be given in experience, and, accordingly, knowable through experience, Kant is adamant to highlight, it is required that what he denotes as the two sources of knowledge, reason and sensibility, are brought together in a synthesis. To specify, what is required for an experience of *x* is that a concept is applied in the structuring of the intuitive manifold given to us through our senses: “Thoughts without content are empty, intuitions without concepts are blind. Hence it is just as necessary that we make our concepts sensible (i.e., that we add the object to them in intuition) as it is necessary that we make our intuitions understandable (i.e., that we bring them under concepts).”<sup>283</sup> Experience and knowledge, then, always refer to an object, which “is that in whose concept the manifold of a given intuition is *united*”<sup>284</sup>. Knowledge is object knowledge insofar concept and intuition are united in an object. For Kant, it is the objectivity of knowledge, thus, what makes knowledge knowledge in the first place.

Transcendental knowledge refers to those structures that enable the givenness of an object as such, i.e., objective knowledge alike. As is well-known, to demonstrate which structures are required for the givenness of an object as such and what role they play is one of the main tasks of the *Critique of Pure Reason*. Over the course of the *Critique*, Kant presents space and time as pure forms of intuition on the side of our sensibility, and the categories as pure concepts of reason on the side of reason. Categories are predicates of objects as such, regardless of their intuitive content. They constitute the general components of any judgement whatsoever. But Kant’s analysis, as is well known, goes further. He asks what else is needed so that judgments can be made at all. Given that in a judgement a predicate is attributed to an object, what is required, is a synthesis. The question, then, is what kind of synthetic activity is necessary so that categories can be applied?

It is here where Kant addresses the question of the self the first time. The main context, in which Kant refers to the self and self-knowledge in his theoretical philosophy, is the question how knowledge as such is possible. As Frank points out, Kant’s interest in the ‘I’ or ‘self’ was wholly motivated by the ambition to provide an explanation of how the objectivity of knowledge as such is possible.<sup>285</sup> It is only because the ‘I’, the ‘I think’ to be more precise, has a logical function in the constitution of objective knowledge that Kant allows the ‘I’ such a central place in his critical system.

---

<sup>281</sup> Kant 1996/1781, 45 (B 2), 64 (B 25).

<sup>282</sup> Kant 1996/1781, 45 (B 2), 64 (B 25).

<sup>283</sup> Kant 1996/1781, 107 (B 75).

<sup>284</sup> Kant 1996/1781, 181 (B 137).

<sup>285</sup> Frank 1991b, 418.

But what exactly is the logical function of the ‘I think’? The ‘I think’ is precisely the answer (or part thereof) to the question what kind of synthetic activity is needed to unify conceptual features with intuitive content in the experience of an object. In other words, it is the ‘I think’ that enables the synthesis of the two components that are linked in a judgement: an object and its features. It is the ‘I think’ that enables the application of the categories to intuitive content.

How does the ‘I think’ fulfill that function? It does so by constituting the unity of consciousness. Put differently, it is by the fact that all what is consciously given to me could be accompanied by the ‘I think’ that what is consciously given to me is given to *me*. The ‘I think’ relates all consciously given content to my consciousness. And only insofar elements of different conscious contents are possibly given in the same consciousness is a synthesis of different features in a concept possible. And only then can an object be given to consciousness in the first place. Similarly, only if the different components of a judgment are present in the same consciousness, i.e., if they are unified in one single consciousness, is a judgment possible to begin with. By unifying consciousness, the ‘I think’ enables the conscious givenness of an object as such, and the possibility of the knowledge of an object alike:

The synthetic unity of consciousness is, therefore, an objective condition of all cognition. Not only do I myself need this condition in order to cognize an object, but every intuition must be subject to it *in order to become an object for me*. For otherwise, and without that synthesis, the manifold would *not* unite in one consciousness.<sup>286</sup>

It is because of the unifying function of the ‘I think’ that Kant also speaks of the “objective unity of self-consciousness”<sup>287</sup>. Objectivity, i.e., experience of an object and objective knowledge, which consists of judgements, requires the synthetic activity of the ‘I think’. No experience and no knowledge, then, without the subject. Whatever object may present itself to me, the ‘I think’ as a structural and logical presupposition “must be *capable* of accompanying all my presentations”<sup>288</sup>. Kant describes the ‘I think’, i.e., self-consciousness, therefore as transcendental apperception. It is *transcendental* insofar as it is a necessary condition of the possibility of experience. It is *apperception* (ad-perception or auto-perception), because any experience of *x* includes the perception of *x* being given to *my* consciousness.<sup>289</sup>

Having laid out the general constraints for any knowledge whatsoever in Kant, and the context in which Kant examines the self, we may now ask what exactly it is that,

---

<sup>286</sup> Kant 1996/1781, 181 (B 138).

<sup>287</sup> Kant 1996/1781, 182 (B139).

<sup>288</sup> Kant 1996/1781, 177 (B 131).

<sup>289</sup> Frank 1991b, 425.

according to Kant, we can possibly know about the self, the consciousness of which seems certain in any moment of consciousness. The question is, therefore, how did Kant conceive of self-knowledge?

One might wonder whether Kant has developed a genuine theory of self-knowledge as a genuine class of knowledge. As has been pointed out in Kantian scholarship, Kant had indeed tried to deliver such but came to realize that his attempt had failed.<sup>290</sup> As a result, Kant conceptualized self-knowledge as a special case of object-knowledge. On this view, the difference between knowing an *object x*, for instance, the notorious table in front of me, and *self-knowledge* only consists in the different objects in question. The general constraints of knowledge as knowledge of an object apply equally to our knowledge of self.

But what, according to Kant, can we say about the self as an object of knowledge? Importantly, Kant highlights that it is “inner sense” through which we perceive ourselves, in an analogue way to how we perceive other objects that appear in space and time by affecting our “outer sense”. To specify, in the case of the latter, things, via the different sensory modalities, affect our perceptual system, which, led by the synthetic activity of pure reason, induces the experience of an object. On Kant’s view, by the same token, our inner sense is affected, not by external things, but by ourselves. For, as Kant’s theory of experience goes, it is only by the synthetic activity of the ‘I think’, the subject’s activity, that the intuitive manifold given to our outer sense may be organized as the experience of an object. In Kant’s words:

[B]y no means does the understanding already find in inner sense such a combination of the manifold; rather, the understanding *produces it*, inasmuch as the understanding *affects* that sense.<sup>291</sup>

The application of categories to intuitive content through the synthetic activity of the subject affects the inner sense, for which reason Kant speaks of self-affection.<sup>292</sup> What Kant, therefore, is saying, I suggest, is that in seeing the table as an object in front of me, I am at the same time experiencing the subject’s synthetic activity. In other words, the fact that an object is given to me bears witness of the unification of consciousness through the subject, for without the latter no object could be experienced in the first place. Object-givenness, then, includes the experience of the subject’s activity.

However, on Kantian terms, the experience of the subject’s activity should not be understood as a *direct* experience of the subject’s activity. We only experience the subject’s activity, i.e., ourselves as intellectual beings, only insofar as it affects the inner sense.

---

<sup>290</sup> Cf. Washburn 1976.

<sup>291</sup> Kant 1996/1781, 193 (B 155).

<sup>292</sup> For a discussion of the role of self-affection in Kant’s theory of experience, see Indregard 2017.

Therefore, we only experience ourselves as we *appear* to ourselves as determined by the conditions of sensibility:

Whatever is presented through a sense is, to that extent, always appearance. Hence either we must not grant that there is an inner sense at all; or we must grant that the subject who is the object of this sense can be presented through it only as appearance, and not as he would judge himself if his intuition were self-activity only, i.e., if it were intellectual intuition.<sup>293</sup>

The crucial point is that, for Kant, the inner sense is structured in a similar way like the outer sense. That is, something—the thing in itself—affects the sense and induces an appearance. The scope of our knowledge about that thing is restricted to how it appears to us and does not include the thing in itself as it *is* independently of our perceiving. That is, in seeing the table in front of me, I may gain knowledge about its perceptual properties and causal connections to other entities of the perceivable world. Yet, what the table is beyond my experience of it I cannot *know*. According to Kant, from an epistemological point of view, I am only allowed to *think* that there is something that subsists my experience of the table.

Similarly, the scope of my self-knowledge is restricted to how I appear to myself in inner sense: my self-appearance in inner sense does not present my ‘self’ as it is independently of my experience of the outer world. I cannot have any knowledge of my being as it is independently of my experience. The general Kantian constraints of knowledge apply to the self. Hence, any claim of knowledge must be legitimized by a synthesis of concept and intuition.

But in the case of the self, such a synthesis is not possible. For, we precisely lack the corresponding intuition of the self. The intuition of self that is available is not pure and mixed up with our intuitions of the objects in the world. In perceiving empirical objects, we may have an indirect and implicit intuition of the activity of the ‘I’ insofar as it grants the objective unity of perceptions, but apart from the implicit manifestation of self in our intuitions of the world we are not capable of intuiting the synthetic activity of our subjectivity. Bereft of any empirical content given to us through our senses, no intuition, an *intellectual* intuition based only on pure thinking as it were, is impossible for us. We are always only aware of our own subjective activity insofar as our subjective activity is directed at something passively given to us.

That is why Kant believes that a pure intuition of self-activity, independently of the extended world, remains a Cartesian fantasy. Notably, in the ‘Refutation of Idealism’ Kant maintains that the Cartesian ‘I think’ is an empirical sentence always already imbued with intuitive content perceived through the outer sense, which is therefore primary to the inner

---

<sup>293</sup> Kant 1996/1781, 100 (B 68).

sense.<sup>294</sup> Moreover, it means that our self-consciousness is not to be conflated with self-knowledge. Just like in the case of perceptual objects of the world, the notion of a division between *appearance* and *reality* applies equally to the self.<sup>295</sup> Hence, our self-consciousness indeed indicates our existence and we are aware of ourselves as existing beings.<sup>296</sup> But we just cannot further determine the nature of the being of self as it is in itself. That is, the object of self-consciousness “is not the consciousness of the *determining* self, but only that of the *determinable* self, i.e., of my inner intuition”<sup>297</sup>.

Moreover, according to Kant, the problem is not only that we lack a pure intuition of our subjective activity. He also utters some doubts about the ‘I’ being a concept: “[W]e cannot even say it is a concept, but only that it is a mere consciousness accompanying all concepts.”<sup>298</sup> It is a “presentation” that is “by itself quite empty” and through which “nothing more is presented than a transcendental subject of thoughts = x”<sup>299</sup>. Independently of thoughts, there is nothing we can possibly say about its nature: “This subject is cognized only through the thoughts that are its predicates, and apart from them we can never have the least concept of it . . . .”<sup>300</sup> Moreover, such thoughts, if they are to present any self as an object at all, need to involve some intuitive, i.e., empirical material. For, “I do not cognize any object by merely thinking, but I can cognize any object only by determining a given intuition with respect to the unity of consciousness in which all thought consists”<sup>301</sup>. The ‘I think’, taken as a single and isolated thought or presentation is nothing but a logical function that does, as such, not present any object. Rather, it is the possibility of any object presentation. It is formally required to apply the category of ‘substance’ in the first place and therefore cannot indicate an object itself.<sup>302</sup>

To sum up, we have no *intuition* of ourselves as thinking subjects independently of our experiences; and, we do not even have any *concept* of ourselves as thinking subjects independently of our thoughts, which, as human beings, are never pure but always already entrenched with experiential material. Moreover, even if one were to analyze the pure ‘I think’ one would find no way to derive any knowledge about the self as an object

---

<sup>294</sup> Kant 1996/1781, 288ff. (B 274ff.), 435 (B 422).

<sup>295</sup> Kant 1996/1781, 194f. (B 156f.): “Hence if concerning the determinations of the outer senses we grant that we cognize objects through them only insofar as we are outwardly affected, then we must also concede concerning inner sense that we intuit ourselves through it only as we are inwardly affected *by ourselves*, i.e., we must concede that, as far as inner intuition is concerned, our own (self as) subject is cognized by us only as appearance, but not in terms of what it is in itself.”

<sup>296</sup> Kant 1996/1781, 195 (B 157).

<sup>297</sup> Kant 1996/1781, 425 (B 407).

<sup>298</sup> Kant 1996/1781, 385 (B 404).

<sup>299</sup> Kant 1996/1781, 385 (B 404).

<sup>300</sup> Kant 1996/1781, 385 (B 404).

<sup>301</sup> Kant 1996/1781, 425 (B 406).

<sup>302</sup> Kant 1996/1781, 435 (B 422).

independent of experience. Hence, independently of experience, nothing can be really known about the subject.

For, all knowledge about ourselves is bound to how we appear to ourselves in inner intuition. To draw any conclusions about the self as it is in itself from how we appear to ourselves or from the structure of thought, for Kant, would be a mistake, as he is adamant to point out against the rational psychologists of his time. As *rational* psychologists, they had tried to show the substantiality, immateriality, simplicity and persistence of the soul by an investigation of the pure structure of thought, i.e., the ‘I think’, bereft of any concrete empirical content. But as Kant concludes from his analysis of the elusiveness of self:

Hence analyzing the consciousness of myself in thought as such does not yield the slightest gain as regards the cognition of myself as object. The logical exposition of thought as such is wrongly considered to be a metaphysical determination of the object.<sup>303</sup>

To demonstrate that the mere structure of thinking does not suffice to determine the metaphysical nature of selfhood, Kant, artificially, considers the case in which pure thinking could become conscious of itself, i.e., ‘appear’ to itself *as* mere thinking bereft of any concrete empirical content. Now the question is *what* exactly would an intuition of pure thinking, if such were possible, include? If in this case, thinking were given to itself as an enduring, simple substance independent of all material, then this, according to Kant, would present a—really, the only—counter argument against his whole doctrine of the elusiveness of self:

It would be for our entire critique a great stumbling-block—indeed, even the only one—if there were a possibility of proving a priori that all thinking beings are in themselves simple substances; and that, being such, they therefore (as a consequence from the same basis of proof) inseparably carry with them personality and *are conscious of their existence as one that is set apart from all matter*.<sup>304</sup>

But Kant further elaborates on that merely hypothetical possibility. Even if such consciousness of pure thinking of itself were possible, he asks, with what kind of predicates could thinking determine itself? The point is, there do not remain any predicates:

Here, however, I want to become conscious of myself only as thinking; I set aside how my own self is given in intuition. And thus the self could indeed be merely appearance to me who think, but not insofar as I think. In the

---

<sup>303</sup> Kant 1996/1781, 426 (B 409).

<sup>304</sup> Kant 1996/1781, 426 f. (B 409). My italics.

consciousness of myself in mere thought I am the *being itself*—of which, however, through this consciousness nothing is yet given for thought.<sup>305</sup>

Thus, the problem is, that even if it were possible to have a consciousness of myself as pure thinking, nothing would show up. The point is that, according to Kant, the Cartesian *cogito* is an empirical proposition. It is only possible insofar something is given in experience through intuition. Take empirical intuition away and nothing remains. If trying to determine the way of existence of pure thinking by building on the consciousness of pure thinking (genetivus subiectivus and genetivus obiectivus) we would be redirected to predicates applied to empirical objects:

For although I would have, through this marvelous power that the consciousness of the moral law first reveals to me, a purely intellectual principle for determining my existence, yet through what predicates would I do this? Through none but those predicates that must be given to me in sensible intuition; and thus I would in rational psychology end up again where I was: viz., in need of sensible intuition in order to provide signification for my concepts of understanding (substance, cause, etc.), through which alone I can have cognition of myself. These intuitions, however, can never help me advance beyond the realm of experience.<sup>306</sup>

Hence, even if we possessed the ability to put us in a state of mere thinking bereft of all empirical content, and if we were to describe such consciousness to determine the way of existence of mere thinking, we would fall back on using predicates that belong to the realm of experience, i.e., of empirical intuition. The hope to determine, in non-empirical terms, the *way of existence* of the thinking being we are, simply by referring to the character of thinking itself, is futile. On Kant's view, the nature of selfhood is elusive.

And, to sum up, the reasons that make up the argument for the elusiveness of selfhood are manifold:

- 1) To know something requires both concept and intuition. But in the case of the self, both are wanting. Neither do we find something that we could intuit that would refer simply and directly to the self, nor is the 'I think' a concept (but rather the *conditio sine qua non* for the application of concepts).
- 2) The only way we experience ourselves as thinking beings is through self-affection manifest in the inner sense. But the self-experience that results from self-affection amounts to nothing much more than a knowledge of how we are insofar as we appear to ourselves. Thus, we are only dealing with a relative and not an absolute self-

---

<sup>305</sup> Kant 1996/1781, 439 (B 429).

<sup>306</sup> Kant 1996/1781, 440 f. (B 431).



knowledge. For, such a self-experience does not present knowledge of how a thinking being *is in itself*, i.e., considered independently from experience. The determination of our being remains, therefore, limited to how we appear to ourselves.

- 3) Self-experience, analogous to the experience of anything else, physical things, which form the objects of our experience, does not capture the true metaphysical nature. We have no knowledge of what and how that what underlies what is given to us in experience is independently from experience, i.e., how the thing is in itself. I don't know what the table is independently from my experience of it, and, equally, claims Kant, we cannot further determine what the nature of selfhood is independently from our self-experience (other than some form of thinking spontaneity). In both cases, table or self, the true metaphysical nature is elusive.

### 3.1.2 Kant's motivation

Clearly, Kant's argument was meant to fend off the rational psychologists, the defenders of the notion of the soul, who maintained one can infer from the way we experience ourselves in thinking to the nature of selfhood, considered as it is in itself and independently from experience: i.e., that our self-experience warrants the belief in the classical soul. Kant wanted to intercept such inference from experience to being, from phenomenology to metaphysics. To do this, he introduced the appearance-reality distinction, which many take for granted in the case of the experience of objects in the world, and applied it to the way we experience ourselves. That, however, as I emphasized with regard to the first two skepticisms, might be a serious mistake. As I will argue now, understanding self-experience along the line of experiencing objects in the world has indeed some conceptual consequences that are problematic in that they distort how we think about the nature of selfhood—and, crucially, in a way that still gives too much credit to the soul-theory. Let me begin my clarification of these conceptual ramifications by granting a few points that I believe Kant succeeds to make.

First, he successfully demonstrates that we cannot conclude the simplicity of an allegedly incorruptible experience-independent soul from the unity of consciousness. The 'I think' has a logical, transcendental function in experience insofar as it is the very condition of the possibility of experience. But why should the unity of *experience* be a guarantee of a feature of something that ought to *transcend experience*? Such an inference is indeed not possible and Kant is right in pointing this out: I cannot describe a feature pertaining to the experiential realm and then ascribe the same feature to something that ought to be independent of experience.

Second, related to the first point, even if we had any possibility to describe the way mere thinking is present to itself and the nature of selfhood that reveals itself in such consciousness, we would always only have predicates that apply to possible objects of experience and that are thus ‘biased’ or pre-structured by experience.

Third, Kant, therefore, is right in claiming that *we cannot determine the nature of selfhood as it is independently of experience*, i.e., we cannot grasp how a pure self-awareness of thinking, independently of all empirical content, could possibly look like. We are incapable of describing such a kind of self-consciousness.

However, crucially, I believe it is an impactful mistake by concluding from that that the *nature of selfhood* is elusive. For, it remains utterly unclear why, given Kant’s arguments, we should at all believe it could possibly make sense to still speak of the nature of selfhood *as it is independently of experience*. Kant had emphasized we cannot apply any predicates through which we experience ourselves to the self considered in itself, i.e., independent of experience. Correct. But why should we then still believe that it makes sense to believe in selfhood considered independently of experience at all? Or to put it differently, if it is true that *we cannot determine the nature of selfhood as it is independently of experience*, couldn't it just be that it is simply impossible to determine *the nature of selfhood as it is independently of experience* because there is just nothing left that could be determined?

Interestingly, that option, namely that the reality of selfhood might essentially be bound to the experiential realm, i.e., that the existence and nature of selfhood *is* experiential and that it makes no sense to speak of selfhood considered independently of experience is something that is not considered in Kant’s approach. Accordingly, I submit, the notion of the elusiveness of the nature of selfhood is less skeptical than it purports to be. For, to speak of the elusiveness of the nature of selfhood as considered independently of experience one implicitly *posits* a nature of selfhood unknown to us, but such a thesis does not follow from the simple fact that we cannot determine our alleged nature as it is independently of experience.

Of course, the reasons that motivated Kant to posit such an unknown nature, I presume, are to be found in the depths of his philosophical system. For one, he wanted to retain the possibility of the belief in an immortal soul—for ethical concerns in the practical context. Moreover, closely related, he was eager to conceive of the human self as free. And, importantly, freedom is something that, according to Kant, cannot possibly be manifest in experience. Accordingly, if we are to be conceived of as free, i.e., if our nature is to be conceived of in terms of pure spontaneity, then either we must claim that we can experience our freedom or that we cannot experience the nature of selfhood.<sup>307</sup> For Kant, to accept the experience of freedom would corrupt his whole notion of experience and of knowledge.

---

<sup>307</sup> On Kant’s theory of freedom, see, for instance, Allison 1990.

The best solution, given the architecture of his system, thus, seems to be a notion of self that is elusive and that cannot be known through our self-experience.

### 3.1.3 Conceptual ramifications of Kant's elusiveness claim and soul-theory

Although noteworthy and generally deserving an own investigation, the reasons why Kant was motivated to describe the nature of selfhood as elusive are not decisive for my purposes here. What is more important are the consequences for the understanding of selfhood *when* it is considered as elusive. And it is in this way, I suggest, that Kant's approach is an example that has broad applicability.

Taking the nature of selfhood to be elusive implies that what the self is to *be* is not only much more than we experience, it is also supposed to be of different kind than our experience, a way of existence that is different to how we experience ourselves. But what exactly are we to think of this 'more'? What is the 'I' or selfhood, considered as spontaneous thinking activity and as independently of empirical givenness, that is, bereft of sensory experiences? Three general options, I believe, remain:

- (a) The self-consciousness of the 'I' as pure thinking (intellectual intuition) has a character that is qualitatively very different from the self-consciousness pertinent to our experience that is pervaded by sensory experience (empirical intuition);
- (b) Or the self-consciousness of pure thinking has no character whatsoever and is a *sui generis* consciousness that has no quality of givenness;
- (c) Or pure thinking is simply non-conscious.

Regarding Kant, I think it is fair to say that it is not absolutely clear what he had in mind. Considering his notion of "intellectual intuition", one might be inclined to say that option (a) gets the picture right, because in intellectual intuition pure spontaneity, pure thinking, is said to be conscious of itself in a completely different way as compared to empirical intuition, the kind of experience we enjoy as a matter of fact. But Kant does not say that we have reason to believe that there is such a consciousness intrinsic to pure thinking activity. He only considers that there might be thinking beings, for instance, god, who could have the capability for such an intellectual intuition, which seems in the end like a desperate attempt to make of the elusive nature of selfhood 'something' that, even if it must be nothing for us, is at least thinkable to be given to some god-like consciousness.<sup>308</sup>

---

<sup>308</sup> Cf. Kant 1996/1781, § 24. See also O'Farrell 1979, Gram 1981.

The attempt to connect the notion of intellectual intuition with a hypothetical god-like consciousness reveals a lot about the other options. What meaning does the nature of selfhood have if it is to be conceived of in terms that make it either non-conscious or conscious in a way that lacks all character of intuition as we know it, i.e., that is a completely non-qualitative consciousness without any what-it-is-likeness? In considering the possibility of an intellectual intuition what Kant seems to acknowledge is that the true nature of selfhood must in one way or another be a form of consciousness. If it is completely different from the consciousness we have and through which we experience ourselves, then at least it must be thinkable as a possible consciousness. If the true metaphysical nature of selfhood were to consist in something bereft of anything that resembles consciousness, one must wonder how such nature should be the *nature of ourselves*? Can I think *myself* as an essentially non-conscious process, activity, experience-enabling structure or something in any other sense completely different to that what I experience about myself in my experience? All three options (a, b, c), I take it, do not seem to give an attractive offer for the understanding of the nature of myself. A doctrine that considers the determination of the nature of selfhood in something that ought to be located somewhere beyond the experiential process is still too much entangled in the idea of an experience-transcendent soul that ought to underlie our experience and self-experience.

As James, rightly in my view, points out against the soul-theory, and which applies equally to the Kantian view:

The Soul, however, when closely scrutinized, guarantees no immortality of a sort *we care for*. The enjoyment of the atom-like simplicity of their substance *in saecula saeculorum* would not to most people seem a consummation devoutly to be wished. The substance must give rise to a stream of consciousness continuous with the present stream, in order to arouse our hope, but of this the mere persistence of the substance *per se* offers no guarantee.<sup>309</sup>

Although this quotation is concerned with the immortality of the soul, the idea of James's criticism can be applied to the Kantian notion of a nature of selfhood that is either non-conscious or conscious in a way not continuous, or worse even, incompatible with our experience. The point is that we usually do not care for such a completely experience-free nature of selfhood, one that, once bereft of all experiential context, lacks everything that makes our life as it is for us as the individuals we experience ourselves to be. What is more, it seems rather more likely that we *cannot* even care, in terms of *self-care*, for such an abstract pure thinking in which nothing else is consciously given than the spontaneous activity of thinking—consciously given in a way we cannot possibly grasp.

---

<sup>309</sup> James 1950/1890, 348.

Evidently, some may object that the question of what we are and the questions with regard to what we care about are to be kept well-apart, and that, therefore, James's interjection does not undermine neither soul-theory nor Kant's understanding of self. I would concede that both questions are certainly not congruent, given that we can care about things that neither form part of ourselves nor could not exist at all. Yet, I find issues less clear when it comes to *self-care*. Although admittedly not a conclusive argument, I am persuaded by James's suggestion that the way in which we care about ourselves does tell us something about our own being, and that self-care involves a—perhaps sometimes only implicit—understanding of what our being consists of. It seems strange to claim that our selfhood ought to consist in an indeterminable *x*, whereby *x* is an allegedly immortal soul, the persistence of which, if not somehow meaningful *for our consciousness*, seems rather irrelevant to us.

Another way to express this notion could be identified in Epicurus's famous advice to Menoeceus, when he says that we shouldn't fear death, given that death is the end of all sensation and consciousness, and that we, as experiencers, thus, won't be present when death has arrived.<sup>310</sup> Or, consider the role of death in the context of traditional spiritual and religious belief. In this context, belief in the immortality of the soul is often associated with some sort of belief in 'Heaven', 'Paradise', or 'Hell'. All of these ideas include some form of consciousness, i.e., the immortality of the soul is *eo ipso* considered to consist in some form of on-going of the stream of consciousness, although the self ought to find itself in a different 'world' that comes with either gratifying celestial experiences or punishing experiences. The post-life world might not really have the structure of a world but comprise only diffuse experiences such as sensations of 'light' and 'peace'—or horrific 'pain'. Yet, the crucial point is that any immortality of the non-empirical soul qua pure 'I' cannot be conceived of unless as involving some form of experience. The empty, alleged soul-substance matters not. This is also reflected in Locke's verdict, when he says that "whatever substance there is, however framed, without consciousness there is no person"<sup>311</sup>. This can also be applied to a non-substantially understood, pure 'I think': without any experience being involved, there is no one.<sup>312</sup>

---

<sup>310</sup> Epicurus 1993, 63f. (124f.).

<sup>311</sup> Locke 2008/1689, 216, II.xxvii, §23.

<sup>312</sup> In a sense, this is precisely what Hume was eager to show and maintained, when he claimed he has "no notion" of the mind "when conceiv'd distinct from particular perceptions" (Hume 2007/1739, 400 (App. 19)), which seems to imply that he does have a notion of self when self is considered in the context of particular perceptions. This supports Strawson's reading that Hume "never claims that the subject of experience isn't encountered in experience" (2017, 253). On this reading, all what Hume denies is the idea of a simple and perfect identical self-substance and the claim that our experience would motivate such an idea. It is with regard to *that* idea that he emphasizes "there is no such idea" (Hume 2007/1739, 164 (1.4.6.2)).

But if it be admitted that self-consciousness, independently of any empirical context in which some *x* is given to us, does form some kind of experience, then I would presume that this kind of experience is not absolutely different to the kind of experience we actually do enjoy within the empirical context, our experience of the world. It wouldn't be elusive.

Once the nature of selfhood, however, is declared elusive, then it seems utterly unclear what such a kind of selfhood has to do with *us*, the experiencers as which we experience ourselves in the empirical context. And most importantly, although Kant argues against the soul-theory, in describing the nature of selfhood as elusive, he restituted one of its characteristic ideas: the idea that there is 'something' behind our self-experience, something on a "non-phenomenal plane"<sup>313</sup>. Kant criticized the soul-theorists for positing that 'something' as a substance-like, persisting object. He took this to be a category mistake, because it applies what is a category of experience to 'something' that ought to be beyond and independent of experience. But the alternative he offers, namely to consider of that 'something' underlying our experience as some form of mere thinking activity, pure spontaneity fares no better, even if Kant acknowledges we have no means to further determine such activity. To criticize soul-theory for describing the alleged experience-transcendent soul based on self-experience, and then saying that the only problem with such an account is that we don't and cannot know what the experience-transcendent soul is like, is still in agreement with soul-theory *that there is* an experience-transcendent soul-like 'something' whether that be a substance or pure activity. Kant, thus, agrees with soul-theory that there is 'something', the 'I' or 'self-consciousness' as it is in itself and not insofar as it is given in experience, which *could* be determined, even though we, as human beings, are essentially incapable to do so. Even worse, Kant has stripped away all the predicates that were attributed to the soul by carving out that these are predicates that only apply to something given in experience. Thus, whatever Kant believes that lies behind our self-experience, the self as it is in itself can only be a very much impoverished version of the classical soul.

As James points out, if we understand the elusive nature of selfhood in terms of some mysterious kind of agency that can exist independently and prior to the experience we enjoy, and which is at least partially responsible for the constitution of that very experience, then we must conclude:

Well, if it be so, Transcendentalism is only Substantialism grown shame-faced,  
and the Ego only a 'cheap and nasty' edition of the soul. All our reasons for

---

<sup>313</sup> See James 1950/1890, 345: "By the Soul is always meant something *behind* the present Thought, another kind of substance, existing on a non-phenomenal plane."

preferring the ‘Thought’ to the ‘Soul’ apply with redoubled force when the Soul is shrunk to this estate.<sup>314</sup>

Such a polemic is entertaining, but its underlying message is also true. However, it should be noted that it is not a mistake that happened to Kant without himself noticing. Quite to the contrary, as I said, he was eager to preserve the possibility for us to *believe* in the soul, though we cannot *know* what it is like and whether we are indeed of substantial kind. And the insight, crucial for my purpose here, is that to preserve that possibility he declared the nature of selfhood elusive. It is the alleged elusiveness that allows believing in the soul without automatically having to defend it as an object-like substance, a claim which Kant, bearing in mind Hume’s skepticism, found impossible to defend on the ground of self-experience.

But again, my point here is not so much that Kant, as an individual person, was eager to safeguard the possibility for spiritual—and, ultimately, practically relevant—belief in the soul and its immortality. The crucial point is that there is an intrinsic relationship between his elusiveness claim and the notion of a soul, the notion of a bearer that is to underlie our experiences and self-experience, functioning as the alleged real self. The elusiveness claim is a vehicle for soul-theory. Whenever we construe selfhood in elusive terms, we open the door to soul-theory. Obviously, Kant did precisely not claim that we are immortal souls, and, more so, he claimed that we couldn’t possibly determine this on the ground of our self-experience. But what might seem as an act of enlightenment and skepticism, really is a sort of inoculation and perennial defense of the soul-theory. For, having posited that, based on self-experience, we cannot determine the true nature of selfhood, by the same token, means the unavoidable incapability to undermine soul-theory based on self-experience. Put differently, by emphasizing, against the rational psychologists, that self-experience couldn’t possibly *verify* the existence of the soul, Kant also highlighted by implication, against empiricist philosophy, that self-experience couldn’t possibly *falsify* the existence of the soul.<sup>315</sup> Moreover, by claiming that the nature of selfhood—qua elusive—lies beyond our self-experience, implicitly, what is posited is some *x* that could generally be determined. Both these last points illustrate how much weight the elusiveness claim carries, although talk of elusiveness is typically attributed a skeptical character. But given the ballast, it is far from clear that the elusiveness view can indeed claim the label of skepticism to begin with. For, to sum up, by disempowering our self-experience with regard to the determination of selfhood, the elusiveness claim turns

---

<sup>314</sup> James 1950/1890, 365.

<sup>315</sup> Although not explicitly referring to the metaphysics of self, this, I suggest, is precisely what is meant when Kant famously exclaimed: “I therefore had to annul *knowledge* in order to make room for *faith*.” (Kant 1996/1781, 31 (B xxx)) With the elusiveness claim, faith in the immortality of the soul is possible, because no empirical claim could possibly undermine a person’s belief in their immortality and substantiality.

out to be the thesis, i.e., the positing, of an indeterminable  $x$  that ought to be our nature—our real self—and which cannot be possibly ruled out by any act of reflection or any other act of our thinking. Let me now give an overview of the reasons why I deem such a move to be problematic and discuss them further.

### 3.1.4 On what's wrong with soul-theory and the metaphysics of elusiveness

I believe the following arguments ultimately render the elusiveness view unattractive. It should be noted that I acknowledge that not all of these arguments will be received as strict knock-out arguments. After all, these issues are precisely what is usually debated. Although I do believe that these arguments render the elusiveness view impossible, I shall note that, here, I only suggest the more modest claim that, considered in combination, they put a high weight on the shoulders of those who are eager to defend the elusiveness view.

#### 1) *The argument from the impossibility of self-care*

First of all, as I have already emphasized, pointing to James, it is hard to see why one would care for oneself as considered independently of experience. If our true nature consists in something completely different than our being as we experience ourselves, then in which way does our self-care concern that kind of elusive being? It seems impossible to care about the indeterminable  $x$  that ought to present our true metaphysical nature, if all our self-concern usually refers to ourselves insofar as we experience ourselves. Our self-care, as we know it from experience, concerns ourselves as experiencers. If the notion of the immortality of the soul is to elicit our interest, James exclaimed, then it must somehow involve the continuance of our experiential process, our stream of consciousness. Any immortality of a substance, however construed, that does not involve us as experiencers does not concern *us*, the experiencers we experience ourselves to be (even on Kantian terms and presumably according to all who accept self-experience but claim that our nature is elusive to our self-experience). I do believe James is right in pointing out that the phenomenology of our self-concern or self-care is revelatory with regard to what we are.<sup>316</sup> The point, I shall highlight, is not that if I care about  $x$ , then  $x$  constitutes my being. That is obviously wrong. Rather, the argument is that the self-care James describes is a constraint for how we think about the nature of ourselves: If  $x$  ought to be our nature of selfhood, then  $x$  must be something we could possibly care about in self-care. But if  $x$  could possibly be only determined and described in such a way that we could not possibly care about it, then  $x$  cannot present our true metaphysical nature. And that, I suggest, is precisely the case with

---

<sup>316</sup> In fact, I will further elaborate on that view, which is already to be found in Locke and other thinkers, in the Chapters 5–7.



Kant's elusiveness claim. Given that he deems the true nature of self to be indeterminable for us, the true nature of ourselves becomes something we cannot possibly care about.

Some may insist that *what we care about (in self-care)* and *what we are* can easily come apart or may indeed never overlap. Couldn't it just be the case that we care about ourselves as experiencers and that we care about our existence insofar as we experience, while at the same time our true and essential nature consists in something other than experiencing? I don't believe James has to rule out this possibility to make his point. The phenomenon of self-care he describes, if we agree it exists, does in itself point to our being as experiencers. That is, those who believe that self-care and being of selfhood are orthogonal to each other would have to provide a further reason why we should be skeptical about what our deepest self-concern is about.

## 2) *The argument from the lacking inter-esse*

I anticipate that many will find talk of self-care unconvincing, given that self-care, to many, will seem as something that may be characteristic of human selfhood or even only to some human individuals, but surely not as something intrinsic to selfhood per se.<sup>317</sup> Even so, I suggest we can make a similar point without reference to the affectively-loaded notion of self-care or self-interest. Instead of talking of the impossibility of self-care for or self-interest in our allegedly true being and metaphysical nature of selfhood, let me describe it as a lacking *inter-esse* between what I experience to be myself and that what ought to be my true being. What I mean by this is that having construed our true being, the 'I think' as it is in itself, as elusive and indeterminable, it becomes unclear how that very same kind of selfhood ought to explain or even form the selfhood that we are. Between the way I experience myself in the empirical context and the way I ought to be really, considered independently of the empirical context, is a hiatus. It becomes unclear how pure self-consciousness, bereft of the empirical context, ought to form precisely the very being of myself as I find myself *in* the empirical context. It should be noted that when we are intentionally directed at the objects of the world that we experience, we don't perceive them to be given to someone—some alleged true bearer of experience—located on a plane beyond and behind our experience. No, we perceive worldly objects insofar as they are given to ourselves as we experience ourselves *here* and *now* in the empirical context of the world.

When Hume claimed that we have no notion of self independently from our empirical consciousness, then this can also be interpreted in the way that we cannot possibly identify with that what ought to be our true nature and being. It is in this sense that one could say that whatever my true being is, it is not *me* whose being that is, because that

---

<sup>317</sup> In Chapter 6 and 7, I will defend the view that self-care might indeed be more intrinsically connected to self-experience and selfhood than many believe.

what is me is known only in the empirical context, which, on the elusiveness view, is precisely not what ought to be *me* considered ‘in itself’, i.e., just for myself, without any interference of something else given to me. But one might wonder, if the nature of selfhood is incomprehensible to me from the experiential standpoint, but I know myself only from the experiential standpoint, how could the non-experientially structured, allegedly true nature of selfhood possibly explain or otherwise elucidate the kind of selfhood associated with the experiential standpoint? The problem, I suggest, is not so much that we, as a matter of fact, as human beings cannot gain any insight into how our true, unknown nature may explain or determine our being as experiencers. The problem is that, rather, it seems impossible that whatever the structure of the pure ‘I think’ ought to be, it will never be a plausible determination of the kind of empirical ‘I’ we find ourselves to be in self-experience. For, there is no inter-esse between the two. By contrast, the way the Kantian position is construed, the one (pure self-consciousness) ought to present the true being, the *esse*, of the other (our empirical self-experience).

The hiatus between the ‘I’ considered in itself and the ‘I’ in (empirical) self-experience, I suggest, must not be unbridgeable. If the way I am given to myself when considered independently of any experiential content were essentially and incomprehensibly different to the way I am given to myself in the context of experiencing some empirical content, then it is unclear how the ‘I’ in the former case could possibly present the true metaphysical nature of the ‘I’ in the latter case.<sup>318</sup>

### *3) The argument from the misplaced analogy between object-experience and self-experience*

How did Kant end up claiming that we couldn’t possibly determine the true nature of selfhood? As I have already emphasized above, for Kant, this claim results from the way he conceived of objective knowledge and its conditions. Crucially, his conception of knowledge is based on the kind of knowledge taken for granted in the natural sciences,

---

<sup>318</sup> Protagonists of German Idealism, notably Fichte, did not accept the Kantian claim that the nature of the pure ‘I’ ought to be indeterminable, as he took this to undermine the Kantian project of a critical analysis of the conditions of possibility for any experience whatsoever. Since, in the Kantian system, all conditions of possibility for any experience whatsoever trace back to the accomplishments of the synthetic activity of the ‘I’, leaving the self-consciousness of the pure ‘I’ indeterminate risks building the transcendental system only on sand (cf. Frank 1991b, 418, 433). It was therefore that Fichte tried to demonstrate that an intellectual intuition *is* possible, i.e., that a determination of the pure ‘I’ is possible. Interestingly, doing so, he described the *being* of the ‘I’ in terms of consciousness, implicitly showing that the claim of the indeterminability of the pure ‘I’ only holds as long as one expects the being of the ‘I’ to be located *behind* and *beyond* experience, elusive to our empirical self-consciousness. Now, Fichte’s reception and development of the notion of ‘intellectual intuition’ is obviously a thorny subject, and I have no intention to engage in an assessment of it (but see Breazale 2013, Ch. 8). The only point I wish to make is that the epoch of German Idealism and its tasks were motivated, at least to a great extent, by the discomfort with Kant’s elusiveness claim with regard to the ‘I’, and that this discomfort not only concerns the relationship between empirical and pure ‘I’, as I have argued, but, crucially, also the validity of Kant’s explication of experience, objectivity, and knowledge.

which is rooted in the perception of objects. The leading paradigm in the way he treats our self-experience, thus, is object-experience. Such a strategy, however, was already something that I had identified in the first two forms of skepticism in Chapter 1 and 2, and found to be misleading and wrong. With Kant's approach, we can register another attempt. Where does it go wrong?

One aspect of it I have just touched upon when I discussed the argument from the lacking inter-esse. Let me elaborate on that a bit further to explain. I criticized that Kant accepted an unbridgeable hiatus between the way we appear to us in self-experience and our true nature as it is in itself. How could Kant possibly accept such an implication? Once again, the reason is that he took self-experience to be a specific case of object-experience. For, considering the experience of a given object in the world, we usually don't have a problem in accepting that how an object appears to us and the true nature of the thing in itself, considered independently of our perceiving it, could be very different. Now, to be sure, I don't want to claim that, in the case of the objects given to us in the world, talk of a thing in itself is correct, while in the case of myself as an experiencer it isn't. For, in this work, I want to remain neutral with regard to the notion of the thing in itself. But what I do want to point out is that in the case of object-experience, i.e., in its phenomenology, there is something that lends support to the idea of a thing in itself in a relevant way that doesn't apply to self-experience.

Consider first object-experience. I see a tree at the horizon. I presume that everybody will agree that the tree presents itself to us in a way that may induce the belief that the tree exists independently of our perceiving it. The experience of the tree involves the notion that whether or not I look at the tree—it is there. I can take a walk and come back to this site to contemplate the tree again or not. In both cases, under normal conditions of our everyday life, we will take the tree to be there. The phenomenology of the perception of the tree is pervaded by the sense that the tree is real independently of us. Accordingly, phenomenologically speaking, whenever we perceive an object  $x$  in the world, we do make an implicit difference between  $x$  insofar as it is given to us and  $x$  insofar as it is independently of our perception. I believe it safe to say that this is how the phenomenology of our perception is organized, and I take this to be a *relatively* uncontroversial claim. Kant goes a bit farther in claiming that we cannot possibly know what the thing in itself is, given that all our knowledge of it is based on experience. Whether or not we find this convincing, it certainly is a claim that seems compatible with the way object-experience unfolds phenomenologically in the unquestioned case of everyday life.

Contrary to this, things are much less clear, to put it mildly, in the case of self-experience. Do we experience ourselves in a way that we implicitly make a difference between ourselves insofar as we experience in self-experience and ourselves insofar as we are independently of our self-experience? I believe the clear answer must be no. Or, to put

it more modestly, if someone might be convinced that this is the case, this would in fact be a very controversial claim.

Now, in favor of that controversial claim, some might make the objection that even with regard to ourselves, we actually do experience a certain kind of transcendence. For, isn't it the case that when we experience ourselves to be the subjects of this current experience, we also take ourselves to be the ones to whom something was and will be given in the past and future? Don't we transcend the current experience? As I have argued in Chapter 2, I do believe that this is true. Even so, there would still not be an analogy with object-experience in the relevant sense. For, anticipating myself in the future is not taking myself transcending experience *as such* but only *this* current experience. Put differently, self-experience is typically and originally not such that I experience myself as transcending *all* experience, i.e., as if I was something beyond experience. My transcendence of experience refers to this *token* experiential episode but not to 'experience' as a certain *type* of event. On comparison, *that* is the case in object-experience. When I see the tree, I do perceive it in a way that implies a sense of the tree being real over and above anyone experiencing it. But, in the case of self-experience, I precisely don't see how we could be said to experience ourselves as being real independently of our own—nor for that matter—any experiential process. Again, Hume's verdict that we have 'no notion' of self independently of experience cannot be overestimated.

The misplaced analogy between the phenomenology of self-experience and object-experience reveals itself also—closely related—on the conceptual level, namely in Kant's construal of the notion of the thing in itself. In the case of thing-perception, the experience of a given object *x*, the thing in itself is construed by thinking *x* considered independently of experience. In the case of object-experience, that means independently of the subject's *spontaneity* and *receptivity*, that is, *x* independently of *x* affecting the subject, which would involve both the synthetic activity of the 'I' and the reception of empirical content to be unified in the synthetic activity of the 'I'. By contrast, in the case of the 'I', the notion of the thing in itself is construed a little bit differently. For, the 'I', considered in itself, does not present a case in which all spontaneity and receptivity are absent. Rather, according to Kant, it is to be considered as *pure* spontaneity. Independent from experience, in the case of the 'I', then, means bereft of empirical consciousness but not bereft of all consciousness. In fact, the thing in itself, in the case of the 'I', is considered as pure self-consciousness.

Hence, if we substitute the notion of *experience* with *consciousness*, it becomes more evident that the cases of 'I' and objects are incongruent. While the thing in itself in the case of object-experience is considered as 'independently of all consciousness', in the case of the 'I', the meaning of the thing in itself refers to the 'I' considered as 'independently of empirical consciousness'.

But why is that a problem? First of all, it shows that the notion of the thing in itself works differently for self-experience and object-experience, casting doubt on whether self-experience should be treated in the same way like object-experience. Second, it severely undermines the idea of applying talk of the thing in itself to the case of the 'I' (or selfhood). For, in order to even construe of it, Kant needs to argue in favor of the notion of 'intellectual intuition', which shows that the 'I' cannot even be thought of without associating with it the notion of some kind of intuition, something that we, if we are following Kant, usually take to be an aspect of empirical consciousness. So, when we construe of the 'I' as it is in itself, we always already imagine something experiential (in the broad, contemporaneous sense), a self-consciousness that is a form of consciousness, but one that is throughout different from the self-consciousness we know from the empirical context. But if Kant is saying that such an intellectual intuition is possible, what he is ultimately saying is that even independently of all empirical consciousness, the 'I' cannot be real other than by being somehow present to itself, by somehow being experienced by itself, even though this form of self-experience is to be absolutely different to the empirical self-experience we have as human beings. The point, thus, is that, for Kant, in the case of the 'I', the notion of the thing in itself does not refer to an 'independence of experience', nor for that matter, 'transcendence'. Rather, what Kant has in mind when he talks about the 'I', considered as it is in itself, is a complete perversion of that what is the thing in itself in our experience of objects. For, what Kant refers to when talking about the pure self-consciousness independently of all empirical consciousness is a kind of self-presentation that ought to occur in absolute *immanence*, allegedly independent of all transcendence, the most direct form of experience (broadly construed). If that is correct, then it is more than questionable why we should accept the Kantian claim that self-knowledge is a specific case of object-knowledge, which lies at the ground of Kant's elusiveness view.

4) *The argument from semantics: Givenness without the given, synthesis without the synthesized?*

The last argument consisted in the problem that the way the notion of the thing in itself works in the case of object-experience and self-experience are different respectively. The problem, I argued, is that Kant didn't yield to offer a notion of the 'I', considered in itself, that is analogous to an appearing object in the world, considered in itself. I now want to discuss whether the notion of the 'I' considered in itself, i.e., independently of empirical consciousness, can make sense at all. That is, whether Kant's construal of a pure self-consciousness independently of the empirical context is a convincing notion to begin with. For, it is far from clear that self-experience considered independently of any empirical content *p* or *q* can possibly make sense at all. I think that there are in fact semantic and phenomenological aspects that render even the notion of a pure 'I' or pure self-

consciousness bereft of any empirical content impossible. At least, as long as we understand self-experience like Kant and those many who have followed him, namely as the experience that *something is given to us*. It is the synthetic unity of our experience, the fact that empirical content is given to me in a unified manner that I experience myself as being the experiencer to whom objects are given. That is the self-experience we take for granted, and which is intrinsic to the empirical context of our experience of the world.

But now Kant aims for a description of the nature of this kind of self-consciousness considered independently of the empirical context, independently of the objects that are given *to me*. But I, myself, insofar as I appear to myself, experience myself in the ‘*to me*’ that pertains to the experience of *x*. It is in virtue of having the experience of *x*, i.e., that something is given to *me*, that I have an experience of *myself*. Stripping away all empirical content that could determine *x* and hoping one could leave behind the pure ‘*to me*’ seems futile. If nothing is given to me, then the ‘*to me*’ cannot remain unaltered, in fact, it cannot be. The notion that one could simply put the ‘*to me*’ as the object instead of all empirical content (= the ‘*to me*’ is given to me and nothing else), and thereby yield knowledge—objective even—of the ‘*to me*’, is hopeless. Not only is such a consciousness not possible *for us*, it does not seem to make sense to begin with. For, the sense or the determination of the ‘*to me*’ is fed through the context of empirical givenness (receptivity) in which something else is given ‘*to me*’. The ‘*to me*’ considered isolated from all empirical content and isolated from empirical givenness as such loses its meaning.

Now, the defender of Kant’s proposal might respond that this problem is precisely the crux with the notion of ‘intellectual intuition’, a form of intuition independent from (empirical) receptivity, a form of self-awareness of mere activity. Moreover, the ‘*to me*’ might characterize selfhood in the empirical context, but selfhood—better the ‘I’—considered as the spontaneous synthetic activity presupposed by all empirical givenness is not essentially characterized by the ‘*to me*’, which bears a passive connotation, but rather by some—to us incomprehensible—form of consciousness of mere activity. In fact, in the empirical context, selfhood might appear as if its sense consisted in the ‘*to me*’, in the being of some sort of passive locus or pole of experience. However, that is just how selfhood appears in experience to itself. *Really*, what the subject *is*, is some form of pure spontaneity and activity that cannot be determined by anything experiential, because any experience comes about only through the very activity of the subject. The real nature of the subject is its synthetic activity and what its structure is we can never know.

Even so, i.e., even if one were to grant Kant that the synthetic activity, through which we can experience something in the first place, presents the real—yet for us indeterminable—nature of selfhood, and that the ‘*to me*’ is only a feature of selfhood in experience (= in the empirical context), this does not mean that the self-consciousness of pure synthetic activity of the ‘I’, i.e., bereft of any empirical givenness, were possible.

Quite the contrary, the mentioned problem of the loss of meaning of the ‘to me’ when considered independently of the empirical context re-occurs with the alleged pure synthetic activity of the ‘I’. For, how could one possibly consider the synthetic activity independently of any synthesis through which an object comes into experience? Or, to use other words, if nothing is being synthesized (= synthesis through which an object is experienced) what meaning of synthesis remains? One cannot understand the synthetic activity through which object-experience is possible without the synthesis of an object. But this is not so much an epistemological problem because we, as human beings, are bound to empirical givenness (and object-experience). The problem is more fundamental, the activity without that what is being produced, the synthesis of experiences, *is* no activity. And again, it wouldn't help to claim that in the non-empirical self-consciousness of the pure synthetic activity of the ‘I’ it is the activity itself that is the object of consciousness. For, in this case, ‘object of consciousness’ means something completely different than what Kant refers to when he talks about the objectivity of the object in experience (= in the empirical context). Moreover, how could the pure self-consciousness make itself the object (however construed) of consciousness if not *in actu*, i.e., in the context of the synthetic activity that enables object-experience? The activity, the ‘I’, can only be self-aware of itself insofar it is indeed active. And it is active insofar as the ‘I’ is constitutive in object-experience. The synthesis qua synthesis makes only sense in the context in which something is in fact synthesized. Synthesis without the empirical manifold to be synthesized is a non-starter. Synthesis without what is synthesized is nothing. The synthetic activity without that what is the product of the synthetic activity, the empirical object, is nothing, for there is nothing to be synthesized.

Without empirical content, without anything that is given in experience, both the experiential feature of the ‘to me’ and the structure of the synthetic activity of the ‘I’, they dissolve. Once one had stripped away the ‘I think’ from the empirical context it loses its meaning. This is not so much the problem of an alleged limited human consciousness that could not yield a gaze on the allegedly pure synthetic activity. It is a semantic or conceptual problem. The synthetic activity considered independently of the synthesis, the synthesized experience, makes no sense.

##### 5) *The argument from the lack of skepticism*

The preceding arguments involve strong claims against the Kantian project and any attempt to consider selfhood in elusive terms. However, it should be noted that in order to criticize Kant's reference to the ‘I’, considered as it is in itself, it is not necessary to make these stronger claims that suggest to interpret the alleged *indeterminability* of the nature of the pure ‘I think’ as a manifestation of the *impossibility* of a pure ‘I think’. A more modest critique might be more attractive to some, and, crucially, it might suffice to successfully

undermine the attractiveness of the elusiveness view. Therefore, the following last argument against Kant's theory of the elusiveness of selfhood, which I wish to propose now, only points to the fact that Kant's position is much less a skeptical view than it purports to be: namely insofar as it makes some metaphysical assumptions that would actually be in need of further argumentation.

Additionally to the assumptions I have already mentioned, I now want to shed light on another. I shall argue that although it is an implication of the others, it still deserves to be carved out. For, crucially, it is precisely the relevant assumption that undermines the alleged skeptical character of the Kantian elusiveness view. *Prima facie*, it seems to be a skeptical position to say that, while self-experience may indeed reveal the reality of selfhood, it cannot provide any insight about its most original, true or metaphysical nature. On closer examination, however, rather than a skeptical position, Kant's view turns out to be a metaphysical thesis. It is a metaphysical thesis insofar as it posits that the nature or being of selfhood is something that ought to lie *behind* or *beyond* our self-experience rather than *in* our self-experience.

However, it should be noted that the elusiveness claim could in fact be formulated in a slightly different way, giving it a stronger skeptical character. Let me describe such a version in order to illustrate and isolate in which sense the elusiveness view of Kant is precisely *not* skeptical. A more skeptical position would consider the possibility that the nature of selfhood in fact consists precisely in the kind of self-experience we enjoy in the experience of the world, i.e., in the empirical context. That is, saying that from self-experience we cannot come to a conclusion with regard to a determination of our nature, on the more skeptical position, means that we cannot rule out that our self-experience actually presents the nature of our being. It might be true that self-experience does not suffice to verify any attempt to determine the nature of selfhood, however, from that it doesn't follow that our being must necessarily be different from empirical self-experience. Perhaps the kind of self-consciousness that we know from the empirical context, structured as we find it in reflection on our self-experience, our being in time, just *is* our nature. Perhaps empirical self-consciousness just *is* the true nature of selfhood. On the more skeptical version, that is indeed an option.

On comparison, when Kant says that the 'I' as it is in itself is indeterminable for us and the nature of the 'I' cannot be known through our self-experience, then he *does* rule out that our self-experience might present the true nature of the 'I', its original being. Hence, Kant's elusiveness claim goes farther than the thesis that any argument based on our reflection on self-experience is inconclusive with regard to the nature of selfhood. Rather, he says that the nature of selfhood, its inner being, is elusive to self-experience because the nature of selfhood lies beyond our self-experience. But that is a metaphysical thesis rather than a skeptical attitude towards one's own self-experience. And, notably, it



is a thesis that is possible for Kant only because the guiding paradigm for his analysis of knowledge and experience are object-knowledge and object-experience respectively. Such a paradigm, however, is a line of reasoning that requires further argumentation, one that would have to dissolve the mentioned worries that adhere to the attempt of conceptualizing selfhood in terms of something underlying our experience and self-experience.<sup>319</sup>

### 3.2 Illusiveness: McClelland on the self as the bearer of self-experience

So far, I have argued that Kant's elusiveness view doesn't offer an attractive theory of our self-experience and selfhood. The core argument was that Kant construed of the nature of selfhood as it is in itself in terms of a genuine and pure form of self-consciousness of which it is unclear how it could possibly present the nature of the kind of self we experience ourselves to be in empirical self-consciousness. Moreover, I emphasized that it is far from clear that such a pure self-consciousness is possible at all, let alone, a coherent notion, which I denied.

I now want to turn to an alternative strategy to cash out an elusiveness claim about the nature of selfhood. On this line of thinking, the nature of selfhood is elusive to self-experience, insofar as our self-experience appears to tell us what we are, when in fact it essentially distorts our nature. In a recent paper, Tom McClelland discusses such an approach under the label of the *illusion model*.<sup>320</sup> According to the illusion model, our self-experience may make it seem like we possess a set of features while none of these describe what the self really is. More specifically, it claims that, according to our self-experience, we may appear to ourselves as a single, immaterial thing which is irreducible to neural processes, as a free agent who has persistent individual characteristics etc., while really the self is something completely different. That is, on the illusion model, our self-experience does have a real referent, namely the *bearer* of our self-experience. But unlike our self-experience suggests, the real bearer of our self-experience is not some alleged immaterial self or soul but the *brain*.

---

<sup>319</sup> In Chapter 5, I will discuss contemporaneous accounts that motivate a concept of selfhood according to which selfhood cannot be defined in terms of self-experience but only in something that underlies self-experience. Typically, a central topos in their argumentation is the notion of personal identity. Before I can enter that field, however, in Chapter 4, I introduce a contemporaneous account of selfhood that actually does define selfhood in terms of self-experience.

<sup>320</sup> McClelland 2019, 28. See also Section 1.1.5 for a comparison to similar yet different approaches.

The illusion model, as McClelland describes it, is meant as an alternative to the hallucination model that I discussed in Chapter 1, which claims that self-experience has simply no referent and is therefore to be considered hallucinatory.

The alternative to this is an *illusion model* of the self according to which self-representations do have a real referent—the self—but systematically and dramatically misrepresent what features the self has.<sup>321</sup>

Against the hallucination model, McClelland argues that it overlooks that the self-representations (or self-experiences) need *some* bearer, whether that be the self or something else. As he points out:

Now, if one posits the existence of mental representations one must also posit a bearer of those representations *B*. To dispense with *B* whilst leaving its representational states behind would be like dispensing with the Cheshire Cat whilst leaving its smile behind: there is an ontological dependence here that cannot be avoided.<sup>322</sup>

Technically, then, McClelland suggests, the question is not so much whether self-representations refer to *something*. The question is rather *what* they refer to. Hence, the hallucination model, which suggests that self-representations have no referent, is too quick in denying the reality of self. More specifically, McClelland emphasizes, to maintain that the brain is *not* the self self-experience refers to, the hallucination model would have to provide compelling arguments. However, the ones provided by defenders of the hallucination model, McClelland stresses, are insufficient. Typically, as he takes it, they argue that our self-experience purports to refer to an immaterial self or soul, while really there is no entity that has the properties our self-experience attributes to ourselves. McClelland objects that, from the fact that there is no entity that had such a profile of properties, we cannot conclude directly to the claim that there is no self at all. After all, he highlights, it could well be that our self-experience simply gives rise to a fundamental misconstrual of the nature of the being of our—in fact, real—self:

[T]he premises of the argument for VST [virtual self theory, P.S.] only entail the disjunctive conclusion that *either* the self is a merely intentional entity (as per the hallucination model) *or* that the self is a real entity that is misrepresented in our self-representations (as per the illusion model).<sup>323</sup>

---

<sup>321</sup> McClelland 2019, 28.

<sup>322</sup> McClelland 2019, 34.

<sup>323</sup> McClelland 2019, 28.

Now, McClelland proposes that the case of the illusion model is stronger than that of the hallucination model. For, even the hallucination model has to acknowledge that there is at least some physical bearer of our self-experience and it would have to give further reasons why it would be wrong to identify the physical bearer, the brain, with self. Crucially, McClelland further notes that, regarding the question whether the brain is the self, the burden of proof is on the side of the hallucination model. This is so because there are, according to McClelland, indeed good reasons to believe that the brain is the self:

It appears to each one of us that we have a self, and it appears that  $p$  when we have defeasible reason to believe that  $p$ . This suggests that the burden of proof is not on advocates of the illusion model to demonstrate the theoretical need for positing selves. Rather, the burden of proof is on advocates of the hallucination model to demonstrate that every appearance is overturned by some defeating consideration.<sup>324</sup>

I believe McClelland has a valid point against the hallucination model (although Metzinger does of course acknowledge that the brain is the bearer of self-representations). However, since I have already argued against the hallucination model in Chapter 1, here, my interest is not with the comparison between the illusion and the hallucination model. I am convinced that the hallucination model is wrong. However, that doesn't necessarily mean that the illusion model presents a convincing account—even though it is admittedly more plausible than the hallucination model. Yet, the scope of McClelland's argument does not involve a defense of the illusion model against all different kinds of accounts but only against the hallucination model.

What I want to do now is examine whether the illusion model, considered in the way McClelland has depicted it, could avoid or respond to some of the objections that I have brought forward against Kant's elusiveness view of self. As I will argue, even though the illusion model is different in many important regards compared to Kant's elusiveness view and thus may offer some alternatives, ultimately, it is exposed and vulnerable to the same issues. To see this, we need to shed light on the core argument of McClelland that is to motivate the illusion model.

### 3.2.1 Who is the bearer? Identifying the self with the brain

McClelland's core argument for the identification of the self with the brain runs via the notion of the *bearer* of experience and self-experience. For him, self-experience and how we talk about self-experience suggests that we should take the self to be the bearer of our experience and self-experience:

---

<sup>324</sup> McClelland 2019, 31.

After all, a whole host of platitudes about the self revolve around it being the bearer of mental states. To say that the self is the agent of one's actions, the thinker of one's thoughts or the subject of one's perceptual experiences is to say that the self is the bearer of those various categories of mental state. One might even make a case for saying that being the bearer of mental representations is *at the very core* of our understanding of what it is to be a self.<sup>325</sup>

Having identified the bearer of experience and self-experience with the self, on the one hand, and having identified the brain as the neurological underpinning of experience and self-experience, on the other hand, McClelland proposes that the brain is the bearer of experience and self-experience, and that it therefore must *be* the self. On this view, it might appear to us that we are some kind of free agent, some immaterial substance that persists over time and in which inhere the different experiences that unfold in the stream of consciousness, when really we are fundamentally brains, given that it is the brain that is the bearer of all our experiences. The argument can be summed up in the following way:

(P<sub>1</sub>):           The self or selfhood consists in being the bearer of (self-) experience.

(P<sub>2</sub>):           The brain is the neurological underpinning of (self-) experience and therefore the bearer of (self-)experience.

Conclusion:     The brain as the bearer of (self-)experience is the self.

Before I assess whether identifying the self with the brain fares better than the elusiveness view with regard to the objections I discussed above against Kant, I want to shed light on whether the argument is valid in itself. *Prima facie*, it does seem compelling. However, on closer examination, I suggest, there is a problem with the element that ought to connect both premises, the notion of the bearer of (self-)experience. For, the expression 'bearer of (self-)experience' can be understood in two different ways. A first understanding takes the bearer to be that what underlies (self-)experience: the metaphysical substance. From such a *substantial understanding* an *experiential understanding* can be distinguished: I, myself, the one who is undergoing different experiences, am the bearer of (self-)experiences insofar as they are all presented as mine. Unlike in the case of the substantial understanding, the experiential understanding of the bearer only refers to the fact that I am the pole of all experiences that I live through. No reference is made to a bearer that is behind my (self-) experience, underlying it, as it were.

Now, the problem with McClelland's argument, I believe, is that, in the two premises, a different understanding of the bearer is operative respectively to the effect that both the experiential and substantial understanding of the bearer are mixed up in the conclusion. The argument, however, can only be convincing if in both premises and the

---

<sup>325</sup> McClelland 2019, 35.

conclusion the notion of the bearer is identical. That, however, would only be possible if in both premises and the conclusion the substantial understanding of bearer would be predominant. For, in case of P<sub>2</sub>, the experiential understanding of the bearer doesn't make sense. Even if it might be true that the brain is involved in the generation of our experiences and can, in this sense, be said to be underlying our experiences, it wouldn't make sense to say that all experiences are given to the brain and that the brain appears to itself as the pole of all experiences.

How about the substantial understanding? Can we make the argument valid by operating only with the substantial understanding? The short answer is that we could *if* we were indeed accepting that self-experience must refer to an underlying substance. It is only then that P<sub>1</sub> would read in a way compatible with a substantial understanding. But is it really clear, let alone a platitude, that the self or selfhood consists in being a *substantial* bearer that underlies (self-)experience?

Not only is this a strong metaphysical thesis that certainly requires argumentation. It is also far from clear that this is the platitude McClelland refers to. Nor for that matter that he himself is operating with a substantial understanding when he says that we usually take the self to be the bearer of (self-)experiences. Finally, even if he did believe that our self-experience is such that we consider ourselves to be substances underlying our self-experience, it is not clear that the phenomenology of self-experience is indeed such.

In fact, I believe that if anything is a platitude about the self being the bearer of experience, then it is a platitude only in the context of the experiential understanding. That is, yes, most people will agree that all their experiences are given *to them* insofar as they are the experiencers. But from this, it precisely doesn't follow that there is some substance—immaterial or material—that would be underlying all our experiences. Moreover, crucially, it doesn't even follow that it appears to ourselves that there is one single substance underlying all our experiences that presents our real being. I can experience myself as being a free, experiencing agent who exists as experiencer, with certain persistent characteristics as an individual without thereby necessarily taking myself to be an identical substance that underlies all my experiences. To conclude, McClelland cannot refer to platitudes if he were eager to motivate P<sub>1</sub> with a substantial understanding of the bearer of (self-)experience. Worse even, I take it that he would have the rather uncontroversial descriptions of self-experience, the platitudes, against himself if he tried to launch an understanding of the self in terms of a substance—even though the alleged substance is identified with the brain rather than an immaterial soul. This becomes more palpable if we consider what it would mean to define the self in terms of the brain.

### 3.2.2 The brain doesn't fare better than the pure 'I think'

In Section 3.1, I argued that Kant's conception of the pure 'I think', the 'I' considered in itself, faces several problems that mostly concern the relationship between self-experience, i.e., the experience of being the one who is living through different experiences, and what ought to be the real nature of selfhood. I argued that the way Kant conceptualizes the alleged elusive nature of the 'I' creates an unbridgeable hiatus between our self-experience and that what we ought to be. Most importantly, I criticized that the elusiveness claim implies that there is some  $x$  behind our self-experience that we ought to be, applying the distinction between appearance and reality, which we know from thing-perception, to the case of self-experience.

Importantly, all the problems that I mentioned arise when defining the self in terms of an indeterminable  $x$ , also emerge when one defines the self in terms of the brain. Accordingly, I suggest, whether one takes the self to be some experience-transcendent soul-substance, some elusive and indeterminable  $x$ , or a material thing such as the brain, it matters not. In all cases, the following issues emerge.

If we define selfhood in terms of the brain, it becomes obscure how I could care about myself *in* caring about my brain. Evidently, we all care about our brains. We rather have a healthy organism and a functioning brain. We all know or somehow feel that our head and the brain in it is of central importance for our being in this world. But the kind of care about the brain is derivative of our self-care that—if James is right—is about the ongoing of our experiential process. We care about our brains because of its importance *for* our experiential life, *for* us. Our care for our brains is secondary in that it ensues from our care about ourselves as experiencing beings, which is primary. Or, to describe the same from another angle, our self-care could never be fully reduced to our care about our brains.

Moreover, one can express the same fact by considering the brain without being involved in the generation of self-experience, for instance, in form of a thought experiment, when the brain is kept in function while only having stripped away its ability to generate experiences. In this case, would we be dealing with a case of selfhood? I don't see how we could make sense of the claim that the brain—even though it cannot produce any experiences any longer—while otherwise functioning could possibly be regarded as presenting a case of selfhood. As Theodor Lipps emphasizes: "Without the immediately experienced 'I' the real 'I'—the substrate—would be nothing more than a nameless something."<sup>326</sup> Allowing ourselves another brief and absurd thought experiment, imagine it turned out that we all lacked a brain, other things being equal. While this would be a shocking revelation, it would by no means undermine the immediate belief that we exist as

---

<sup>326</sup> Lipps 1901, 41, my translation: "Ohne dies [dem unmittelbar erlebten Ich, P.S.] wäre das reale Ich nichts als ein namensloses Etwas."

experiencers, and there is a great chance our whole life as we know it would proceed unharmed. All it would give rise to is the counterfactual insight that brains are not required for experience. But it would not necessarily lead to a relevant doubt about the reality of ourselves and the social and cultural lifeworld that is founded on our experience.

Most importantly, there is also a fundamental problem with translating or transferring the semantics of the notion of the bearer of (self-)experiences in its experiential understanding to a substantial understanding. Again, to stress, the kind of self-experience McClelland takes for granted, in its most formal and uncontroversial formulation, is the experience of being the one who is undergoing different experiences, i.e., the experience that *p* and *q* are given to me. But how could the brain ever play this role? In the absence of experience, neither are there any experiences that could be the experiences of anyone, and the unconscious brain cannot play the role of selfhood. Or, if there is indeed experience, it is the experiencer in experience, it is *me* who is already playing the role. The defender of the illusion model may object that it may seem to you that it is you—the experiencing self—who is undergoing the experiences, when it is really the brain. But to speak like that “requires infiltrating intentional vocabulary into the description of subpersonal processes”<sup>327</sup>:

I am not my brain—for my brain is certainly not married, not a psychiatrist, and it has no children. Even worse, it does not see nor hear anything, it cannot read or write, it cannot dance or play the piano, and so on. Thus, I am rather glad not to *be* my brain, but only to *have* it.<sup>328</sup>

Accordingly, I believe it is correct to say that it would be a category mistake to insist that the brain is the true experiencer, while the experiencing self is only an illusory appearance—a misrepresentation—of the real facts. The self as we experience it and the brain are ‘located’ on different semantic planes. And, importantly, the relationship between self and brain is not comparable with the relationship between our perception of a tree and the tree as it is independently of our perception of it: self-experience/brain is not parallel to tree-perception/tree. Again, there is no inter-esse between myself as experiencer and the brain which is an object. The fact that my brain stands in the intimate relationship with my experiences in that, qua neurophysiological underpinning, it is involved in their generation, doesn’t change the fact that I cannot identify myself with the brain in a way that would make sense. Once we accept that we have self-experience in terms of an experience of being the owner of different experiences that are phenomenally lived through by us, there is no object, not even the brain that is so crucial for our experience, which could either

---

<sup>327</sup> Fuchs 2017a, 44.

<sup>328</sup> Fuchs 2017a, 44.

logically or phenomenologically play the role of the self.<sup>329</sup> I believe the analytic philosopher Ingmar Persson is right when he emphasizes:

Our selves must logically be that to which our experiences are non-derivatively attributable—the owner aspect—and the perception of ourselves that is necessary for the attribution of experiences to ourselves must be full in the sense of revealing the essence of what is perceived—the phenomenal aspect.<sup>330</sup>

*If* we—and I think we should—acknowledge that we have a self-experience of being the owner of our experiences insofar as they are lived through by us, then we also have an awareness of what could possibly play the role of the self: an experiencing—and, that means—experiential self. Persson, in a similar manner, concludes:

Now, there can scarcely be any other kind of material thing to which that which satisfies the owner and the phenomenal aspects of the notion of the self essentially belongs. ... So, as used by us, tokens of 'I' do not identify things of any kind that can take the place of our human organisms. ... [W]e exist, but are not identical to things forming any kind.<sup>331</sup>

As we can see, thus, identifying the self with the brain may seem to solve the referential issue of self-experience. However, similar problems as in the case of Kant's pure 'I think' emerge, when we try to understand in which sense the brain could be the real self or could present the real nature of selfhood. The illusion model doesn't offer sufficient advantages over the Kantian elusiveness view.

Finally, I am not fully convinced it would even have major advantages over Metzinger's hallucination model. For, part of the advertisement of the illusion model, one may take it, is that it fits better with our everyday intuition of our existence—the fact that we exist, whatever we are. Even though not all of our self-experience is correct, the illusion model tells us, at least the fundamental experience of existing is not a mere hallucination. Yet, does the view that we exist as selves but that we are really brains provide a conciliating view? One simple ramification of such a view, one may take it, as I have already mentioned, is that all parts of our lifeworld that involve ourselves as experiential selves—notably, the social and cultural lifeworld—would all lose their status of reality and would have to be considered illusory. Rather than free agents that interact, we are all really natural things—brains—that somehow stand in relations to each other that are purely structured by physical laws. Everything that rests on the semantics of ourselves as experiential selves would be

---

<sup>329</sup> It is a different thing to say though that the experiencing self as we experience it amounts to a certain kind of object. That, however, requires an alternative conception of what 'object' means. Cf. Strawson 2017, ch. 3. Here, however, my interest is with the more widespread, traditional understanding of what an object is.

<sup>330</sup> Persson 2004, 407.

<sup>331</sup> Persson 2004, 408.



threatened by the illusion model just as it is the case with the hallucination model. If we were really eager to adopt a position with such a drastic implication, I propose, the case made for it would have to be much stronger than that of the illusion model.

### 3.3 Elusiveness II: Two takes on why the self must be elusive

In the last sections, I have argued that if we accept the claim that our experience involves the self-experience of being the experiential center to which the different experiential contents are given, and, if we take such a self-experience to present a consciousness of one's own existence, then we encounter problems if we were also to hold that we couldn't determine the nature of selfhood by examining our self-experience. If we acknowledge that we have self-experience, I suggest, we cannot declare selfhood to be elusive. As a consequence, if we acknowledge that we have self-experience then we should also come to accept that self-experience can tell us what is essential to the nature of selfhood. Before I turn to an approach that, in fact, *defines* selfhood *in terms of* self-experience rather than in terms of something that ought to be beyond experience, in the remainder of this chapter, I want to briefly discuss two further elusive views. Both argue that the self must be elusive so as to enable the structure of intentionality that is characteristic of our experience of the world and pertinent to any object-experience.

#### 3.3.1 Howell: Non-objective self-awareness means elusiveness of the self as the bearer

Let me begin with Robert J. Howell's account that resembles in many regards McClelland's illusion model. Similar to the illusion model, Howell emphasizes that our experience involves self-experience: "Even in the first-person perspective, with mental states transparent as you please, there is an indubitable sense that the things presented are presented to *me*."<sup>332</sup> And, like the illusion model, Howell is eager to identify the self with the brain: "[A]s far as I can tell, the best candidates for selves and subjects are brains . . . ."<sup>333</sup> But, unlike the illusion model, he doesn't claim that our self-experience is illusive but rather elusive. For, he doesn't hold that we experience ourselves as objects of any kind, at least not in the most original form of self-experience:

While conscious mental states do generate a form of self-awareness, they still do not present the self to perception as, say, the roundness of a record presents

---

<sup>332</sup> Howell 2010, 478.

<sup>333</sup> Howell 2010, 480.

the record to perception. When I am having a conscious mental state, the awareness generated is an awareness that the mental state is mine, but it does not present an awareness of my self. In this sort of self-awareness—when it is pre-reflective—the self is not an object of an intentional state.<sup>334</sup>

Put differently, Howell highlights that our experience involves the self-awareness of the experience being ours without us being presented with an object ‘self’. In fact, he points out that if the self were to be presented as an object, that would jeopardize our intentional directedness at the objects in the world:

It is the nature of subjectivity that the subject itself is not an object of awareness from the subject’s own point of view. . . . Rather, subjectivity necessitates that the subject escapes the first-person point of view, because it is necessary that a subject’s mental properties be intentional and therefore do not present their bearers. If this were not the case they would constitute veil thicker than any imposed by modern philosophy’s ‘way of ideas,’ effectively shutting off thought and destroying the sense in which the ‘subject’ could accurately be called a subject at all.<sup>335</sup>

Now, as mentioned, Howell does believe the self or subject is a form of object, the brain. But he simply takes it that for the intentional structure to be possible, we cannot be given to ourselves as objects while being directed at some worldly objects. If we appeared to ourselves as objects, then, for him, the objects of the world couldn’t be given as objects at the same time. Building on Sartre, Howell emphasizes the self, if it appeared as an object, would block our view on the world. And this is why he considers the self to be necessarily elusive to our self-experience. Ultimately, thus, Howell comes to make three important claims:

- (1) Mental states involve self-awareness, which is “propositional” but “not objectual”<sup>336</sup>.
- (2) Because we don’t appear to us as objects, the self which *is* nonetheless an object, can be rightly described as elusive: “So although mental states—by virtue of their indubitably being mine—do generate a sort of self-awareness, it is not the sort that contradicts the elusiveness thesis.”<sup>337</sup>
- (3) The self is elusive in that it doesn’t appear as an object, although it is best conceived as an object, more specifically, the brain.

---

<sup>334</sup> Howell 2010, 479.

<sup>335</sup> Howell 2010, 480.

<sup>336</sup> Howell 2010, 479.

<sup>337</sup> Howell 2010, 479.

Does Howell's account provide an advancement over the illusion model? I find Howell's phenomenological conception of self-experience more accurate than that of the illusion model in that Howell highlights that we don't experience ourselves as objects. However, given that this is so, the question arises as to why he still considers the self to be an object. After all, he emphasizes "that a subject's mental properties be intentional and therefore do not present their bearers"<sup>338</sup>. But if the mental properties *of* the subject are intentional in a way that includes an absence of their bearers, why should the bearer, the brain, *be* the self? Just like in the case of Kant and the illusion model, a problematic duplication of self occurs rendering the issue of how the self, considered as it is operative *in* experience, and the self, considered as the underlying bearer of experience, ought to come together. Once admitting that the self or subject's role consists in being intentional, i.e., in being the center around which different experiences of *x* gravitate, then introducing an object that ought to be the real self evokes the same issues as in the case of Kant and the illusion model.

### 3.3.2 Evans: Identifying the elusive self with unprojected consciousness

Let me finally come to a last elusiveness view that I want to discuss: C. O. Evans's suggestion to identify the elusive self with what he calls "unprojected consciousness"<sup>339</sup>. Evans's proposal is significantly different to the positions I have discussed in this chapter so far, and, as it will turn out, cannot be rubricated as a skeptical position.

First, unlike the illusion model and Howell's approach, Evans emphasizes that the self should not be considered as an object. Evans advertises his approach as an alternative to the traditional "Pure Ego Theory" and the "Serial Theory of the self"<sup>340</sup>. With the former, he refers to the view "that the self must lie outside experience, and must be unknowable in itself" and that "our knowledge of the self is essentially inferential"<sup>341</sup>. With the latter, he refers to the bundle view often ascribed to Hume, the view "that the self is constituted by experiences"<sup>342</sup>. Now, according to Evans, despite their heterogeneity, both conceptions share a common presupposition, which render them infelicitous depictions of selfhood. He says:

The presupposition is that the type of account that can be given of material objects can also be given of the self: in other words, these theories of the self in effect transpose to the self a theory formulated in the first instance as an account

---

<sup>338</sup> Howell 2010, 480.

<sup>339</sup> Evans 1970, 104, 149.

<sup>340</sup> Evans 1970, 29.

<sup>341</sup> Evans 1970, 29.

<sup>342</sup> Evans 1970, 29.

of the relation between a material object and its properties. By implication, therefore, they treat the self as an object.<sup>343</sup>

Supporting my argumentation throughout the chapter, Evans goes on to dismiss the attempt to conceive of the self as an object:

The fault I find with both the Pure Ego Theory and the Serial Theory—thesis and antithesis, as it were—is that for all the difference it makes they might as well not be about ourselves at all. The theories take no cognizance of our native knowledge of ourselves. They show no awareness that it might beg the question simply to assume that selves can be given the same type of analysis as objects.<sup>344</sup>

Secondly, as these quotations already reveal, Evans's critique not only concerns the classical soul-theory or the illusion model, it is also directed against Kant's understanding of the elusiveness of self. For, crucially, Evans takes the self to be knowable through and in experience. This is so, because he takes the self to *be* experiential, although it doesn't appear as an object:

My identification of the self is an identification of the *subject* of consciousness. The position is that the self is experiential, but is not on that account an *object* of experience. ... The major implication of the position is that from the fact that the self is not an object of experience it does not follow that it is non-experiential.<sup>345</sup>

Bearing in mind that Evans also doesn't buy into the view that the self is composed of its experiences—as a bundle, as it were—he further highlights: “It does not follow therefore that the experiential self is the same as the empirical self.”<sup>346</sup> For him to say that the self is experiential neither means that the self is an object nor that it is constituted by the contents of experience. Evidently, when we turn our attention towards ourselves as experiencers, then we can make ourselves the object of attention. But in that case, we experience ourselves only considered as an object, but not as the subject that is performing the attention, the attending subject:

[I]ndeed the individuating of an experience by a subject ... just is to make of it an object of attention. Bearing this in mind it is evident that an experience of the self is an experience *qua* object of attention and as such we are dealing with the self as object (the empirical self). But, as will soon become apparent in the course of the exposition, the self *qua* subject—the experiential self—cannot as

---

<sup>343</sup> Evans 1970, 29.

<sup>344</sup> Evans 1970, 30.

<sup>345</sup> Evans 1970, 145.

<sup>346</sup> Evans 1970, 145.

a matter of logical necessity become the object of attention; cannot therefore become an object.<sup>347</sup>

For Evans, rather than a deficient mode of self, the fact that the self qua subject cannot be given as an object is a real feature of the self. That is, it isn't the case that the real features of self were elusive—as the other accounts discussed earlier had it. Rather, according to Evans, the real feature of the self *is* its elusiveness, which is an experiential property characteristic of unprojected consciousness, i.e., the “elements of consciousness that together make up the background of consciousness when attention is paid to an object”<sup>348</sup>. Hence, instead of saying that we cannot see what the self really is like because we are usually busy seeing objects, Evans says that it is just the nature of self to remain in the background of our consciousness of the world. It is never meant to be seen in a way that we see objects, but only resides in our experience by forming the background of our attention. In fact, the self forming the background of attention “is one of the factors that accounts for the view that the self lies behind its experiences”<sup>349</sup>, when in fact it isn't anything outside experience but part of it as unprojected consciousness. Defining the self in terms of the elusive background of attention has some advantages, as Evans emphasizes:

This enables it to be asserted in all consistency both that the self as subject is experiential, and that it is never presented as an object of experience. Furthermore, it obviates the necessity of treating the self as something unknown in itself.<sup>350</sup>

The message, then, is that in order to determine the nature of self, we don't need to look anywhere behind or beyond experience. And, we shouldn't be misled by a wrong interpretation of the specific phenomenology of our self-experience that—for so many philosophers and thinkers guided by the perceptual model of object-experience—has given rise to the idea that the self must be some kind of thing that underlies—in which way ever—our experience.

As it shouldn't come with surprise, I think Evans makes a strong argument: It is not our self-experience of being an experiencer that would motivate the idea of an experience-transcendent self that—substantial or not—underlies our experiences on a non-experiential plane. Such an idea is rather rooted in the tendency to apply notions that stem from the context of object-experience to the case of self-experience. For Evans, to speak of the elusiveness of self, thus, is to deny both that the self qua subject does or could possibly appear as an object. Both claims I find accurate. By calling it elusive, Evans describes the

---

<sup>347</sup> Evans 1970, 145.

<sup>348</sup> Evans 1970, 104.

<sup>349</sup> Evans 1970, 150.

<sup>350</sup> Evans 1970, 150.

particular phenomenology of selfhood and means that this is what the phenomenology and the being of selfhood is like. This is a positive—and, as I take it, correct—claim. At the same time, this is a negative claim, which consists in the view that the elusiveness of self means precisely *not* that we couldn't determine the phenomenology and being of selfhood. Again, I agree. However, I do find the label of *elusiveness* problematic. At best, it is inviting misunderstandings. At worst, it implies that there is still room for the idea that, if one only could *per impossibile* shed light on the self qua subject *while* it is intentionally directed at a given object, one could find out more about the self and its nature. And, accordingly, that there is more to the nature of the self than we can fathom while being directed at an object. That, I argued, is the case with Kant. Moreover, to call the self elusive suggests that there is not much more to say about its phenomenological part in experience other than that it is pervasive and somehow always forming the background of attention. By contrast, I suggest, not only is there much more to say about how self-experience figures in our experience of the world, it is also required to make more explicit in which sense something like the background of attention could be identified with the self. Put differently, what we need is a clearer phenomenological account of how self-experience *reveals* the nature of our selfhood, and, crucially, in which sense self-experience could be said to *present* the nature of selfhood. In which sense is the self experiential? To conclude, although Evans's proposal points us in the right direction, I take it, by connecting his account with an elusiveness view, he obfuscates the further phenomenological tasks that wait if one accepts that the self is essentially experiential. Therefore, I suggest, we need a different point of departure, one of the kind I will discuss in the next chapter.

### 3.4 Conclusion: The elusiveness view still gets the picture wrong

In this chapter, I presented a third form of skepticism about self-experience. Unlike the two forms of skepticism that I discussed in the previous chapters, the third skeptical approach does acknowledge the existence of self. It doesn't deny the reality of selfhood, it doesn't deny that we have self-experience, and it doesn't deny that self-experience involves a consciousness of our existence. However, it denies that our self-experience could possibly reveal what the nature of selfhood is. According to the third form of skepticism, selfhood is elusive.

As it turned out, proposals falling in that third category all, in some way, are guided by the paradigm of object-experience. Their claim that the nature of self cannot be determined through self-experience is rooted in the observation that the self qua subject cannot be given as an object in experience without losing the specific subjective character that the self has when it is directed at an object other than itself. It simply slips away from

our reflective gaze. Kant held that we cannot determine the structure of pure self-consciousness as it is in itself, i.e., without object-consciousness. Therefore, he also concluded that we cannot determine whether our nature is that of a substance. Crucially, however, as I carved out, he petrified the possibility of the belief in that the self is a soul-substance. The illusion model, in a straightforward manner, identifies the material underpinning of our experiences, the brain, with the bearer of our experiences and takes the brain to be the real self. Howell similarly holds that the brain is the material object to which our self-experience refers, and argues that it lies just in the intentional structure that the bearer of the intentional experience cannot itself appear as an object as this would precisely hinder the self to be intentionally directed at something else.

In my view, all these attempts to construe self-experience as something that may make us aware of our existence but not of our being, i.e., our nature, must fail. They all introduce a certain duplication of self:

- the kind of *experiential being* that I am insofar as I experience myself in experience as the one who is undergoing different experiences, and
- the *alleged real self* that, on the elusiveness view, ought to be essentially different from the experiential self.

Such a dualism creates an unbridgeable hiatus to the effect that it becomes throughout unclear how the latter could possibly present the nature of the former, given the infelicitous implications, I argued, any ousting of selfhood's nature from the experiential domain will evoke. If one agrees that we have self-experience, therefore, the only way to make sense of selfhood and to avoid these circumstances is to construe selfhood in experiential terms. By identifying the self with unprojected consciousness, Evans has gestured towards such a direction: a lead I will follow up on in the next chapter.





## Part II



## Chapter 4

# Minimal Self and the Ontological Interpretation of Self-Experience

I can test a theory of the self in terms of my own experience, but I do not judge the suggested theory by comparing it with the theory of the self I know to be true. My judgement is not based on any theory at all: it is based on experience of being a self—quite another matter.<sup>351</sup>

The consciousness in which I am conscious of myself, is my consciousness, and put concretely, consciousness of myself and I myself (similarly in full concretion) are identical. To be a subject means to be in the way of being self-aware.<sup>352</sup>

The preceding chapters discussed three main forms of skepticism about self-experience and rejected all of them. In this chapter, I address an alternative take on self-experience. One which considers self-experience to be informative with regard to the metaphysical nature of selfhood, its being, and not just its existence or reality. Such a line of thinking is probably most prominently to be found in the phenomenological tradition, which I will now discuss. Thinkers such as Edmund Husserl, Martin Heidegger, Maurice Merleau-Ponty, Emmanuel Lévinas or Jean-Paul Sartre, to name just a few, despite the significant differences of their conceptions, share the view that selfhood is experientially constituted. And so, that the meaning of what selfhood could possibly be is to be found *within* experience. They all, more or less, reject the idea that the self is something that is really located outside and beyond experience and self-experience—and thus object to any attempt to define selfhood in a non-experiential way. The heterogeneity among their phenomenological accounts concerns particularly how they take experiential selfhood to be experientially constituted, i.e., how the experience of being oneself is phenomenologically described in an accurate manner, but also how to interpret the experiential constitution of selfhood in a more general ontological framework. The different attempts to carve out the phenomenology of selfhood have resulted in a plurality of proposals that emphasize certain aspects of self-experience and their relevance for the right understanding of selfhood: for instance, the role of other people, embodiment, temporality, world and its structure, but also the notion of being.

---

<sup>351</sup> Evans 1970, 31.

<sup>352</sup> Husserl 1973a (Hua 14), 151, my translation: “Das Bewusstsein, in dem ich meiner bewusst bin, ist mein Bewusstsein, und konkret gefasst ist Bewusstsein meiner selbst und ich selbst (ebenfalls in voller Konkretion) identisch. Subjektsein ist, in der Weise seiner selbst bewusst zu sein, zu sein.”

In the following analysis, I will focus on one specific account of experiential selfhood as a proxy for the phenomenological tradition: Dan Zahavi's "experiential self"<sup>353</sup> and the minimal self-view it is based on, which I already introduced in Chapter 2. His account is particularly interesting and lends itself to a close and focused discussion for several reasons. First of all, his account is probably the most—and also most controversially—discussed phenomenological view of selfhood in the current debates that involve all the different philosophical paradigms, research traditions, and scientific disciplines, such as analytic philosophy of mind, the cognitive sciences, hermeneutics, and the social sciences. Moreover, his account presents what is an interpretation of the phenomenological tradition as a whole, considering its different developments, arguments, and inner differentiations as well as disagreements. As a consequence of his rather synoptic perspective on the phenomenological tradition, he is less tied to a specific ontological framework but rather focuses on the experiential aspects of selfhood. Crucially, despite his focus on the experiential aspects of selfhood, his account is meant as an approach to what selfhood *is*. This more ontological claim about selfhood, however, is often overseen and not sufficiently understood in the debate on the metaphysical nature of selfhood, notably in the philosophy of mind. One reason is that Zahavi himself does direct only relatively little attention to these issues, while focusing more on the issues of the phenomenological structure of phenomenal consciousness, self-consciousness, and their structural entwinement. Consequently, in the philosophy of mind, his minimal self-view is predominantly received as a *merely* phenomenological theory of self-experience and its role for experience in general. By contrast, without taking into account the ontological claim of the minimal self-view, one risks to misunderstand also its exact phenomenological point regarding phenomenal consciousness and self-consciousness. A clarifying discussion of how Zahavi's theory of phenomenal consciousness and self-consciousness is by the same token a metaphysical approach to the being of selfhood and not *just* its phenomenology is urgently needed. As it will turn out, Zahavi's ontological interpretation of self-experience amounts to a very particular understanding of self-experience, according to which self-experience is not only the way by which we gain insight about what selfhood is, but also the way of being of selfhood proper. Put differently, the ontological claim about the minimal self, to be carved out in this chapter, consists in the view that self-experience is both the *modus cognoscendi* and the *modus essendi* of selfhood. I will suggest to call this *self-experientialism*<sup>354</sup>. My aim in this chapter is to reconstruct this position, explain its key

---

<sup>353</sup> Zahavi 2011b, 2014.

<sup>354</sup> Similar and related ideas can, for instance, be found in Evans 1970, Fasching 2009, 2011, Strawson 2017, Ch. 17. The approach that I develop here on the basis of Zahavi's approach, however, makes a specific claim, namely, that it is the reflexivity of experience—pre-reflective self-awareness, which amounts to an experience of reflexivity—that constitutes the most original meaning of 'self'. *That* is to which all talk of self traces back. And, as I will try to show, that is the rationale for self-experientialism to hold that self-experience constitutes selfhood. The phenomenological self-experientialism that I develop here, thus,

claims, and bring afore how it is motivated and preceded by a phenomenological analysis of a rather uncontroversial kind of self-experience: reflective experience.

#### 4.1 The three conceptual roles of the minimal self

Zahavi's minimal self-view, which he has been defending for already more than twenty years,<sup>355</sup> is widely received and discussed in the current debates in the philosophy of mind, broadly construed, but also in the empirical sciences, notably the cognitive sciences. Typically, presentations of his position focus on the claim that “*all* my conscious experiences include a minimal form of pre-reflective self-awareness, in that they are all non-reflectively experienced as *mine*”, that is, that they “are always and necessarily given to me in the *first* person, immediately and directly”<sup>356</sup>, in short, that “all conscious experience involves self-consciousness”<sup>357</sup>. Accordingly, it is often also stressed that the minimal self-view provides an account of “a sense of self intrinsic to any phenomenally conscious state”<sup>358</sup> or “a primitive form of self-awareness that is integral to all experiences”<sup>359</sup>. Some take the claim to be “that phenomenal consciousness entails minimal self-consciousness”<sup>360</sup>. Yet, other describe it as “all conscious states are characterized by an inner awareness”<sup>361</sup> or “[s]elf-consciousness is an invariant, always present, ubiquitous feature of our consciousness”<sup>362</sup>.

While all these descriptions, which serve only as few examples, capture well what are essential aspects of the minimal self-view, they do not exhaust the whole scope of it. For, as Zahavi is adamant, his approach is not only an account of experience and the role

---

involves the claim that any understanding of self and its meaning is bound to and initially motivated by experience. A related though different application of the term ‘experientialism’ has been associated with the view that “for a subject to have a SBO [sense of bodily ownership, P.S.] is for her to be aware of (A) the experience-dependency of the properties involved in the content of somatosensory experiences; and (B) the relevant experiences as being her own” (Serrahima 2019, iii). Lakoff & Johnson 1980 and Lakoff 1987 also have proposed to call their philosophical position ‘experientialism’. Although their understanding of experientialism might bear some resemblance to what I develop here under the label of self-experientialism (building on the phenomenological tradition), these similarities in some points are entirely accidental. Cf. Lakoff 1987, 14, 267, 344.

<sup>355</sup> The main arguments are to be found in his books Zahavi 1999, 2005, 2014. For more recent elaborations, see e.g., Zahavi 2017b, 2018, 2020b, Zahavi & Kriegel 2016.

<sup>356</sup> Colombetti 2011, 303.

<sup>357</sup> Schear 2009, 95.

<sup>358</sup> Bortolan 2020, 68.

<sup>359</sup> Ratcliffe 2017, 3.

<sup>360</sup> Grünbaum 2012, 274

<sup>361</sup> Farrell & McClelland 2017, 5.

<sup>362</sup> Kiverstein 2020, 559.

of self-experience in it, but also of selfhood as essentially experiential.<sup>363</sup> Accordingly, the “minimal self” as a notion has three conceptual roles<sup>364</sup> to play that, despite their interrelatedness, need to be well distinguished as well as do the philosophical claims that are associated with these roles.

1. Conceptual role [*phenomenal consciousness*]:

Provide a phenomenologically adequate description of what is required for phenomenal consciousness of  $x$  (experience) to be possible, i.e., elucidate the essential structural features of experience.

2. Conceptual role [*self-consciousness*]:

Provide a phenomenologically adequate description of what is required for self-consciousness (self-experience) to be possible, i.e., elucidate the essential structural features of self-experience.

3. Conceptual role [*selfhood*]:

Provide a theoretical account of what selfhood essentially *is*, i.e., to clarify what is minimally required for selfhood to be prevalent or real.

Evidently, all the descriptions of the minimal self-view above—which reflect the dominant reception of the minimal self-view—refer only to the first two conceptual roles insofar as they address issues with regard to the phenomenology of experience and how self-experience features in it. They leave out the question what selfhood essentially *is*, i.e., its wider metaphysical implications. Yet, it is not clear whether a separate treatment of phenomenological and metaphysical issues is always possible—let alone if at all. When it comes to the perception of some object in the world, for instance a bottle on the table, it seems possible to focus on the phenomenology only and refrain from any ontological judgement about whether the bottle is really there and what it is independently from my experiencing it. Phenomenological issues, it seems, in this case, can be addressed without further engaging in metaphysical inquiries. Such a view is paradigmatic for the distinction, often operative in the current philosophy of mind, between phenomenology and metaphysics. However, in case of selfhood, it is far from clear that such a bifurcation of

---

<sup>363</sup> See particularly Zahavi 2014, Ch. 2.

<sup>364</sup> Note that with ‘conceptual role’ I don’t mean to refer to the importance of the minimal self for the development of concepts and conceptual mastery, which would be a topic on its own. Rather, I mean that the theoretical *concept of the minimal self* has three roles in the philosophical debate in that it features in the understanding of three distinct though entwined phenomena or topics: phenomenal consciousness, self-consciousness, and selfhood. Cf. Clowes & Gärtner 2020, 625.

issues goes through as clinically sterile as some might think. For, asking whether all experience involves a sense of self, doesn't this require that one at least has some idea what selfhood is, what one is looking for when examining the phenomenology of experience in search of self? Trivially, whether a sense of self is intrinsic to experience depends on the notion of selfhood. And yet, the specific notion of selfhood and its ontological implications is often ignored when Zahavi's minimal self is discussed. This, however, is problematic when his account is being rejected based on what is a stunted and distorted understanding of the scope of the minimal self-account.

Zahavi doesn't seem to be completely innocent in this. At least, if we accord credibility to Barry Dainton who states that his ontological commitments are not clear. For, as Dainton argues, "while there are passages where Zahavi does seem to be offering an account of what selves *are* ... on other occasions he can be construed as merely arguing that all our experience is accompanied by *me-ish* quality"<sup>365</sup>. In fact, while Zahavi does allocate efforts to analyzing the phenomenological structure of phenomenal consciousness (1. conceptual role) and of self-consciousness (2. conceptual role), remarks of the following kind that reveal his ontological interpretation of self-experience are rather sparse (3. conceptual role):

[E]xperiential selfhood should be defined in terms of self-experience rather than vice versa.<sup>366</sup>

Ultimately, Husserl's thesis concerning the existence of a prereflective self-awareness is connected to a general claim concerning the *being* of subjectivity. To be a subject is to exist for-itself, that is, to be self-aware.<sup>367</sup>

[T]he phenomenological notion ... strikes me as being of pivotal significance. It is fundamental in the sense that nothing that lacks this dimension deserves to be called a self. Thus, in my view, this experiential sense of self deserves to be called the *minimal self* or the *core self*.<sup>368</sup>

As a consequence of this asymmetrical treatment of the different roles that ought to be played by the "minimal self", it is often overseen that the minimal self-view in fact amounts to a theory of the metaphysical nature of selfhood and not *only* its appearance in phenomenology—as if these two issues were orthogonal to each other.<sup>369</sup> That said, it doesn't go completely unnoticed by his critics either. Marie Guillot, who has delivered what

---

<sup>365</sup> Dainton 2008, 145 Fn. 5.

<sup>366</sup> Zahavi 2014a, 74. See also Zahavi 2005, 106.

<sup>367</sup> Zahavi 1999, 62.

<sup>368</sup> Zahavi 2005, 106

<sup>369</sup> Zahavi is well-aware of the fact that taking selfhood to be essentially experiential is a metaphysical thesis insofar as it is a thesis with regard to what the being of selfhood consists in.

is probably the most impactful criticism of Zahavi's minimal self-view in the recent years, recognizes in her influential paper "*I Me Mine: on a Confusion Concerning the Subjective Character of Experience*" that Zahavi endorses in fact three theses:

- (1) "an *epistemic* thesis (we have an awareness of the self in having experiences)"
- (2) "a *phenomenal* thesis (this access to the self is of the phenomenal kind)"
- (3) "a *metaphysical* thesis (this 'self-experience' is the self, or at least a form of selfhood)"<sup>370</sup>.

Crucially, she denies that Zahavi has sufficiently shown how these different claims could be launched and how exactly they interrelate. She claims the reason for this is that he had failed to distinguish between different aspects of the subjective character of experience and so didn't reach far beyond the "uncontroversial observation that experiences are special to their subject, in that they are an object of 'first-personal' access to her"<sup>371</sup>. Put differently, what she casts doubt on is that Zahavi has sufficiently demonstrated that experience in fact involves a sense of self and that such a sense of self allows for a metaphysical theory of selfhood according to which selfhood is essentially experiential selfhood. I would grant that the third conceptual role of the minimal self is (and, importantly, also in its significance for the first two conceptual roles) underdeveloped in Zahavi's otherwise outstanding and comprehensive work. But I believe Guillot (and with her many interpreters of Zahavi) misses a point when she takes Zahavi to attempt to establish the phenomenal and metaphysical thesis "simply on the basis of our noticing that there is something it is like *for us* to experience"<sup>372</sup>, as she concludes her verdict. I think the verdict is in need of a reassessment and I will try to deliver one in the following. Notably, doing so will allow me to reconstruct the minimal self-view as a metaphysical theory of selfhood and elucidate the ontological interpretation of self-experience that, as I further want to develop, amounts to an alternative to the skeptical views about self-experience.

## 4.2 On the 1<sup>st</sup> conceptual role: phenomenal consciousness

Before I turn to Guillot's criticism that concerns the 2<sup>nd</sup> and 3<sup>rd</sup> conceptual role of the minimal self, let me provide a sketch of how the notion of the minimal self is said to fulfil the 1<sup>st</sup> conceptual role. This can be brief, as I have already described how the minimal self-view affects and determines the understanding of phenomenal consciousness in Chapter 2.

---

<sup>370</sup> Guillot 2017, 49.

<sup>371</sup> Guillot 2017, 49f.

<sup>372</sup> Guillot 2017, 50.



In Chapter 2, I introduced Zahavi's notion of the minimal self as a concept that presents a voice against the Hume-inspired No-Self-Picture. The point, departing from an argument by McGilvary against the view that experience can be exhaustively described by *object*-experience, was that the experiential flow is manifest itself in the experience of an object. To provide a positive and phenomenologically sound account of such a self-manifestation of the stream of consciousness, Zahavi builds on Husserl's notion of inner time-consciousness. The central claim is that any phenomenally conscious state unfolds over time and that, whatever object is experientially given, it is so given insofar as it is present in a *temporal* manner. Phenomenal presence of *x* means that *x* is given *now*, whereby the phenomenological character of 'now' presupposes the temporal fringe that is also constituted by retentional and protentional consciousness. On Husserl's view, that means, any phenomenally conscious state has a *streaming* character implying that the stream of consciousness is itself manifest in experience insofar as it presents the form of experience of any kind whatsoever. Crucially, identifying this self-manifestation of the stream of consciousness with the minimal self, Zahavi can maintain that the pre-reflective self-awareness that is intrinsic to experience presents a kind of minimal self. And, notably, he can claim that the minimal self so-conceived is part and parcel of phenomenal consciousness, namely insofar as the subjective process of experiencing experiences itself. *This* is the way how the minimal self features in the understanding of phenomenal consciousness.

Most importantly, Zahavi interprets the pre-reflective self-awareness of experience not as an anonymous kind of self-experience, but as the "for-me-ness"<sup>373</sup> of experience that constitutes per se a form of "mineness"<sup>374</sup>. Put differently, he suggests to conceive of the pre-reflective self-awareness, which is phenomenologically manifest as the temporal self-givenness of the stream of consciousness, as *egological*. It is this move, he believes, that allows him to claim that even on its most basic level, phenomenal consciousness incorporates a sense of self.

However, in Chapter 2, I haven't discussed the reasons for this suggestion. I will do so in the following. For, his egological interpretation of pre-reflective self-awareness precisely concerns the 2<sup>nd</sup> and 3<sup>rd</sup> conceptual role of the minimal self, without which any understanding of the 1<sup>st</sup> conceptual role will remain incomplete. And, notably, it is precisely the egological interpretation that Guillot brands as an unwarranted maneuver.

---

<sup>373</sup> Zahavi & Kriegel 2016. Cf. Zahavi 2005, 122, Zahavi 2014, 19.

<sup>374</sup> Zahavi 2005, 124, Zahavi 2014, 19.

## 4.3 On the 2<sup>nd</sup> conceptual role: Self-consciousness and sense of self

Before I discuss Zahavi's and other arguments with regard to the egological character of the pre-reflective self-awareness that ought to play the 2<sup>nd</sup> conceptual role of the minimal self, namely to elucidate the most basic form of self-experience, let us hear out the accusations that are directed at him by Guillot.

### 4.3.1 Guillot: Zahavi's account is guilty of a confusion concerning 'subjective character'

In her impactful paper, Guillot utters dissatisfaction with the current debate on what has been termed the "subjective character"<sup>375</sup> of experience. More specifically, she is discontent with the heterogeneous and undifferentiated manner the terminus is being used. On her view, three notions need to be distinguished that all refer to the subjective character of experience in a slightly different way: *for-me-ness*, *me-ness*, and *mineness*. Failing to discern these, as she claims, will likely result in significant confusions that may undermine any argument concerning subjective character in a malignant way. This, she believes, is precisely the kind of shortcoming adhering to Zahavi's proposal, wherefore she takes the minimal self-view to be flawed and unable—at least as it stands—to buttress all three conceptual roles it purports to play.

How does she define the three notions of subjective character and in which sense has Zahavi failed to recognize their distinctness on her view?

<i>For-me-ness</i>	"the object of awareness is the experience itself" <sup>376</sup>
<i>Me-ness</i>	"the subject is somehow aware of herself" <sup>377</sup> , i.e., that the subject is "an object of phenomenal awareness" <sup>378</sup>
<i>Mineness</i>	"a phenomenal awareness that my experiences are <i>mine</i> " <sup>379</sup>

For Guillot, these three notions need not only be conceptually distinguished, they are also both conceptually and phenomenologically independent from each other. That means, neither do they mutually imply each other conceptually nor are they phenomenologically inseparable from each other. The latter, she argues, holds because *for-me-ness*, *me-ness*,

---

<sup>375</sup> "Subjective character" has been distinguished from "qualitative character". While the latter refers to the specific qualitative how of an experiential state (tasting chocolate vs. tasting chili pepper), the former refers to the fact that an experiential state has what-it's-likeness at all. Cf. Levine 2001, Kriegel 2009.

<sup>376</sup> Guillot 2017, 27.

<sup>377</sup> Guillot 2017, 29.

<sup>378</sup> Guillot 2017, 35.

<sup>379</sup> Guillot 2017, 31.

and mineness can come apart and be present in experiences independently, as she claims to have shown by emphasizing psychopathological conditions, such as depersonalization and schizophrenic thought insertion.<sup>380</sup> Consequently, although she admits most experiences involve all three phenomenal features,<sup>381</sup> she maintains that the presence of for-me-ness does not *per se* amount to the fact that me-ness and mineness were present too.

To wrongly infer that, of that precisely, however, she accuses Zahavi. Because, as she believes, he doesn't distinguish between for-me-ness, me-ness, and mineness, by simply demonstrating that for-me-ness is intrinsic to experience, he mistakenly believed to have shown that me-ness is involved too. That is, the accusation is that, while Zahavi might have successfully shown with Husserl that the stream of consciousness is self-manifest in experience, he has failed to clarify how the self-manifestation of the stream amounts to a sense of self (=awareness of the subject of herself). Since Guillot emphasizes for-me-ness does not conceptually imply me-ness, she holds, such a clarification would be needed.

She acknowledges that an inference from for-me-ness to me-ness could be warranted, namely in one specific case: “if the self *is* the experience (or a part or intrinsic feature thereof), then to be phenomenally aware of the experience (for-me-ness) is to be phenomenally aware of the self (me-ness)”<sup>382</sup>. In fact, as she further recognizes, Zahavi’s minimal self-view operates indeed with such a notion of self. However, as she highlights, this specific notion of self and the minimalist interpretation of selfhood that it embodies is “far from trivial” and needs “the additional support of a substantive extra premise”<sup>383</sup>. Crucially, then, as she, rightly in my view, carves out, to show how the minimal self plays the 2<sup>nd</sup> conceptual role, its 3<sup>rd</sup> conceptual role needs to be considered. That is, a reflection on the notion of selfhood and its metaphysical nature is required before one can assess whether the minimal self presents a form of self-consciousness in terms of me-ness:

Again, even on a minimalist approach, the implication from for-me-ness to me-ness depends on the minimalist view of the self being independently proved true. To be allowed to draw the implication, we need a separate defence of the thin metaphysics of selfhood. Instead of which, minimalists typically *assume* that the implication from for-me-ness to me-ness holds, because they often use an undifferentiated notion of subjective character to cover both aspects, and

---

<sup>380</sup> I already want to state here that I am not convinced by Guillot’s depiction of depersonalization. After having discriminated, with high effort, the three notions, she then goes on and interprets depersonalization as a case in which only for-me-ness is retained. However, she doesn't sufficiently demonstrate, in my view, that the expressions of persons so-affected and their experiences really match her “tripartite framework” (Guillot 2017, 44). As I will argue in Chapter 7, I think we must come to a different phenomenological conclusion. However, in this chapter, I will leave out the issue, as it is not necessary for the discussion of Guillot’s critique of Zahavi’s account and possible replies by the latter.

<sup>381</sup> Guillot 2017, 45f.

<sup>382</sup> Guillot 2017, 37.

<sup>383</sup> Guillot 2017, 37.

then go on to rely on the alleged conceptual connection to defend a thin theory of the self. This is getting it backwards, and the move is unwarranted.<sup>384</sup>

Whereas I believe Guillot is right in her reconstruction of the general plan that is needed to make the case for the minimal self-view (in its three conceptual roles), I am less persuaded by her depiction of Zahavi's move to consist in a merely conceptual transition from for-me-ness to me-ness. By contrast, the transition is motivated, as I want to show in the following sections, by phenomenological reasons of which many can be found also in Zahavi's work, albeit to see this an accurate understanding of the terminology operative in phenomenology and Zahavi's work will be required. Most importantly, it is far from clear Guillot has appreciated the particular terminological framework in which the minimal self-view is formulated. Evidently, as long as this is not clarified, Guillot's narrative of the latter's insufficiency will run the risk of rather adding to than reducing the confusion that prevails in the debate on the subjective character of experience.

#### 4.3.2 Terminological clarifications: Zahavi's response to Guillot

Not surprisingly, Guillot's critique did not pass Zahavi's attention. In his direct reply to her, he emphasizes that

[g]iven how Guillot is defining me-ness and mineness, I would also dispute that for-me-ness entails either [me-ness or mineness, P.S.]. Being aware of one's experiences when they occur is neither tantamount to being aware of oneself as a (secondary) object, nor equivalent to being thematically aware of the experiences as one's own.<sup>385</sup>

As this quote shows, Zahavi doesn't disagree with Guillot's verdict, but he doesn't find it applies to his own proposal. The point is, he highlights, although he equates for-me-ness with mineness, when talking of mineness he is not, contrary to Guillot, referring to the awareness one has of one's own experience *as* one's own experience. Rather, he uses mineness "interchangeably"<sup>386</sup> with for-me-ness to the effect that transiting from for-me-ness to mineness no inference or any further thesis is performed. The transition from for-me-ness to the phenomenon Guillot labels with 'mineness', according to Zahavi, requires an act of reflection and is as such certainly not taken by him to be identical with the pre-reflective self-awareness that amounts to for-me-ness.<sup>387</sup>

---

<sup>384</sup> Guillot 2017, 37.

<sup>385</sup> Zahavi 2018, 706.

<sup>386</sup> Zahavi 2018, 706.

<sup>387</sup> Zahavi 1999, 142–152.

But how about me-ness? Isn't Zahavi's claim precisely that in being self-aware the subject of experience is aware of itself? And does the pre-reflective self-awareness of experience (for-me-ness) amount to the self-awareness of the subject (me-ness)?

Again, the problem lies with Guillot's description of me-ness, when she speaks of it as a case in which the subject is the '*object* of phenomenal awareness'. For, Zahavi, quite explicitly, denotes pre-reflective self-awareness (for-me-ness) as "non-objectifying"<sup>388</sup>, i.e., being pre-reflectively self-aware of itself consciousness doesn't take itself as an object. The self-objectification occurs only when one reflects and explicitly directs one's attention at one's own experiences.<sup>389</sup> Given that this is so, Zahavi can easily and rightly reject Guillot's accusation. In *this* sense, he never claimed for-me-ness entails me-ness.

But a more charitable reading of Guillot is possible, as she originally defined me-ness as an awareness in which 'the subject is *somehow* aware of herself'. Hence, she is not focused on self-awareness as a kind of object-consciousness and allows for any sort of awareness however conceived. That would include the specific way Zahavi suggests selfhood to be real in any kind of experience, namely as "first-personal *givenness*"<sup>390</sup>. Crucially, Guillot's demand for an account for the transition *from for-me-ness to me-ness* or *from phenomenal consciousness to first-personal givenness* or *from pre-reflective self-awareness to egological pre-reflective self-awareness* remains ultimately unharmed by her lapse to have ascribed to Zahavi the view that any experience involves the awareness of the subject as an object.

Zahavi is well-aware of that and complements his response with a few arguments in favor of the "deflationary or thin notion of self" which denies "that the self and experience are distinct particulars"<sup>391</sup>. For, it is only when *self* and *experience* are the same particular that *self-experience* per se amounts to me-ness. That is, only when self and experience are inseparable from each other is it the case that the pre-reflective self-awareness of experience presents a form of *egological* pre-reflective self-awareness. Put differently, on the pre-reflective level, I can only be aware of *myself* in being aware of my *experience* when I myself and my experience are intrinsically interwoven (= not distinct particulars).

However, the arguments Zahavi gives, at least those in his response to Guillot, do not suffice to fully launch the thesis that selfhood is essentially experiential (and that self and experience are not distinct particulars). For, they are primarily oriented against accounts that take self and experience to be *entirely* distinct. To give an example:

---

<sup>388</sup> Zahavi 2007a, 276.

<sup>389</sup> Zahavi 2007a, 276.

<sup>390</sup> Zahavi 2005, 106.

<sup>391</sup> Zahavi 2018, 707.

There is something counterintuitive to the claim that the subject, self or I is entirely non-experiential, such that I would remain a self even if I were zombified and ceased having experiences, simply because a brain, a body or a living organism continued to exist.<sup>392</sup>

While I agree very much with the assessment, it doesn't knock-out the right kind of approach currently attacking the minimal self-view here, it seems to me. For, one could maintain that the self is not a particular distinct from experience, but still hold that experience doesn't require any aspect of selfhood. This, for instance, seems precisely to be the case in Sartre's famous non-egological theory of self-awareness which Zahavi has discussed in his works.<sup>393</sup> But it is arguably also the case in more recent accounts such as Bayne's interpretation of the self as an "intentional entity"<sup>394</sup> that has its reality only in experience. Crucially, while defining self in terms of experience, Bayne doesn't take self and experience as the same particular, given that he is "not convinced that *de se* representation must structure the experience of any possible subject of experience"<sup>395</sup>.<sup>396</sup> That is, when self is present it might be inseparable from experience and thus experiential throughout, but that doesn't mean by itself that experience was inseparable from self.

Hence, more is needed to dissolve the rumors in the debate—quite successfully spread by Guillot's paper—that the minimal self-view hasn't yet found a way to secure its own argumentative basis.<sup>397</sup> Before I will address possible arguments in more detail, let me first briefly give an overview of the terminological decisions by Zahavi that we need to bear in mind, when discussing the minimal self-view. It is important to make sure everybody is on the same page.

<b>Guillot</b>	<b>Zahavi</b>
<i>For-me-ness</i>	Non-egological pre-reflective self-awareness (anonymous self-awareness of experience)
<i>Me-ness*</i> (subject is <i>somehow</i> aware of herself)	Egological pre-reflective self-awareness / for-me-ness / mineness
<i>Me-ness</i> (subject is object of awareness)	Reflective egological self-awareness
<i>Mineness</i>	Reflective egological self-awareness

As this overview shows, Zahavi did not fail to distinguish between the different phenomena and concepts Guillot describes. However, the way his distinctions work is not parallel to

<sup>392</sup> Zahavi 2018, 708.

<sup>393</sup> Zahavi 1991, 139ff.

<sup>394</sup> Bayne 2010, 292.

<sup>395</sup> Bayne 2010, 290.

<sup>396</sup> I will discuss Bayne's position in more detail in Chapter 5.

<sup>397</sup> Guillot's view that Zahavi hasn't delivered an argument for his notion of selfhood has inspired others to formulate similar doubts. Cf. McClelland forthcoming b, López-Silva 2019.

hers. This is so, for one, because he precisely doesn't believe that Guillot's for-me-ness, an anonymous, egoless self-awareness of experience, is possible. Moreover, given the phenomenological nature of his account, the rationale of his argument doesn't consist in a conceptual deduction starting from for-me-ness to me-ness\*. Rather, the phenomenological goal of his stance is to show that for-me-ness and me-ness\* cannot be separated to begin with. That is primarily a phenomenological claim and the conceptual choices follow what is taken as the phenomenal facts. These, as described by Zahavi, shall be elucidated now.

#### 4.3.3 A phenomenological model: Three levels of experience and egocentricity

Although Zahavi's defense of his experiential notion of selfhood is spread over numerous papers and books, when it comes to his key argumentation with regard to the egological character of pre-reflective self-awareness, it is his book *Self-Awareness and Alterity* (1999) that deserves special attention. Unfortunately, and surprisingly, it is hardly if at all recognized in the more recent debate on the subjective character of experience in which Zahavi's account is being discussed. However, it is precisely there that he provides arguments that pose serious challenges to those doubtful about the egological character of pre-reflective self-awareness. My aim, therefore, is to focus on the argumentation Zahavi gives in that book and to reconstruct and re-introduce it in the debate. Moreover, as it will turn out, it provides the right background for the discussion of the metaphysical aspect of the minimal self-view (3<sup>rd</sup> conceptual role).

The systematic point of departure of the argument is the—*relatively* uncontroversial—phenomenological notion of egological consciousness. Let me illustrate it by describing a fictitious experiential episode that most of us will find easy to identify with:

I go through a street and pass by the parked cars of which I do not take further notice other than avoiding to touch them with my huge bag on the back. Without thinking about it—I am thinking about a friend I am looking for and who I am meeting for dinner—I try not to scratch the cars with the zipper of my bag. I keep walking and try to cross the street and look left and right for approaching cars. I barely notice that a bird is flying over my head only one or two meters away. I reach the opposite side of the street and see my friend around hundred meters in front of a tailor. I remember she told me she wanted to get her trousers done when we were sitting in the café this morning. I hurry so I wouldn't miss her.

According to Zahavi, three “levels of egocentricity”<sup>398</sup> can be distinguished:

- 1) The act-transcendence of the ego
- 2) The ego as a principle of focus
- 3) The egocentricity of first-personal givenness.<sup>399</sup>

I will now explain each of the level on the basis of the example.

*1) The act-transcendence of the ego*

According to Zahavi, certain experiences, such as imagination, recollection or reflection, are characterized by a “fission of a rather peculiar kind”<sup>400</sup>. In the example, if I recollect my friend telling me this morning that she wanted to get her trousers done, my experience involves a form of duality: the act of recollection and the presentiated act of listening to my friend in the morning that is recollected. Crucially, both acts are experienced as being performed by the same identical ego: *I* remember how *I* was listening to my friend telling me about her trousers; the “duality, difference, or distance” is “bridged”<sup>401</sup>. For Zahavi, these are cases of “truly egological self-awareness”<sup>402</sup> in which an awareness of the identity of the present consciousness with a presentiated consciousness exists. This is the kind of awareness Guillot labels mineness and which for Zahavi is only possible in the context of a presentiating act that requires some form of reflection.

We can only speak of an experience as being owned by an ego if we operate with a difference between the experience and the ego, and we only need to do that when we realize that the ego retains its identity through different experiences. But in order for that realization to occur, it is necessary to relate and compare different experiences, and this is exactly what takes place in presentiating acts.<sup>403</sup>

Evidently, this is a form of awareness that we can easily generate ourselves when we reflect on our past experiences. When we do so, most will agree, we find that we, the reflecting consciousness, are identical with the consciousness of a past experience, i.e., the consciousness being reflected on. In engaging in such a kind of reflection, we have egological consciousness.

---

<sup>398</sup> Zahavi 1999, 138ff.

<sup>399</sup> Zahavi presents the three levels in the opposite order, following the foundational order of the levels. However, here, my interest is primarily with the *argument* in favor of that foundational order. And, as I take it, it departs from the more uncontroversial phenomenon of reflective consciousness and descends or regresses from there to the more basic levels of experience.

<sup>400</sup> Zahavi 1999, 150.

<sup>401</sup> Zahavi 1999, 150.

<sup>402</sup> Zahavi 1999, 150.

<sup>403</sup> Zahavi 1999, 150.



## 2) *The ego as principle of focus*

The above-mentioned example, however, does not only comprise reflective and recollecting experiences. After all, I am hurrying and eager to find my friend and so I attend only specific objects in my perceptual environment. My focus is on some of them rather than on others. Those I focus on are given to me in more detail. For instance, I see that the shop my friend stands in front of is a tailor, whereas I do not notice what kind of shops are near the tailor or which color the cars have that are parked somewhere in the middle between my current location and that of my friend. As Zahavi emphasizes, “that which we are conscious of does not present itself to consciousness as a ‘flat surface’ but in a relief”<sup>404</sup>. How an object is experienced by us depends partly on whether we focus our attention on it or not. This attentional directedness that is manifest in experience and characterizing our experience of the world presents, as Zahavi emphasizes with Husserl, a form of egological consciousness. For, it is precisely in this attentional directedness that our experience of the world is revealed *as being centered around someone*: the experiencer.<sup>405</sup>

## 3) *The egocentricity of first-personal givenness*

According to Zahavi, there is a third level of experience. It is characterized by the fact that the ego remains rather passive and doesn't show the activity that pertains to the other two levels in which we encounter the ego in its reflective or attentional activity. Consider the experiences in the example that form rather the experiential background: the parked cars that I circumvent without further notice so I wouldn't scratch them or the short noise and the soft blast of wind made by the bird above my head. My attention is with finding my friend and I am seeing into the distance to detect her. Hence, I don't engage in a great part of my perceptual environment that is nonetheless experienced.

The decisive question then is whether the ego is also involved at this third level even though not actively engaging in the experience? In the example, the experiences—or experiential moments (cars, bird)—that are not reflected on or attended do still form part of an experience that is egological, as they are the background of attentional or reflective activity of the ego. However, we can also conceive of more minimal experiences that do not occur in the context of any attentional and reflective egological activity. For instance, very dim feelings of warmth or unspecified and undetermined experiences of red. Now, according to Zahavi's minimal self-view, even these experiences—by virtue of their first-personal givenness—are egological though they are admittedly not characterized by attentional or

---

<sup>404</sup> Zahavi 1999, 148.

<sup>405</sup> Some have argued that “attention precedes self in the explanation of what it is to be human” (Ganeri 2017, 4), suggesting that attention does not yet imply selfhood. See Zahavi 2017a for a response.

reflective activity of the ego. Let me now shed light on his main lines of argument for this view.

#### 4.3.4 Phenomenological arguments: That pre-reflective self-awareness is egological

To understand the rationale of the arguments Zahavi offers, it is important to acknowledge their phenomenological nature. Most importantly, the phenomenological nature does *not* mean that the task was to simply read off the phenomenal components that are to be found in pre-reflective experiences. After all, these experiences are *pre*-reflective and this alone prevents the possibility of simply detecting their phenomenal composition, because any detecting of such kind would be a reflective endeavor. That the arguments in favor of the minimal self-view are phenomenological, here, means that we depart from a phenomenological description of experiences that are less hard to grasp and then ask how these are phenomenologically possible. As I have mentioned, the point of departure for Zahavi is the relatively uncontroversial: egological consciousness ascribed to level 1 and 2. The phenomenological task that follows is to examine how such an egological consciousness is possible and whether carving out the possibilities will help to describe level 3, which seems to be more basic than the others. That is the first main line of argument to which Zahavi has devoted an extensive number of publications. The second main line of argument refers to the phenomenon of intersubjectivity, the experience of others. Again, the phenomenological task is to carve out how an experience of others is possible and what is required on the end of the experiential self in order to be capable of recognizing another self.

##### *1) Argument from the possibility of reflective experience*

Let me begin with reflective experience and what is required for it to be possible. In Chapter 2, I argued that a temporal horizon is needed for any object-consciousness to be possible. The same applies for reflective experience, in which one's pre-reflective experience is taken as an object, as Zahavi lets us know:

Ultimately, reflection presupposes the constitution of a temporal horizon. It would be impossible without the ecstatic unity of the flow. When reflection sets in, it initially grasps something that has just elapsed, namely, the motivating prereflective phase of the flow. The reasons why this phase can still be thematized by the subsequent reflection is that it does not disappear, i.e., it is

not cut off from the Living Present but remains united with it through retention.<sup>406</sup>

But reflective experience compared to any perceptual object-experience is peculiar. In it, as I mentioned when describing the experience of presentiating acts at level 3, we have a consciousness of an *identity* which presents egological consciousness in the strict sense. It is the consciousness of the identity between the reflecting consciousness *R* and the reflected consciousness *r*. For, “it is not sufficient that the act in question is reflexively thematized and made into an object”<sup>407</sup>. It must also be “grasped as being *identical* with the thematizing act”<sup>408</sup>. That means that *r* is experienced as belonging to the same experiential process in which *R* takes place. Recollecting the past experience of getting a haircut and reflecting on what it was like to undergo this procedure, *I*, the reflecting consciousness, identify myself with the consciousness that underwent the procedure. And this is where Zahavi’s argument in favor of an egological interpretation of pre-reflective self-awareness kicks in. The identification between *R* and *r* in reflective experience can only occur because both experiences are unfolding in the same temporally structured stream of consciousness and so are united. That is, the reflection is only possible because I have already been aware of *r* as forming part of *this—my—stream of consciousness*. When turning my reflective gaze *R* towards *r*, I don’t learn for the first time that I have been undergoing *r* but acknowledge—now more explicitly—that I have been undergoing *r*. Thus, the egological interpretation of the temporally understood pre-reflective self-awareness explains how a transition from a pre-reflective experience *r* to a reflective experience *R* which is directed at *r* is possible. The suggestion is that we need to conceive of this transition as one from an *implicit* egological to an *explicit* egological consciousness. Egological consciousness in reflection, on Zahavi’s conception, therefore, is not a *product* of reflection, but rather a “thematization”<sup>409</sup> of an aspect that is otherwise tacit in pre-reflective consciousness. Finally, then, Zahavi’s position is that we need to acknowledge that even pre-reflective experiences have *for-me-ness* and are egological (even though the ego doesn’t explicitly show up as attentive or reflective agent)<sup>410</sup> because otherwise we wouldn’t be able to explain how a transition from a pre-reflective experience to reflective experience could be conceived of.

The position derives its dialectical force to a great part from what Zahavi takes to be the incapability of the opposite position to provide an explanation of precisely that transition. Non-egological theories of self-awareness operate with a notion of pre-reflective

---

<sup>406</sup> Zahavi 1999, 79.

<sup>407</sup> Zahavi 1999, 17.

<sup>408</sup> Zahavi 1999, 17. Cf. Hicks, Laird, and Dorward 1928, 196.

<sup>409</sup> Zahavi 2005, 52.

<sup>410</sup> Cf. Zahavi 1999, 152.

experience that ought to be anonymous and only become egological *through* reflection. That is, these theories argue that pre-reflectively, given that no ego shows up, experience is egoless throughout. On such a line of thinking—Zahavi usually mentions Aaron Gurwitsch and Jean-Paul Sartre as proponents,<sup>411</sup>—it is only in the process of reflection, that is, when an explicit egological consciousness emerges, experience can be legitimately called egological too. But the problem is: How can an egological consciousness, on such a view, come about to begin with? How is it possible that reflective consciousness *R* identifies with *r* whereby *r* is supposed to be anonymous? On which basis is the identification performed? How could such an *appropriation*, which takes the place of Zahavi's *thematization*, occur? After all, the pre-reflective experiences ought to belong to no one. How does it become mine in reflection?

Even William James, who holds “the unity into which the Thought ... binds the individual past facts with each other and with itself, does not exist until the Thought is there”<sup>412</sup>, admits when he famously compares past experiences to the beasts and the ego with a herdsman: “No beast would be so branded unless he belonged to the owner of the herd. They are not his because they are branded; they are branded because they are his.”<sup>413</sup> The point is when a reflecting act turns its gaze onto a past experience, in doing so it doesn't *make* the past experience its own experience. Rather, in reflection one finds the past experience as one's own experience: “The Thought does not capture them, but as soon as it comes into existence it finds them already its own.”<sup>414</sup> Although James attempts to give an account of an appropriation of past experiences through an arising Thought (a reflective consciousness of past experiences) at a given time,<sup>415</sup> he must admit that “[t]he only point that is obscure is the *act of appropriation* itself”<sup>416</sup> and ultimately also has to concede that the “present moment of consciousness” “may feel its own immediate existence”<sup>417</sup>. This, at least, in line with the minimal self-view, is much more compatible with what seems to be a phenomenal fact: that “[e]ach thought is thus born an owner”<sup>418</sup>.

Finally, to conclude, two major challenges emerge for a non-egological theory of pre-reflective self-awareness. It must tell a convincing story why it is that

---

<sup>411</sup> Cf. Zahavi 1999, 139ff.

<sup>412</sup> James 1950/1890, 338

<sup>413</sup> James 1950/1890, 337.

<sup>414</sup> James 1950/1890, 338.

<sup>415</sup> See McGilvary (1907) for a critical and, as I find, convincing discussion.

<sup>416</sup> James 1950/1890, 340.

<sup>417</sup> James 1950/1890, 341.

<sup>418</sup> James 1950/1890, 339.

- (1) when we reflect on an experience *r*, we already take this experience to belong to the reflecting consciousness *R*, and are aware of an identity of the consciousness to which both acts belong [identity = ego as act-transcendent pole];
- (2) when we reflect, we never become witness of any process of appropriation [appropriation doesn't lie comfortable with the phenomenology of reflection].

However, chances of finding such a story that is plausible are, I side with Zahavi, very low. Consider a critique by Joseph K. Schear who compares reflection with opening the refrigerator and stating that the light is on. As we all know, it is wrong to conclude from the light being on that it is also on when the door is closed. The same fallacy, on his view, applies to Zahavi's minimal self: "But just as it doesn't follow from the light being on when we open the refrigerator that the light is always on, so it doesn't follow from our being able to report knowingly on our conscious lives when asked that our conscious lives always includes self-consciousness."<sup>419</sup> Yet, abstracting from the fact that the metaphor is misplaced,<sup>420</sup> what such an objection implies is that pre-reflective experience is taken to be throughout different from reflective experience. In fact, Schear explicitly states this when he asks why Zahavi doesn't acknowledge the possibility that reflection "might bring on a transformation of conscious experience—a shift of mental posture—rather than merely trigger the revelation of what had always been at work?"<sup>421</sup>

But this is exactly the point. The phenomenological model of the three levels presented above suggests there *is* a transformation from pre-reflective experience to reflective experience. And the argument is that, for such transformation to be possible, we need to conceive of pre-reflective self-awareness already as implicitly egological. Thus, the proposal is that the transformation is an articulation—which also acknowledges the phenomenological difference between pre-reflective and reflective experience. That is, how the ego is involved in experience may differ heavily, and in reflective experience the ego is explicit, while it remains implicit on the pre-reflective level. If one wants to doubt that the ego is implicitly involved on the pre-reflective level, then one would have to show how an egological reflective experience can emerge out of a throughout non-egological pre-reflective experience. But as long as there is no compelling case which could elucidate an alleged transition from a completely non-egological pre-reflective experience to an

---

<sup>419</sup> Schear 2009, 101.

<sup>420</sup> There is no such conclusion in Zahavi's approach. The idea is not that since every time I look (reflect) that I find myself being self-conscious that I assume it is always the case, even when I don't look (reflect). Rather, the idea is that for being able to look (reflect), I must have been self-aware in the first place. That is a whole different story and so the metaphor seems ill-chosen.

<sup>421</sup> Schear 2009, 101.

egological reflective experience, I find myself convinced Zahavi holds the better cards in his hands.<sup>422</sup>

## 2) *Argument from the possibility of social experience*

Let me now turn to another main line of argument that has a similar rationale. Here, the phenomenological explanandum is social experience, the experience of another experiential self. Again, the question is how it is possible that we can experience others. How must consciousness be structured prior to its encounter with others so that the other can be experienced as other in the first place?

As Zahavi emphasizes, Husserl originally had defended a non-egological theory himself but encountered difficulties “when it came to a phenomenological analysis of intersubjectivity”<sup>423</sup>. The difficulties that emerged in that context refer to the problem that if experience were non-egological the encounter with an other would not give rise to an *experience* of intersubjectivity. To see this, we need to bear in mind how Zahavi understands intersubjectivity:

Intersubjectivity is a relation between subjects; it is a subject-subject relation. But for me to relate to another as subject is for me to relate to somebody with a first-person perspective of his or her own. We encounter others as such when we encounter them as experiencing subjects, and this means as subjects that have a perspective not just upon the world of objects, but upon us too.<sup>424</sup>

Essential to the experience of an other is that I experience her *as experiencing*. The question is, therefore, how I can possibly experience that there is someone else experiencing?

The first trivial insight is that to do so, I obviously need to somehow discriminate between the experiences at stake, that is, between the experience of the other and my own experience. If both kinds of experience were given on a par in the same manner, the other’s experience would immediately cease to be the experience of an other.<sup>425</sup> But wherein lies the phenomenological difference between my experiences and the experiences of the other? What marks their distinctiveness?

Zahavi points one more to Husserl’s analyses. Crucially, for Husserl as for many other phenomenologists, the experience of the other is characterized by two seemingly contradicting aspects:

- 1) [Experienceability]: On the one hand, perceiving the other’s behavior we experience it “as expressive of mental states that transcend the behavior that

---

<sup>422</sup> Cf. Zahavi 1999, 139–152, Zahavi 2007a.

<sup>423</sup> Zahavi 1999, 143.

<sup>424</sup> Zahavi 2014, 193.

<sup>425</sup> Cf. Zahavi 2014, 154.

expresses them”<sup>426</sup>. Hence, we have an experience of the other’s experience, but not in a direct manner. We experience her experiences from our perspective and not from hers.

- 2) [Inaccessibility]: On the other hand, and precisely because we don’t experience the other’s experience from her perspective, her experience is in a certain way inaccessible to us.

Both aspects, now, are what “makes the experience in question an experience of an other”, thus, the inaccessibility of the other’s experience “is not an imperfection or a shortcoming” but is “constitutional”<sup>427</sup>: “The otherness of the other is exactly *manifest* in his elusiveness and inaccessibility.”<sup>428</sup>

For an experience of the *inaccessibility of experiences* to be possible, a few things are required.<sup>429</sup> My focus here lies only on the aspects relevant to the issue of the egological character of pre-reflective self-awareness. The question is what is the role of egological pre-reflective self-awareness for the experience of the inaccessibility of the experience of the other?

The most decisive point in that regard is that the inaccessibility of the other’s experience is a contrast phenomenon.<sup>430</sup> As such, it is contrasted with something. What is this something? It is my own experience it is contrasted with. That means, first of all, that in order to be aware of other’s experiences, I need to be self-aware of my own experience.<sup>431</sup> Otherwise, I couldn’t contrast the other’s experiences with any other experience. Without being self-aware of one’s own experience, it would be inexplicable how any behavior of somebody else were experienced as expressive of mental states. The second aspect of the contrast is that unlike my own experience, the other’s experience is inaccessible. That means, that the experience of the inaccessibility of the other’s experience is determined by my own experience being accessible. It is against the experience of direct accessibility (or first-personal givenness, as Zahavi would have it) of my experiences that an experience of inaccessibility is possible. Thus, not only is it the immediate acquaintance with my own experience that distinguishes *my* experience from the experience of every other subject, it is only the immediate acquaintance that makes possible the peculiar kind

---

<sup>426</sup> Zahavi 2014, 155.

<sup>427</sup> Zahavi 2014, 154.

<sup>428</sup> Zahavi 2014, 155.

<sup>429</sup> As Husserl and—after him—all phenomenologists have emphasized, to experience another experiencing self, an *alter ego*, requires that the other appears in an embodied manner. It is the lived body of the other through which we can become aware of the other’s experiences in the first place. See, for instance, Taipale 2014, 79ff.

<sup>430</sup> Notably, the first-personal givenness of experience, its mineness “is an aspect that can be characterized without any contrasting others”, Zahavi 2014, 84.

<sup>431</sup> That makes it very hard for those who are eager to defend the transparency thesis (cf. Section 2.5.4).

of experience of somebody else's experiences. For, the inaccessibility of the other's experience is a secondary mode of experience derived from and contrasted with the accessibility of my own experience. Only against the experiential background of the first-personal givenness of my own experience can the elusiveness of the other's experience amount to an experience of the other. Finally, what distinguishes my experience from the experience of the other is that the former is given in a first-personal mode, whereas the latter is not.<sup>432</sup>

And it is from this, that Zahavi concludes that the first-personal givenness of experience already suffices to individuate my experience. No further feature is required. By being self-aware of my experiences I never mistake them for somebody else's.<sup>433</sup> And it is also in this sense that the pre-reflective self-awareness is egological because it provides the basis for any encounter with an *alter ego*.<sup>434</sup>

On Zahavi's view, pre-reflective self-awareness wouldn't be able to play such a kind of role if it were non-egological resp. anonymous: "If the perceiving subject is anonymous, so is the perceived other, and to try to reintroduce a plurality of subjects into this anonymous collectivity is hopeless."<sup>435</sup> The problem with an anonymous self-awareness lacking any sense of self is that it becomes unclear how it could explain the appearance of another self. How should an experience of an other arise out of a throughout anonymous consciousness. Zahavi, rightly in my view, emphasizes that endorsing a notion of "a fundamental anonymity prior to any distinction between self and other obscures what must be clarified", "it does not solve the problem of intersubjectivity, but dissolves it"<sup>436</sup>. But evidently, we *do* have an experience of others and we *are* related to others. The price for the non-egological theory, hence, seems quite high, as it fails to account for something that most, if not all, accept is the case.<sup>437</sup>

This can be further illustrated by an objection of the following sort: Couldn't it be the case that when distinguishing my and the experience of the other that "one of these experiences is *this* experience, occurring in *this* particular mental life, and the other is *that* experience, occurring in *that* other particular mental life"<sup>438</sup>? Again, the point would be that if that was the case, it would become incomprehensible how from there we could possibly end up with the experience of an *intersubjective* encounter. If that was the basis

---

<sup>432</sup> Zahavi 2014, 126.

<sup>433</sup> Zahavi 2014, 23ff. See also Chapter 6.

<sup>434</sup> Zahavi 2009, Zahavi 2020b.

<sup>435</sup> Zahavi 2014, 86.

<sup>436</sup> Zahavi 2014, 170.

<sup>437</sup> A similar argument can be found in a discussion between the British philosophers G. Dawes Hicks, John Laird, and Alan Dorward. See Hicks, Laird, and Dorward 1928, 196. More recently, despite other fundamental differences, such a view can also be found in Peacocke 2019, 136.

<sup>438</sup> Parfit 1987, 517.



of experience, how would one ever come to identify oneself with one of the mental lives and ascribe the other to somebody else? Moreover, even if one allowed oneself to describe the situation like Parfit does, wouldn't one also have to admit that *this* and *that* mental life respectively are given in very different ways? Isn't there a significant difference between a pain experience in *this (my)* mental life and the experience of an experience of pain in *that (your)* mental life? And how could one describe this asymmetry without referring in some way to the immediate, direct, first-personal or self-reflexive character of *this* mental life and the distinctive character of *that* mental life? Any attempt to conceptualize *this* and *that* mental life in a purely *de re*-kind of consciousness lacking any character of *de se* will fall short to account for the involvement that is felt by the experiencer of *this* mental life.

For all these reasons, I believe Zahavi is justified in the different labels that he uses to denote this involvement of the experiencer even on a pre-reflective level: the *for-me*-ness of experience, the egological character of pre-reflective self-awareness or the first-personal givenness. With all these descriptions, the point is always to emphasize that over and above the pre-reflective self-awareness nothing else is needed for an—albeit—minimal sense of self. Note, that this doesn't amount to the claim that on top of an allegedly anonymous pre-reflective self-awareness there is some specific ingredient *x* that makes it a minimal sense of self. Rather, the claim is that the most basic form of pre-reflective self-awareness that Zahavi had identified with the temporal self-givenness of the stream of consciousness just *is* a sense of self. Therefore, it would be wrongly described in terms of a non-egological form of consciousness to begin with. Coming back to the criticism of Guillot, we can now recognize a decisive point of disagreement in the debate. Guillot assumes the possibility of a non-egological kind of consciousness (which she somewhat confusingly terms *for-me*-ness) and then concludes that this entails no sense of self (or *me-ness*\*).<sup>439</sup> This involves two theses:

- (a) that a non-egological kind of consciousness is possible
- (b) that a non-egological kind of consciousness does not entail a sense of self.

Most importantly, Zahavi agrees with the latter. In fact, as I have tried to show in reconstructing his main lines of argument, once we take pre-reflective self-awareness to be non-egological or anonymous, the experience of reflective experience (egological consciousness) and the experience of intersubjectivity (which is also a form of egological consciousness) become impossible. But since we *do* have reflective and intersubjective

---

<sup>439</sup> Guillot's argument in favor of experiences that lack a sense of self (or *me-ness*\* and *me-ness*) is reliant on Alexandre Billon's (2016) conceptualization of the psychopathological phenomenon of depersonalization. Billon argues that in depersonalization a person's experience lacks subjective character. In Chapter 7, I argue against this interpretation and defend a what I suggest is a more differentiated view.

experiences, there is no motivation to assume (a). The question, therefore, arises as to whether Guillot hasn't introduced a wrong premise to begin with. Given that her phenomenological analyses of for-me-ness (as a form of non-egological) self-awareness are rather sparse and *therefore* alone unconvincing, it is dubitable whether for-me-ness is best described as lacking any sense of self. Finally, we must recognize, I believe, Guillot owes an alternative theory of how an egological consciousness (as found in reflection and social experience) can emerge out of an allegedly anonymous pre-reflective experience. Zahavi, emphasizing Husserlian phenomenology, by contrast, has offered a positive phenomenological account of how egological consciousness can indeed be accounted for if we grant pre-reflective self-awareness amounts to a sense of self or is acknowledged as already egological (even though the ego is not explicit). As long as this is the case, I find myself convinced that the dialectical balance will swing towards the side of Zahavi's phenomenological account of pre-reflective self-awareness as egological, i.e., as a consciousness that involves a minimal sense of self or amounts in itself to me-ness\*.

Let me now turn to some other challenges for such a view and further assess the power of Zahavi's notion of the minimal self. This, I shall argue, will also shed light on the 3<sup>rd</sup> conceptual role of the minimal self: to present a notion of selfhood and its being.

#### 4.3.5 Challenge I: McClelland and the elusiveness of mineness

In a recent paper, Tom McClelland argues that proponents of a theory like the minimal self-view, which he identifies as the “universal mineness thesis”<sup>440</sup>, need to find a way to overcome what he calls the “epistemic impediment”<sup>441</sup>. The challenge consists in an epistemic problem. On his view, a phenomenal property can only be isolated and demonstrated by a phenomenal contrast, i.e., comparing two phenomenal states, whereby one shows the phenomenal property, the other not. But the minimal self as an alleged phenomenal feature is *ex hypothesi* “contrast-resistant”<sup>442</sup>, because it ought to be present in any kind of experience. Hence, one cannot build contrast cases, whereby one phenomenal state lacks the feature of minimal self the other not. Accordingly, the minimal self taken as a phenomenal property is elusive.<sup>443</sup> As McClelland emphasizes, the contrast-resistance of elusive phenomenal properties doesn't necessarily mean they didn't exist. It only means that “*knowing* that a particular contrast-resistant property exists is problematic”<sup>444</sup>. Although

---

<sup>440</sup> McClelland forthcoming b, 3.

<sup>441</sup> McClelland forthcoming b, 6ff. Note that McClelland discusses in total four impediments for theories like the minimal self.

<sup>442</sup> McClelland forthcoming b, 6.

<sup>443</sup> A similar argument can be found in Schear 2009.

<sup>444</sup> McClelland forthcoming b, 8.

McClelland admits that we might be able to determine *that* different experiences share certain aspects, further argumentation is required to carve out *what* exactly it is they share. He goes on and applies this to the case of mineness and asks how do we know “that what they share is the sense of mineness. Why not, for instance, adopt the more modest view that experiences share for-me-ness [sc. in the Guillotonian, non-egological sense, P.S.]”<sup>445</sup>? Can the minimal self-view and the lines of arguments I presented above provide an answer? As a response, I want to make four points:

First, although phenomenal contrasts are an effective way of isolating and determining certain phenomenal properties, it seems bit of an oversimplification that all properties would have to be carved out like that. How about, for instance, the attempt to describe the experiential structures at work in very basic experiences such as the perception of a melody? If we can succeed in doing so, we have good reason to believe that the structures at work at such basic ways of being intentionally directed at a temporal object may also be required for more complex experiences.

Second, of course, this is precisely the line of argument followed by Husserl and Zahavi, as I have argued. If it is true that the temporal self-givenness of the stream of consciousness is even required for a very basic object to appear in the first place, then this is also a structural feature of all other experiences.

Third, should the temporal self-givenness of the stream of consciousness be understood in terms of a sense of self? Attempting to explain the possibility of reflective and social experience, the minimal self-view does provide an argument why we should interpret the temporal self-givenness of the flow as a sense of self.

Fourth, and that is a point that I find needs to be reassessed in the debate, I don't see why interpreting pre-reflective experience as non-egological and anonymous should be a ‘more modest view’. Once we agree that pre-reflective experience involves self-awareness, we can ask whether it amounts to a *sense of self* or whether it is best described as *anonymous*. Guillot and McClelland believe that construing for-me-ness in anonymous terms is harmless and innocent, whereas only the minimal self-view is in need for argumentation. I believe that is a mistake from the outset. Rather, *both* options are non-trivial phenomenological interpretations of pre-reflective experience that require justification. The sense of self as well as the anonymity is a phenomenal property that when claimed to exist must be argued for. However, neither Guillot nor McClelland, as well as others critics of the minimal self-view, as far as I know, have made a case for the non-egological interpretation.

---

<sup>445</sup> McClelland forthcoming b, 8.

### 4.3.6 Challenge II: Peacocke's creature at Degree 0

A more positive argument in favor of a non-egological theory in the recent debate can be identified in Christopher Peacocke's *The Mirror of the World. Subjects, Consciousness and Self-Consciousness* (2014) and his example of a "creature at Degree 0" (or in short 'D0') that serves as an "example of Sartrean nonpositional consciousness"<sup>446</sup>. The creature Peacocke describes "has a (partial) atlas of the world around a location, and may have a history of the world in the form of a time series of time-indexed atlases, specifying how the world was in the past"<sup>447</sup>. However, what D0 lacks is any *de se* content whatsoever. As such, it does not even "represent what is in fact one's current location as one's current location"<sup>448</sup>, as "[t]he location may be represented simply as *here*"<sup>449</sup>. On Peacocke's view, *here* and *now*, as the possible case of D0 demonstrates, are indexicals that do not require the involvement of the 'I'.<sup>450</sup> Although the notions of '*de se* content' and self-consciousness, as used by Peacocke, are not congruent, his description of D0 ultimately amounts to the view that such a creature would have an experience of the world without any self-experience whatsoever.<sup>451</sup>

Peacocke's argument in favor of D0 forms part of a theory very different and opposed to the minimal self-view.<sup>452</sup> It would be an interesting project to bring both accounts into discussion and possibly thereby connect still segregated though indubitably related strands of the debate on self. Here, however, my aim is limited and focused: to determine whether Peacocke's depiction of a non-egological D0 is phenomenologically plausible.<sup>453</sup>

Evidently, my preceding defense of the temporal self-manifestation of the stream of consciousness would also apply to D0: any experience of 'now' involves the manifestation of the stream of consciousness, the process of experiencing. This might already be enough to reject Peacocke's thesis that the world-experience of D0 may contain an experience of *here* and *now*, or "de hinc content"<sup>454</sup>, but no self-consciousness

---

<sup>446</sup> Peacocke 2014, 34f.

<sup>447</sup> Peacocke 2014, 32.

<sup>448</sup> Peacocke 2014, 33.

<sup>449</sup> Peacocke 2014, 33.

<sup>450</sup> Peacocke 2014, 33f.

<sup>451</sup> As Peacocke puts it: "If conscious subjects at Degree 0 are possible, then there can be consciousness of objects without self-consciousness, without even the most minimal form of nonconceptual *de se* content." (Peacocke 2014, 35).

<sup>452</sup> This is not only so because Peacocke is a proponent of a Sartrean-like non-egological theory of self-awareness, but also because he advocates a non-experiential approach to the metaphysics of the subject. Cf. Peacocke 2019, Ch. 4.

<sup>453</sup> On Peacocke's view, such a creature would still qualify as a subject, given that he even accepts the idea of "subjects capable only of nonrepresentational mental states" (Peacocke 2019, 38).

<sup>454</sup> Schellenberg 2016.

whatsoever. However, I want to complement my rejection with a further objection with regard to the experience of *here* and the role of self-experience, as broadly construed, in it.

Consider Peacocke's depiction of visual perception at D0:

In ordinary visual experience, for instance, there is a location from which the world is experienced. An experience can be as of the world from that location without having a *de se* content. As we might put it: the experience can have a centre, internal to the subjective character of the experience, without the subject thereby having to engage in self-representation.<sup>455</sup>

This creature at Degree 0, however, because it never represents anything as standing in certain relations to itself. None of its perceptual states have *de se* contents of such forms as *that thing is that direction from me*. Rather, they have *here*-contents, such as *that thing is that direction from here*.<sup>456</sup>

Experience of D0, according to Peacocke, may even have a center that he admits might be intrinsic to consciousness as such, without amounting to self-consciousness. Yet, what he thinks is possible, is that D0 may experience *x* as rather *there* than *here*, involving both an experience of *there* and *here*. Both points (*there* and *here*) refer to real locations in the world. But, as I want to suggest, the way they are experienced are significantly different. *There* are the objects, but what is *here*? The *here* is the *from where the visual experience* is taking place. How does Peacocke want to account for this difference phenomenologically? The experience of *there* consists in seeing an object *there*. By contrast, at the real point that corresponds to the *here*, there is precisely not a second object that was also witnessed (which would require yet another point of view from which the two objects would be seen). But what then is at the real point of *here*—if it ought to be neither an object nor the subject as self-aware? Denying that self-experience is involved in the experience of *here* and *there* leaves the experience of *here* phenomenologically indeterminate. A possible solution could be to deny that *here* is experienced as corresponding to a real point in the world. However, this would strike me as failing what *here* and *there* ought to denote. If *here* ceases to present a real point in the world, so does *there*, and both lose their meaning. Hence, I believe Peacocke's proposal cannot account for the phenomenology of *here* and *there*, without accepting that the point of view is experienced as one's own perspective. That, however, would amount to a low-level, non-conceptual form of perspectival self-consciousness.

Yet, this is precisely what Peacocke tries to avoid, when he locates perspectival self-consciousness on a level of consciousness so high that it requires the mastery of concepts:

---

<sup>455</sup> Peacocke 2014, 31.

<sup>456</sup> Peacocke 2014, 30.

[T]he subject must be capable of coming to know propositions of the form I'm  $\varphi$  where  $\varphi$  is not anchored in the subject, where the rationality of making these specifically first person judgements in the cases in which the subject comes to know them is not fully explained by what it is from something to be the subject's body, and *by my body/I* equivalences.<sup>457</sup>

But that would amount to the view that animals, such as dogs, who arguably have non-conceptual *de se* content but lack conceptual capacities (Peacocke would likely locate them at Degree 1)<sup>458</sup>, do not possess perspectival self-consciousness. I find that proposal unconvincing and not fitting well with the experience most pet-owners make with their dogs.

Contrary to Peacocke, then, I suggest he would have to acknowledge that, even on a pre-conceptual level, perspectival self-consciousness is in play, and, even in D0, provided it ought to have an experience of *here*. Of course, that would mean to give up the non-egological theory Peacocke aims to defend. But an egological interpretation, which would ascribe a sense of self even to D0, seems to fit better with the phenomenology of *here*, of which Peacocke has not convincingly shown that it can be disentangled from perspectivity and perspectival self-consciousness. Consequently, I suggest we better stick with the egological interpretation, as the separation between *de hinc* and *de se* content doesn't seem to go through as clean as Peacocke wants to make us believe.

#### 4.3.7 Challenge III: Dainton's notion of co-consciousness and the Isolation Thesis

Let me finish this section with a brief discussion of another non-egological theory that further helps clarify and, as I argue, vindicate the claim of the minimal self-view—and prepare the shift of focus from the phenomenology of self-experience to the notion of selfhood.

On Dainton's view, "it is not obvious" that "we need *mineness* to explain whether an experience is experienced as mine"<sup>459</sup>. As he proposes, an alternative consists in the following view:

Any sense I have that a typical experience is *experienced by a subject* when it occurs is due to the fact that this experience is co-conscious with certain other experiences, namely those comprising the inner component of the phenomenal background. The inner background largely constitutes *what it feels like to be me* ... . If so, then when the inner background is present, so too am I, phenomenologically speaking. Consequently, any experience which is co-

---

<sup>457</sup> Peacocke 2014, 200.

<sup>458</sup> Peacocke 2014, 38.

<sup>459</sup> Dainton 2008, 242.

conscious with the inner background will seem as though it is occurring to a subject (=me).<sup>460</sup>

But is this truly a different view than the minimal self-view? Dainton thinks so, because he believes that Zahavi posits a “primitive ‘ownership’ quality”<sup>461</sup> on top of the co-consciousness that he takes to already present the “phenomenology of *mineness*”<sup>462</sup>: “If an experience is *co-conscious* with my other experiences does it not clearly and unambiguously belong to me?”<sup>463</sup>

On first sight, it might seem Dainton rejects the minimal self simply because of a wrong understanding. For, Zahavi explicitly denies that mineness is a phenomenal quale.<sup>464</sup> In fact, much of his work consists in the attempt to show that a specific quale could not possibly bring about a primary form of self-awareness.<sup>465</sup> Accordingly, Gallagher, a no less prominent proponent of the minimal self-view, emphasizes:

If we accept Dainton’s argument, then there does exist a sense of mineness or SO, but it is not a special or additional quality, or a primitive pre-reflective self-awareness added to the phenomenal background. ... It’s difficult to see how it would count against the phenomenological conception since the phenomenologists, including Zahavi, describe the sense of mineness as an *intrinsic* aspect of experience ... .<sup>466</sup>

---

<sup>460</sup> Dainton 2008, 243.

<sup>461</sup> Dainton 2008, 242.

<sup>462</sup> Dainton 2008, 243.

<sup>463</sup> Dainton 2008, 242.

<sup>464</sup> Zahavi 2014a, 22. To be fair, this point easily lends itself to producing confusion, the reason of which lies in the diverging understanding of phenomenal consciousness and the notion of qualia. Particularly in the analytic philosophy of the mind, qualia have often been understood as referring to the sensory qualities that we find in the phenomenology of perception, for instance, the redness of tomato or the wetness of water. Moreover, following Hume, in analytic philosophy of mind, phenomenal consciousness has usually been reduced to qualia, i.e., sensory phenomenology. That means, if something is conscious, it is so because it is given in a sensory manner. It is only in the past two decades or so that this dogma has been questioned and discussed (Strawson 2010, Kriegel 2015, Pitt 2004, Bayne & Montague 2011, Chudnoff 2015, Montague 2016)—at least in analytic philosophy of mind. For, by contrast, in the phenomenological tradition Zahavi draws from, it has long been acknowledged that experience or phenomenal consciousness contains more than qualia, i.e., that something can be experientially given without being given in a way as given through our senses. The latter, liberal or “more inclusive” (Prinz 2011, 174) understanding of phenomenal consciousness, however, needs to be considered, if Zahavi’s claims are to be understood. For, when saying mineness is not a quale, he is not saying mineness is not experienced. Nor for that matter does saying that mineness is experienced, on his terms, mean that mineness ought to be a quale. For those, however, who equate qualia with conscious givenness, will likely see in Zahavi’s position an unresolvable paradox: that mineness ought to be an experiential feature (qua first-person mode of experiencing) and that mineness is not to be a quale (qua givenness of a sensorial content). It is a paradox because it will appear to them as if he was saying mineness is an experiential feature that is not experienced. However, as I hope has become clear, this is *not* what he is saying.

<sup>465</sup> Zahavi 1999, Ch. 2 & 5, Zahavi 2005, Ch. 1, Zahavi 2007a.

<sup>466</sup> Gallagher 2017, 4.

However, there remains a fundamental disagreement. It concerns the understanding of *phenomenal background*. For Dainton, most if not all our (human) experiences do possess a phenomenal background that consists of all the things that we are co-conscious of at a given point in time and so come with a sense of mineness. However, the phenomenal background, on his view, is *not* an essential structural feature of experience, given that it can, as he believes, change:

If the inner background can change and shrink, why can't it vanish altogether? The possibility of isolated and *phenomenologically ownerless* (or impersonal-seeming) experiences is thus difficult to rule out.<sup>467</sup>

The point, thus, is that for him the phenomenal background is contingent—and so is the sense of self. For such a sense is “not the product of a single simple form of experience, but rather the joint product of several different sorts of (quite ordinary) experiences”<sup>468</sup>. Without plurality of experience, no co-consciousness, no background, no self. And, according to Dainton, the Isolation Thesis, the view that a single and isolated experience of *x*, whereby *x* is absolutely singular, is plausible.

However, the Husserlian understanding of the experiential background as a temporal horizon, which Zahavi endorses, differs significantly from such an interpretation. The temporal horizon as experiential background is taken to present a minimal condition for something to be given—even if we artificially consider a single, isolated experience—in the first place. That is, even if a concrete stream of consciousness consisted in nothing but a Jamesian pulse-like moment of consciousness, a stroke of the gong and nothing more, it would involve the retentional consciousness of there not having been a sound before and the protentional consciousness of the gong to go on (which it doesn't, given that the single experience *ex hypothesi* immediately perishes and with it the stream). Hence, even an isolated experience would have a background.

Most importantly, Dainton does not seem to disagree when acknowledging “the existence and significance of fringe feelings”. For, “these fringe-components”, he explains,

together provide us with an intuitive understanding of our undeployed potential, they make us aware that there is more to our minds (to ourselves) than is occurring in current consciousness—this ‘sense of there being more’ is quite definite, and continually present in ordinary experience, albeit in a non-sensory form.<sup>469</sup>

---

<sup>467</sup> Dainton 2008, 243.

<sup>468</sup> Dainton 2008, 243.

<sup>469</sup> Dainton 2008, 44.



Do isolated experiences lack such fringe-feelings? I don't see how Dainton could possibly make such a claim if he were to. For, if he were to present an account of experience alternative to the Husserlian description, he would have to account for the now-ness character of something being phenomenally present without relying on the streaming character of experience. But not only do I find it hard to find such an alternative proposal (also in Dainton's texts), as far as I can see, although indeed explicitly discontent with Husserl's analyses of time-consciousness mostly because of their obscurity,<sup>470</sup> it still seems he agrees on the central point: that "most contents possess an intrinsic temporal organization"<sup>471</sup> to the effect that "the flow or passage in experience is included in the phenomenal content of experience"<sup>472</sup>. Could there be an exception to this? Dainton ultimately seems skeptical himself, when he says he is "inclined to think" that "durationless experience is impossible"<sup>473</sup>. But if that is so, then he won't find a way to construe a non-streaming experience either.<sup>474</sup> And, accordingly, neither one that would lack a background or temporal horizon. But then, on his own terms, a sense of self would feature in it too.

#### 4.4 On the 3<sup>rd</sup> conceptual role: The notion of selfhood

In the preceding sections, I have reconstructed two main lines of argument of Zahavi in favor of the view that pre-reflective self-awareness is best interpreted as egological. These consisted of the claim that a non-egological interpretation is incapable of explaining the possibility of reflective and intersubjective experience as we, as a matter of fact, frequently undergo it. By the same token, this resulted in the claim that the temporal self-giveness of the stream of consciousness amounts to a sense of self. Assuming from now on that this argumentation does indeed go through, two important questions remain.

(1) Provided pre-reflective self-awareness indeed consists in the temporal self-manifestation of the flow and we do have good reasons to call it egological harboring a minimal sense of self: In which phenomenological sense exactly is the temporal self-manifestation of the stream of consciousness a sense of self? How can we describe it and determinate it qua self?

---

<sup>470</sup> Dainton 2000, 161.

<sup>471</sup> Dainton 2000, 175.

<sup>472</sup> Dainton 2000, 176.

<sup>473</sup> Dainton 2008, 245.

<sup>474</sup> Cf. Zahavi 2007b.

(2) In which sense is the minimal self a minimal requirement for *any* selfhood whatsoever? How do we know the self is *essentially* experiential, as Zahavi also suggests?<sup>475</sup>

Both questions concern the 2<sup>nd</sup> as well as the 3<sup>rd</sup> conceptual role. The first question thematizes the relevance of the 3<sup>rd</sup> for the 2<sup>nd</sup> conceptual role. For, to determine the temporal self-manifestation of the flow as *sense of self*, we need to take a stance regarding what we take selfhood to *be*. But crucially, Zahavi, as I will explain, takes the being of selfhood to consist in the sense of self, when he follows Husserl's understanding of selfhood: "To be a subject is to exist for-itself, that is, to be self-aware."<sup>476</sup> This suggests that there is a relevance of the 2<sup>nd</sup> for the 3<sup>rd</sup> conceptual role, also expressed in the following statement: "[E]xperiential selfhood should be defined in terms of self-experience rather than vice versa."<sup>477</sup>

But doesn't such a position present a vicious circle? First, we want to determine the temporal self-givenness of the flow as a sense of self. To do so, we ponder on what selfhood is. Then, when trying to define what selfhood is, we come to identify it with self-experience, of which the most basic form is seen in the temporal self-givenness of the flow. Guillot thinks such a transition is unwarranted.<sup>478</sup>

However, I think there *is* a way to interpret Zahavi's position in a manner that is not viciously circular. My aim in this section is to offer such an interpretation. Although Zahavi hasn't explicitly engaged in the clarification of these two questions, I shall argue, that we can find most of what we need in his works and arguments. Bringing this afore will elucidate how in fact the question of the phenomenology of self-experience (sense of self) is intertwined with that of the metaphysics of selfhood, the notion of what selfhood *is*. Accordingly, it will also allow me to lay out the ontological interpretation of self-experience associated with the minimal self-view, which I will denote as *self-experientialism*.

---

<sup>475</sup> Zahavi 2005, 106, 125, 128, Zahavi 2014a, 73, 88.

<sup>476</sup> Zahavi 1999, 62.

<sup>477</sup> Zahavi 2014a, 74. See also Zahavi 2005, 106. It should be noted that *experiential* in "experiential selfhood" in the quote is not an adjective that ought to restrict the predicate of the proposition. Zahavi is not saying 'experiential selfhood is to be defined in terms of self-experience', while 'non-experiential selfhood does not need to be defined in terms of self-experience'. For Zahavi, the case is pretty clear. There is no non-experiential selfhood.

<sup>478</sup> As Guillot emphasizes: "As to the metaphysical thesis that the 'me-ish' quality of experience *is* the self, it succumbs to the same objection as the phenomenal thesis in so far as it presupposes it." (2017, 50) What Guillot is saying, thus, is that to posit the phenomenal thesis (that pre-reflective self-awareness is egological and presents a sense of self), Zahavi must first justify the metaphysical thesis and show what selfhood is. But the metaphysical thesis presupposes that the phenomenal thesis is true.

#### 4.4.1 From the phenomenology of reflection to a sense of self

The first thing to note to understand the argument behind Zahavi's experiential notion of selfhood is that once we start analyzing reflective consciousness in the manner as mentioned above, we first have to give up any theoretical assumptions about what selfhood is. In Zahavi's work, that is more or less a silent step he takes by building on Husserlian phenomenology and the methodological paradigm of the *phenomenological reduction* and *epoché*.<sup>479</sup> Most importantly, this step should be uncontroversial, as the goal is to determine what selfhood is and what could possibly count as a sense of self. Moreover, I think it is a step Guillot and those who follow her skepticism about Zahavi's position would find unproblematic. After all, the goal is to develop a notion of self and *not* presuppose it.<sup>480</sup>

The next step is to direct one's attention to reflective experience and to examine it: Do we find anything in reflective experience that could give rise to any idea about 'self' (without having any particular notion of selfhood in mind)?<sup>481</sup> And the answer is: Yes, a certain kind of *self*-relatedness. As I have argued earlier with Zahavi, reflective experience is of the following structure: a reflective act *R* is directed at a reflected act *r*, whereby there is an awareness that *R* and *r* belong to the same consciousness (or experiential process). Following Zahavi, this was denoted as egological consciousness. Again, this is, I take it, a relatively uncontroversial step.<sup>482</sup> This form of reflective consciousness presents a kind of self-consciousness, whereby for now nothing more is meant than that it is a form of consciousness that is conscious of itself (no reference yet to *the* self, selfhood or subject). But reflective consciousness not only happens to be a consciousness of itself, it is also *aware* that it is conscious of itself.<sup>483</sup> Hence, in reflective experience we make an experience of reflexivity, the awareness of *r* belonging to *R*, i.e., to *me*, the ego performing the reflective act. Reflecting on how it was yesterday to eat watermelon, I—whatever I ultimately am—take myself to be the consciousness who had watermelon and who is

---

<sup>479</sup> Cf. Husserl 1976 (Hua 3/1), § 32, §§ 56ff. On the concept of epoché and phenomenological reduction, see, for instance, Moran 2000, 124–164.

<sup>480</sup> Most importantly, that doesn't mean the skeptics would necessarily have to perform or accept the phenomenological reduction in its full scope. It suffices to drop assumptions about the self and reflect on the phenomenology of experience.

<sup>481</sup> Note the parallels but also significant differences to Hume's empiristic reflection. Hume had a notion of self in mind corresponding to classical soul-theory. But if we start the reflection without a given notion of self, we might simply try to describe the phenomenology of experience and then examine what, if anything, motivates talk of self. It is such a method, I suggest, we are in need at this point.

<sup>482</sup> Obviously, this is an assumption one could in general argue against. However, as far as I know, most would agree that (a) reflective experience has the structure indicated, and (b) that this indeed is a form of self-consciousness and can be called egological. If there were views that propose something different, that would be an object for future debate. Here, I take it for granted.

<sup>483</sup> If reflective consciousness happened to be only conscious of itself without by the same token being aware of being conscious of itself, we would deal with a kind of *de re* self-consciousness. However, reflective experience is just not like that. It is *de se*.

reflecting on this now. This identification is *part* of the reflective experience and thus experienced. *This reflexivity presents a sense of self.*

Reflecting on our experience, therefore, we can develop a notion of self (thereby, for now, only referring to the phenomenon of reflexivity, the awareness of *r* belonging to *me*, i.e., the consciousness of the reflecting act *R*). I still consider this relatively uncontroversial. More controversial it becomes only when we ask whether such a sense of self is also present prior to the reflective act, which, most will agree, indeed has a self-reflexive structure. As I have argued extensively in Chapter 2 and the last section, we have good reasons to believe that *some* kind of minimal sense of self is also involved in *pre*-reflective experience, the temporal self-givenness of the flow. But now the question is what this kind of minimal self ought to *look* like on the pre-reflective level. Of course, I have already given Zahavi's answer: It consists in the first-personal givenness of experience or for-me-ness. Yet, it is one thing to argue that a minimal sense of self must be involved on the pre-reflective level of experience for reflective experience to be possible, but it is another thing to claim this minimal self lies *in* the first-personal givenness or for-me-ness. That the first-personal givenness is in itself a minimal sense of self. However, a proponent of a Guillot-like skeptic might object: 'We might follow Zahavi's position in so far as it shows *that* we have to assume a pre-reflective minimal sense of self but we simply cannot determine it in its phenomenological *how*. After all, it's *pre*-reflective and whenever we try to turn to it in order to describe it, we are in a reflective mode.' In this spirit, consider, for instance, Schear's doubts:

[T]o reflect on the structure and character of our *own* experience is an intensely self-conscious enterprise. ... How is pre-reflective self-consciousness to be described, if it must be recognized? ... Appealing to a moment of reflective consciousness is doubtfully a promising route for those of us who need convincing.<sup>484</sup>

Is Schear right? Is it impossible to draw any phenomenological conclusions from reflective to *pre*-reflective consciousness to the effect that we could at best *infer* that on the pre-reflective level we also must have *some* kind of sense of self? And, if that is so, does this mean we are not justified in calling the first-personal givenness a sense of self? Wouldn't it be better for Zahavi to make the more modest claim that, pre-reflectively, there is a minimal sense of self but we cannot further make any statement about it apart from the fact that it must exist?

All these questions, I suggest, can be answered in the negative. Zahavi *can* describe the pre-reflective minimal sense of self presupposed for reflection in terms of first-personal givenness. To see this, first, we need to give up an important background assumption that

---

<sup>484</sup> Schear 2009, 101f.

is operative in the kind of skeptical attitude put forward by Schear: Schear treats pre-reflective experience as if it was some kind of black box to be contrasted with the more illuminated experience of reflection.<sup>485</sup> Crucially, this is a repetition of the higher-order or reflection theory of consciousness: On this view, a reflection is required to bring something into the light of consciousness. But this was exactly the doctrine Zahavi has led ad absurdum by highlighting that if it were true, reflective consciousness wouldn't be possible. Hence, once reflective consciousness is accepted as a phenomenon of self-consciousness we cannot go back to defending a reflection theory of consciousness and pretend pre-reflective experience is something we have no access to. Hence, the argument of Zahavi in favor of a pre-reflective sense of self is *not* a mere inference about something that otherwise remains in the dark.

Most importantly, the fact that we are pre-reflectively self-aware is not only something that we can theoretically argue for but is backed by the phenomenon of reflection itself. It is part of the phenomenon of reflection. For, in reflection we thematize a certain experience  $r$ , the eating of a watermelon. But in focusing on this past experience, we are aware that we are reflecting. Reflecting on a past experience of eating watermelon, we don't undergo the experience of eating a watermelon once again. We reflect on it and recollect how it was like. We are aware of the reflective mode in which we examine the eating watermelon. Crucially, we don't thematize the reflecting act  $R$  but in being aware of reflecting on  $r$ , which we thematize, we are nonetheless aware of  $R$ . Otherwise our reflection on  $r$  wouldn't have the character of reflection, which it has. The phenomenon of reflection is thus itself a testimony of pre-reflective self-awareness: in reflecting on  $r$ , we are pre-reflectively self-aware of  $R$ . We could also engage in a reflection on  $R$ , i.e., our reflecting on  $r$  (and thus ponder on how it is to reflect on  $r$ ). But we don't have to in order to be aware of  $R$ .

Hence, for the reflexivity that holds between  $R$  and  $r$  to be possible, not only is it theoretically required that I have been pre-reflectively self-aware of  $r$  while  $r$ , it is also necessary that I am also pre-reflectively self-aware of  $R$  while  $R$ . And most importantly, the latter is manifest *in* reflective experience. Otherwise, I couldn't experience that  $R$  and  $r$  belong to the same consciousness. For,  $R$ , given that the reflection is directed at  $r$ , is not the object of reflection. Hence, in  $R$  we have a non-objectifying and pre-reflective awareness of  $R$ .

Most importantly, that means, we do not only have one token of reflexivity in reflective experience: the one that holds between  $R$  and  $r$ , but also the reflexivity of  $R$  with  $R$  (and that of  $r$  with  $r$ ). In reflection, we experience:

---

<sup>485</sup> His emphasis on the “refrigerator light fallacy” (Schear 2009, 101) is the most direct expression of this presupposition.

- (1) *r* in a way that suggests that *r* had been self-aware prior to our reflecting (*R*) on it;
- (2) *R* in a way without reflecting on it;
- (3) *R* and *r* belonging to the same consciousness.

Accordingly, in reflective experience, we *do* have a sense of a reflexivity that isn't reducible to the reflective reflexivity between *R* and *r*. In fact, the reflexivity of *R* being pre-reflectively self-aware (2) is just a presupposition as is the 'inferred' pre-reflective self-awareness of *r* (1) for reflective reflexivity (3). Now, we can ask how is it, phenomenologically, that we are aware of *R*? How is it that we are aware of reflecting (*R*) on eating watermelon (*r*)? I think most will agree now that nothing else is required apart from the fact that we are aware of the act of *R*. It suffices that we are aware of reflecting to experience the reflecting act as *our* act. The first-personal givenness of undergoing the reflective act *R* suffices to be self-aware of it as performed by *me*. Why should it be any different in the case of *r*? Just as we are pre-reflectively self-aware of *R* while *R*, we can phenomenologically grasp how it must have been to be *self*-aware of *r* while *r*: namely simply by being aware of *r*, i.e., by first-personal givenness. Awareness of awareness is *self*-awareness. It presents a form of reflexive experience that involves the experience of reflexivity: a sense of self.

#### 4.4.2 From the phenomenology of a sense of self to a notion of selfhood and back?

In the previous section, I have argued, in reflection, we can grasp what kind the reflexivity or sense of self must be like in *pre*-reflective experience. Let us assume I indeed succeeded in showing this. Still, one might ask, even if we were to admit that some kind of experience of *immanent reflexivity*<sup>486</sup>, some *ipseity*<sup>487</sup>, *de se*, *sense of self*, or what have you, is indeed involved in pre-reflective experience: In which sense does this present Guillotonian me-ness\*? In which sense does this amount to a sense of self in terms of the self-awareness of the subject of experience? All what you have shown, at best, is that consciousness is admittedly intrinsically conscious of itself and that this comes with a sense of self, some experience of reflexivity akin to that that we know from reflective consciousness. How does this amount to an awareness of the subject of experience itself and not only of its experiences that admittedly might have some reflexive character?

Again, we already have Zahavi's answer (and metaphysical thesis): Because the subject of experience *is* itself experiential and nothing over and above the stream of consciousness.<sup>488</sup> Hence, to make the claim that for-me-ness presents me-ness\*, Zahavi

---

<sup>486</sup> Frankfurt 1987, 30.

<sup>487</sup> Translated from "ipséité", Sartre 1956, 102.

<sup>488</sup> Zahavi 2014a, 18, 36, 72.

must make a metaphysical commitment which consists in an ontological interpretation of self-experience:

	Sense of self	=	Sense of self
	(=experience of reflexivity)		(=experience of selfhood)
Zahavi		For-me-ness	
Guillot	For-me-ness	?	Me-ness*

Zahavi’s notion of for-me-ness involves the identification of sense of self (=experience of reflexivity) with the sense of self (=experience of selfhood). As Guillot rightly emphasizes, such an equation is not tautological nor trivial.<sup>489</sup> And she is also right that this step is rather silent but operative in Zahavi’s approach. But from it being silent, one cannot infer that it is unwarranted nor that Zahavi hasn’t implicitly provided arguments in favor of that view.<sup>490</sup> Still what is needed, of course, is an explication.

Let me first elucidate the silent step. Recall that I said that when engaging with reflective egological consciousness, no notion of selfhood was involved in the examination other than the reflexivity referring to the identity between *R* and *r*, the ego being directed at one of its own past experiences. All other notions of self, for instance, a substantial understanding of the subject as metaphysically underlying experience, were bracketed. In fact, to describe and understand the phenomena we find in reflective egological consciousness as well as their phenomenological conditions, no reference to an alleged consciousness-transcendent subject-entity is needed (nor helpful).

Crucially, however, having dismissed or bracketed all notions of selfhood, how does Zahavi come back from an examination of selfhood as it is experienced in consciousness to an ontological claim about what selfhood *is*? Wherein consists his silent step back to the metaphysical level that allows him to claim that the awareness of experience, the reflexivity, is *self*-experience? How does he proceed from a sense of self (=experience of reflexivity) to a sense of self (=experience of selfhood)? I suggest we find a hint in the following quotes:

[T]his entails that self-consciousness is a constitutive feature of phenomenal consciousness. But what does all of this have to do with subjectivity and selfhood? Quite a lot, in fact. As Galen Strawson has argued, if we wish to understand what it means to be a self, we should look at self-experience, since self-experience is what gives rise to the question in the first place by giving us a vivid sense that there is something like a self.<sup>491</sup>

---

<sup>489</sup> Guillot 2017, 36.

<sup>490</sup> Guillot 2017, 37.

<sup>491</sup> Zahavi 2014a, 17.

[I]f we wish to answer the metaphysical question concerning whether or not the self is real, we will first need to know what a self is supposed to be. In order to establish this, our best chance will be to look at self-experience.<sup>492</sup>

But how do we get from pre-reflective self-consciousness to a sense of self? ... As I read Sartre, his proposal is that rather than starting with a preconceived notion of self, we should let our understanding of what it means to be a self arise out of our analysis of self-consciousness. Put concisely, the proposal is to identify the self with the subject of experience...<sup>493</sup>

The rationale, thus, is to examine self-experience in order to *develop* a notion of selfhood, because it is self-experience that motivates the issue of self in the first place.<sup>494</sup> Hence, the phenomenological task, when trying to conceptualize selfhood, is to determine where and how in experience something comes up that could motivate the issue of self. What kind of experience is it that makes us speak about selfhood to begin with? Even though Zahavi doesn't pose this question in this precise way, I think it is possible to reconstruct an explicit answer from his proposal. I suggest it could look like the following:

1. Egological consciousness in reflection comes with an experience of reflexivity: *I* reflect on a past experience of *mine as mine*. Whenever we reflect, we have an experience of ourselves as an act-transcendent pole of experience: *I* am the one who is undergoing all these different experiences. In reflection, our experience non-controversially involves a sense of self-familiarity, a sense that in reflecting on one's own experiences one is reflecting on oneself.

2. The next step then is to ask where in experience this kind of reflexivity arises *originally*. Which is the primordial moment in which 'self' comes up first? Of course, as I have extensively discussed, Zahavi argues that it is not only in reflective experience that reflexivity is involved but also in any experience whatsoever: pre-reflective self-awareness or for-me-ness.

Having engaged in something like a phenomenological genealogy of the issue of selfhood as it arises *in* experience, the preliminary answer is: Pre-reflective self-awareness is the most basic form of self-experience (= experience of reflexivity), which makes it a promising *candidate* for a sense of self (= experience of selfhood). As Locke says,

When we see, hear, smell, taste, feel, meditate, or will any thing, we know that we do so. Thus it is always as to our present Sensations and Perceptions: And by this every one is to himself, that which he calls *self* ... [C]onsciousness

---

<sup>492</sup> Zahavi 2005, 128.

<sup>493</sup> Zahavi 2020b, 640.

<sup>494</sup> Interestingly, in a sense, I believe, a somewhat similar point is made by Guillot herself, when she claims that the concept of self is phenomenally grounded and based on an experience of reflexivity, Guillot 2016.



always accompanies thinking, and 'tis that, that makes every one to be, what he calls *self* ....<sup>495</sup>

The suggestion, thus, is that it is the reflexive character of experience that gives not only rise to a sense of self (= experience of reflexivity) but in doing so constitutes selfhood and therefore presents a sense of self (= experience of selfhood).<sup>496</sup> The appeal to reflexivity and its interpretation is partly motivated by the fact that we—bearing in mind Hume's introspective verdict—indeed do *not* encounter any self-object in experience nor for that matters in the world but yet come to speak about selfhood: We still engage in sentences like 'I am happy to see you' and we still are deeply affected and find stark resistance against the idea that all our self-experience ought to be hallucinatory or illusory. Zahavi's phenomenological position provides a possibility to understand why this is so and offers an alternative way of understanding selfhood that is based on the phenomenology of our experience and thus experientially motivated. It fits with our intuitions of being someone without gesturing towards entities that are elusive and beyond our experiential access.

However, he is not the only one who has turned to experience to develop an understanding of selfhood. In the next chapter, I compare his account with some other classical and recent experiential approaches of selfhood to better assess which reasons may speak in favor of his view and which challenges it has to overcome. But first, let me provide a short overview of the claims constituting his ontological interpretation of self-experience which amounts to what I suggest to call *phenomenological self-experientialism*.

#### 4.4.3 Minimal self and phenomenological self-experientialism

Phenomenological self-experientialism consists in two general claims. The first one is methodological:

- (1) Self-experience is the *modus cognoscendi* of selfhood: It is only through self-experience that we come to speak about selfhood in the first place. Hence, we should examine our self-experience to investigate what selfhood *is*.

Investigating self-experience in reflective experience, Zahavi suggests the most basic form of self-experience consists in the temporal self-manifestation of the stream of consciousness. Rather than referring to anything beyond and external to the stream of

---

<sup>495</sup> Locke 2008/1689, 208.

<sup>496</sup> Unlike it is sometimes thought of, the step is not to add a further aspect to the experience of reflexivity but to identify the experience of reflexivity with the experience of selfhood. In identifying the two, no further phenomenal feature is posited. Therefore, arguing that the experience of reflexivity does not yet amount to a sense of self is not more parsimonious than identifying the two. It is simply a different interpretation that requires the same amount of argumentation. Cf. O'Conaill 2019.

consciousness, the most basic experienced reflexivity concerns the self-givenness of experience. This motivates the second, metaphysical claim:

- (2) Self-experience is the *modus essendi* of selfhood: It is in self-experience that selfhood is real. The reflexivity or self-awareness of experience just is what constitutes the being of selfhood, it is selfhood.

And this is, finally, the reason why all three conceptual roles need to be considered to understand the full scope of the meaning of the minimal self. They are linked through the experience of reflexivity that I have described throughout the chapter. The experience of reflexivity forms part and parcel of *phenomenal consciousness* (1<sup>st</sup> conceptual role). It provides the most basic form of *self-consciousness* (2<sup>nd</sup> conceptual role), which amounts to a sense of self in terms of me-ness\*. This is so, because we identify with the most basic experience of reflexivity that “what makes every one to be, what he calls *self*”<sup>497</sup>, i.e., the most basic form of selfhood, a minimal *self* (3<sup>rd</sup> conceptual role). This identification is motivated by the fact that the kind of reflexivity found in reflective experience—and which many are prone to accept as egological—traces back to a more fundamental kind of reflexivity: pre-reflective self-awareness.

## 4.5 Conclusion: Self is sense of self

In this chapter, I reconstructed the ontological interpretation of self-experience that underlies Zahavi’s phenomenology-based minimal self-view. To this aim, I described the three conceptual roles that are ascribed to the minimal self and their structural entwinement. Doing so revealed the metaphysical commitment that is necessary to make the claim that for-me-ness is inseparable from Guillotonian me-ness\*, i.e., the view that any experience involves a sense of self proper, even though the ego is not yet explicit or thematized. The metaphysical commitment consists in the identification of the experience of reflexivity with selfhood. This allows Zahavi to interpret the experience of reflexivity as self-experience or a sense of self.

Is this identification of the experience of reflexivity with the minimal self warranted? So far, I have only tried to reconstruct the ontological interpretation of self-experience operative in Zahavi’s minimal self-view and to show the motivation for such an interpretation. Doing so brought afore that the identification isn’t due to an overseen conflation of different notions of the subjective character of experience but rather

---

<sup>497</sup> Locke 2008/1689, 208.

phenomenologically motivated: a phenomenological thesis about the metaphysics of selfhood that departs from a phenomenological analysis of reflective (and intersubjective) experience. I labelled this position *phenomenological self-experientialism*, which is characterized by the view that self-experience is not only the *modus cognoscendi* but also the *modus essendi* of selfhood. To further assess whether phenomenological self-experientialism is an attractive alternative to the skeptical views about self-experience, in the following chapter, I want to contrast it with competing experiential approaches to selfhood.



## Chapter 5

# Gaps, Identity, and the Primacy of the Minimal Self

[A]t the core of all awareness, there is the self which *is* aware and which must itself be understood to be an entity, as much as entities in the world around us, but with the peculiarity that it cannot be known or described like external objects in terms of special attributes, but only in the awareness of himself that each one has in his own case  
in any experience whatsoever.<sup>498</sup>

Is *self* the same with *substance*? If it be, how can that question have place, concerning the subsistence of self, under a change of substance? If they be distinct, what is the difference betwixt them? For my part, I have a notion of neither, when conceiv'd distinct from particular perceptions.<sup>499</sup>

In the previous chapter, I presented self-experientialism to denote the ontological interpretation of self-experience intrinsic to Zahavi's minimal self-view. I suggested that such an experiential approach to selfhood may provide an alternative to the three skeptical views about self-experience. Crucially, however, self-experientialism is not the only experiential approach to selfhood. In this chapter, I contrast self-experientialism with these competing theories and examine how self-experientialism could answer to relevant challenges that have motivated other philosophers to develop experiential approaches to selfhood different to Zahavi's phenomenological self-experientialism. This shall allow me to further clarify the commitments that attach to self-experientialism and their tenability.

To set the stage for the discussion of self-experientialism in light of recent competing experiential accounts of selfhood, I begin with a brief description of how self-experientialism relates to two classical accounts: Descartes' account of the ego as *res cogitans* and Locke's account of *personal identity*. Considering two relevant differences between self-experientialism and these classical philosophers, I will argue, will facilitate bringing self-experientialism into discussion with more recent experiential approaches to selfhood:

---

<sup>498</sup> Lewis 1982, 175.

<sup>499</sup> Hume 2007/1739, 399f. (App. 18).

- (1) the rejection of the Cartesian substantial understanding of selfhood
- (2) the rejection of Locke's fusion of the definition of selfhood with issues of personal identity.

Against the background of these two features of self-experientialism, I will discuss four recent experiential approaches: Martine Nida-Rümelin's rendition of pre-reflective self-awareness that implies a substantial understanding of selfhood (5.3); Barry Dainton's "C-Theory"<sup>500</sup> that defines selfhood in terms of a capacity for consciousness (5.4); Galen Strawson's "Pearl view"<sup>501</sup> that identifies any given uninterrupted period of consciousness each with an ontologically distinct self (5.5); and Tim Bayne's "virtual phenomenalism"<sup>502</sup> that conceives of the self as a virtual object that is constituted within experience based on a narratively structured reflection on the stream of consciousness (5.6). As I will argue, all four alternative experiential approaches suffer from weaknesses questioning whether they are to be preferred over self-experientialism. Most of these alternative accounts are based on the assumption that a notion of selfhood must be capable to accommodate issues of personal identity. And, each of these accounts presents their own sui generis solution to the challenges of personal identity, thereby rendering the definition of selfhood differently in each case. Finally, I discuss the question whether self-experientialism needs to provide a convincing account of personal identity to make its case. Having distinguished self and person, I conclude that if anything it is the notion of selfhood that constrains the discussion of personal identity—and not vice versa (5.7).

## 5.1 Self-experientialism is not Cartesianism—or is it?

Self-experientialism takes self-experience as *modus cognoscendi* and *modus essendi* of selfhood. This amounts to a view diametrically opposed to the three skeptical views about self-experience. However, it also places self-experientialism somewhere near Cartesianism. For, Descartes similarly held that the being of selfhood consists in the experiencing (*cogitare*) and reflexivity concerning one's own experiences. And, as the most famous proponent of rationalism, he also defended that we can yield an insight about the essential character of our being by reflection. In fact, the whole argument in his *Meditationes* in favor of substance dualism and the associated determination of our being as *res cogitans* rests on a reflection on our self-experience, the *cogito*.<sup>503</sup> Indeed, the

---

<sup>500</sup> Dainton 2008, 112, 113.

<sup>501</sup> Strawson 2017, 37.

<sup>502</sup> Bayne 2008, 289ff.

<sup>503</sup> Descartes 1997/1641.

phenomenological tradition Zahavi's minimal self-view is based on usually did not take Cartesianism *in toto* as a label for a position that was to be avoided by all means. Unlike in analytic philosophy of mind, in which any relevant similarity with Cartesianism of a position would often risk a *reductio ad absurdum*, phenomenology "found in Descartes much more a *causa belli* than the usual *bête noire*"<sup>504</sup>. The relationship of phenomenology to Cartesianism is thus complex, which is also reflected in the following quotation by Husserl:

[O]ne might almost call transcendental phenomenology a neo-Cartesianism, even though it is obliged—and precisely by its radical development of Cartesian motifs—to reject nearly all the well-known doctrinal content of the Cartesian philosophy.<sup>505</sup>

Phenomenology, thus, just as much as it is Cartesian is decisively *not* Cartesian. To carve this out in detail would require a thorough investigation, given that there is not only heterogeneity to be found across phenomenologists regarding their stance towards Cartesianism but also because the label of Cartesianism refers to quite a few different claims ascribed to Descartes.<sup>506</sup> However, such an investigation is not necessary for the current purpose, namely to determine whether self-experientialism amounts to the Cartesian substantial understanding of the self as a thinking thing (*res cogitans*) that is to be conceived of in separation of the body (*res extensa*). The following four interrelated aspects of self-experientialism render it unambiguously anti-Cartesian in the relevant sense.

### *1) Reflection vs. pre-reflective theory of self-consciousness*

As Frank points out, Descartes was a proponent of the reflection theory of consciousness and self-consciousness, according to which consciousness consists in the reflection upon some cognitive content.<sup>507</sup> This is exemplified in the following statement by Descartes: "Conscium esse est quidem cogitare et reflectere supra suam cogitationem."<sup>508</sup> Quite to the contrary, the phenomenological account of self-awareness has been described as a decisively *pre-reflective* theory of self-awareness.

---

<sup>504</sup> Martin 2008, 496.

<sup>505</sup> Husserl 1982, 1.

<sup>506</sup> See, for instance, Alweiss 2010, Becker 2001, Crowell 2002, Landgrebe 1961, Luft 2004, MacDonald 1999, Romano 2012, Soffer 1981, Taipale 2015.

<sup>507</sup> Frank 2015, 35f.

<sup>508</sup> Descartes 1982, 12.

## 2) *Non-substantial understanding of selfhood*

On Descartes' view, self-consciousness consists in the reflection upon cognitive, i.e., experiential, content. If we, according to Descartes, reflect upon the *cogitare* that is revealed in the hyperbolic doubt of the *Meditations* we recognize it, i.e., the thinking or experiencing, as the essence of our being. Crucially, Descartes proposes to understand this essence of being in substantial terms to the effect that he determines our being as *res cogitans*. By contrast, one central aim of the phenomenologists and particularly of the self-experientialism implied in the minimal self-view is to conceptualize selfhood without reifying it and conceive as a mental or experiential *thing*. This is also the reason why Zahavi puts much effort and emphasis on the fact that the most basic form of self-awareness does not consist in some consciousness of an object. The pre-reflective reflexivity of experience doesn't refer to an object but to the experiential process itself in which any object can be given in the first place.

## 3) *Bodily selfhood*

Unlike Descartes, most phenomenologists have insisted on the essentially bodily constitution of selfhood.<sup>509</sup> One important aspect of this bodily constitution is that it is a precondition that allows us “to understand how we can eventually appear to ourselves as mundane objects and interact with Others in a common world”<sup>510</sup>. It is only through our body that we can be part of the world and are visible to others. Similarly, we only encounter others insofar as they are embodied. Without the body of the other, we wouldn't have any access to the other as an experiencing being at all.

However, according to the phenomenologists, the role of the body is even more fundamental: bodily self-awareness is also part and parcel of perceptual intentionality and object-experience. If I look at the notorious cup, I only see, *in sensu strictu*, the front side of the cup, while the back side remains hidden. Crucially, the hidden back side is manifest in experience, namely as absent. Part of this experience of absence is the experience that *if* I moved in this or that manner, *then* I would see the back side. The experience of this capability is what Husserl calls “kinaesthesia (*Kinästhesie*)”<sup>511</sup> or kinesthetic experience. It presents a sense of one's own bodily possibilities, more specifically, of the covariation between one's own possible movements and the expected perceptions. For Husserl, that means that we cannot possibly have an experience of an object without being at the same time aware of one's own body. For, the awareness of the absent sides of an object (which are indispensable aspects of any object-experience) just *is* the awareness of the necessary

---

<sup>509</sup> See, for instance, Gallagher 2005a, Legrand 2007, Moran 2017, Staiti 2015, Summa 2014, Taipale 2014, Zahavi 1999, Ch. 6 and 9.

<sup>510</sup> Zahavi 1999, 91.

<sup>511</sup> Husserl 1973b (Hua 16), 154–203. See also Claesges 1964, 117–132, Drummond 1984.



movements that would be required to attain a perspective from which the currently absent sides would be visible. And that just *is* a fundamental bodily self-awareness.<sup>512</sup>

#### 4) *Solipsism and intersubjectivity*

Cartesian subject philosophy has often been accused of leading into an unsurmountable solipsism that results as a consequence of having set up the philosophical stage by methodical and hyperbolic doubt. Given that only one's own experience qua experience survives the doubt and comes with apodictic evidence, the ontological status of the world and the other obtains a dubitable character. The reality of world and the other consequently is in need of further proof, while one's own being is immediately secured. Referring to pre-reflective self-awareness, as an immediate reflexivity that preconditions any encounter with the other, as I have argued in the last chapter, might then seem to equally entail a scepter of solipsism that cannot be gotten out.

Yet, emphasizing pre-reflective self-awareness as a condition for intersubjective experience doesn't mean to reduce the other to one's own experience. Husserl—and with him many phenomenologists—pointed rather out that the way of the other's givenness just *is* the characteristic transcendence that adheres to his presence. The other could not possibly be given to me in a way akin to my own experience. It couldn't possibly have the same kind of immediacy without ceasing to be the experience of someone else, the *other*. Husserl and the other phenomenologists, accordingly, never tried to overcome the hiatus that lies between oneself and the other which would dissolve the transcendence of the other.<sup>513</sup> Rather, they tried to describe it in a phenomenologically adequate manner so as to carve out the precise structure and meaning that it has for the experience of the world. From the perspective of the phenomenologists, thus, the task would not be to find ways to provide infallible knowledge of the existence of the other. Rather, it consists in acknowledging that it is essential to the experience of the other that it is presumptive and fallible.

Now, some might not be persuaded by this and find that such a line of argument precisely reveals that the phenomenologists just don't find a way that could bridge one's own and the other's being. But such a response would, I take it, quite to the contrary, simply betray that one's own point of departure is much more Cartesian than is the phenomenological position.<sup>514</sup>

---

<sup>512</sup> Zahavi 1999, 95ff. See also Claesges 1964, Summa 2014, Welton 2002, Zippel 2014.

<sup>513</sup> Zahavi 1996, 29.

<sup>514</sup> On the issue of intersubjectivity in Husserlian phenomenology, see, for instance, Gallagher & Zahavi 2008, Ch. 9, Kjosavik, Beyer & Fricke 2020, Römpf 1992. Cf. Mehl 2021. Emphasis on the importance of the interpersonal for much of our experience of the world, which undermines a solipsistic view of selfhood, can also be found in recent views in applied phenomenology. See, for instance, Fuchs 2016, 2020, Ratcliffe 2017.

Even these few points suffice to show that the phenomenological project and notably the self-experientialism associated with the minimal self-view come with a set of quite different claims that oppose and contradict Cartesianism. Contrary to Cartesianism, self-experientialism consists in or is consistent with a pre-reflective rather than a reflective theory of self-awareness, a non-substantial rather than a substantial understanding of selfhood, an embodied rather than a disembodied approach to selfhood, and finally a theory of intersubjectivity which acknowledges the transcendence of the other as the way a social encounter is possible.

## 5.2 A modest reading of Locke: Distinguishing self from person

Descartes is not the only classical thinker whose views resemble in many regards phenomenological self-experientialism.<sup>515</sup> Locke similarly proposed an experiential understanding of selfhood that is grounded in the reflexivity of experience.<sup>516</sup> And, just like self-experientialism, Locke emphasizes that the issue of selfhood (and notably personhood) is to be treated independently from the issue of substantiality: “The Question being what makes the same *Person*, and not whether it be the same Identical Substance, which always thinks in the same *Person*, which in this case matters not at all.”<sup>517</sup> Moreover, crucially, he takes both person and personal identity to be exclusively dependent on consciousness, rendering ‘self’ and ‘person’ throughout experiential notions.

Although these features render his account very similar to self-experientialism, there are a few important and significant differences that need to be taken into account. Carving these out does not only further clarify the ontological commitments of phenomenological self-experientialism and facilitate the following comparison with other more recent experiential accounts of self, it also makes it easier to see that some of the issues some thinkers take with Locke’s position may not apply to phenomenological self-experientialism. Let me therefore briefly point out the most important difference: the distinction between *self* and *person*.

For Locke, *self* arises out of the fact of self-familiarity with one’s own consciousness: the reflexivity of experience.<sup>518</sup> With *person*, Locke refers to the “thinking

---

<sup>515</sup> In fact, Locke’s position could itself be said to be a form of self-experientialism as well as the Cartesian. However, when I talk of self-experientialism, if not stated otherwise, I am referring to phenomenological self-experientialism, different to Cartesianism and Lockeanism.

<sup>516</sup> “[I]t is by the consciousness it has of its present Thoughts and Actions, that it is self to it self now” (Locke 2008/1689, 209).

<sup>517</sup> Locke 2008/1689, 209.

<sup>518</sup> “[C]onsciousness ... alone makes what we call *self*”, Locke 2008/1689, 216. See also Locke 2008/1689, 208, 209.

intelligent Being, that has reason and reflection, and can consider it self as it self, the same thinking thing in different times and places”<sup>519</sup>. This seems to suggest a slight difference between both notions, whereby *self* could also be found in pre-reflective experience whereas the being of person consists precisely in the capability to and the actual performance of reflecting on one’s own experiences, i.e., on oneself. However, not only does Locke not speak of *pre*-reflective experience, his further deliberations on personal identity also rather suggest such a close relationship between self and person that it becomes hard to make a clear-cut conceptual distinction. This is so because for Locke it is the consciousness that is for oneself, i.e., that what is given to oneself as self, in which “alone consists *personal Identity*”<sup>520</sup>. “For it is by the consciousness it has of its present Thoughts and Actions, that it is *self* to it *self* now, and so will be the same *self* as far as the same consciousness can extend to Actions past or to come”<sup>521</sup>.

One ramification of a lack of distinction between pre-reflective and reflective experience, thus, is that self and person seem to collapse into one and the same. This seems to be an important presupposition for Locke’s theory of personal identity. For, it is this coincidence between self and person that makes him claim that whatever past experiences are given to me as mine make me identical with the self that corresponds to these past experiences, that is the same person. Self-experience, then, ends up as being the decisive criterion: it is not only what makes *selfhood* but also *personhood* and *personal identity*.

By contrast, self-experientialism *does* distinguish between pre-reflective and reflective experience. This has the repercussion that it may differentiate between self and person.<sup>522</sup> While conscious animals that lack any ability to reflect on their experiences may be said to have self-experience in terms of the pre-reflective reflexivity of experience that gives rise to *selfhood*, they fail in what would make them persons: they cannot thematize their experiences *as* their own. They cannot identify themselves with the consciousness of past experiences and so lack *personhood*.<sup>523</sup>

An important implication of this view is that self-experience as such doesn't yet amount to personhood and personal identity. The philosophical importance of this can

---

<sup>519</sup> Locke 2008/1689, 208.

<sup>520</sup> Locke 2008/1689, 208.

<sup>521</sup> Locke 2008/1689, 209.

<sup>522</sup> This is not to say that *self* and *person* is parallel to *pre-reflective* and *reflective experience*. It only means that not all selfhood need to amount to personhood. For, it is possible to hold that all pre-reflective experience includes selfhood, while personhood is something that also requires forms of higher conscious acts that involve different kinds of reflective experiences.

<sup>523</sup> Evidently, the issue of how to assess the status of animals as selves or persons is a very complex one and requires careful deliberation. I therefore do not intend to formulate a statement on these matters here. My only point is to highlight that, with the distinction between pre-reflective and reflective experience, one could ascribe selfhood to conscious animals without thereby necessarily having to claim that they are persons. No further implications or conclusions are insinuated here, given that further discussion would be needed. See, for instance, Beaudry 2019, Tooley 2012.

hardly be overestimated but easily overseen. The point is that if self-experience as such doesn't yet present personhood and personal identity, neither does the notion of selfhood or self-experience give rise to personhood and personal identity per se. According to self-experientialism, pre-reflective selfhood, the minimal self, is primordial but not tantamount to personhood and personal identity. Rather than a *guarantee*, the minimal self is a *presupposition* for personhood and as such a *constrain* for personal identity.<sup>524</sup> That means, that selfhood is a *criterion* of personal identity. Accordingly, on this view, it would be a mistake to think an experiential notion of selfhood must account for personal identity. Quite to the contrary, it would suggest that no conception of personal identity will work that involves a construal of selfhood that doesn't involve self-experience.

Self-experientialism, thus, could be understood as a *modest* interpretation of Locke's account of personal identity. It is modest because it doesn't purport to have provided all what is needed to explain personal identity. Given that Locke's theory focuses so much on defending self-experience as a *sufficient* condition for personal identity there is the risk to oversee that independently of whether he succeeds in this defense, he certainly makes also a case for experience (and self-experience) as a *necessary* condition for personal identity:

Nothing but consciousness can unite remote Existences into the same Person, the Identity of Substance will not do it. For whatever Substance there is, however framed, *without consciousness, there is no Person.*<sup>525</sup>

I suggest this is the more modest and compelling case that one can indeed make with Locke and which is compatible with self-experientialism. The most important repercussion of this—relevant for the discussion of the competing experiential theories of selfhood—is that, to undermine self-experientialism, simply pointing out that self-experience doesn't secure personal identity will not do. Showing that something isn't sufficient doesn't mean it isn't necessary. Crucially, however, as I will show in the next sections, proponents of competing theories of experiential selfhood seem to have just endorsed such a—wrong-headed—argumentative strategy to launch their own views regarding experiential selfhood that differ from Locke and self-experientialism in a relevant sense.

Before, however, I first want to specify what is meant with the claim that self-experience, while not a *sufficient*, is a *necessary* condition for personal identity. For Locke, it is clear what it means that self-experience (or consciousness, as he would put it) is a sufficient condition for personal identity. If I remember what are the past experiences of the consciousness that I am now, then these past experiences form part of *myself*, and I can be said to be identical with the consciousness that has had these past experiences: “Any

---

<sup>524</sup> Cf. Zahavi 2005, 129, Zahavi 2014a, 51ff.

<sup>525</sup> Locke 2008/1689, 216, my emphasis.

thing united to it [the present thinking Being, P.S.] by a consciousness of former Actions makes also a part of the *same self*, which is the same both then and now.”<sup>526</sup> But memory is fallible, and I might as well be delusional when I vividly though wrongly ‘remember’ myself having a conversation with Bismarck and take myself to be Kaiser Wilhelm I. Consciousness, i.e., memory, as mentioned earlier, cannot be said to be a guarantee for personal identity. It doesn't suffice.

On a first reading, to say that, instead of a sufficient condition, consciousness could be considered a *necessary* condition, could mean that even though I can be completely wrong when remembering a past experience *x*, I—the present consciousness—can only be identical with a consciousness that I can remember. Accordingly, if I undergo—perhaps in the last decade of my hopefully long life—a form of dementia based on which I lose the memory of the first three decades of my life, I would be no longer identical with the person whose experience I fail to recollect. In fact, Locke is indeed making this claim:

[I]f it be possible for the same Man to have distinct incommunicable consciousness at different times, it is past doubt the same Man would at different times make different Persons ...<sup>527</sup>

Locke would readily concede, truly emphasize, that we need to distinguish between the identity of the *man* and the identity of the *person*.<sup>528</sup> Hence, if I undergo dementia at the age of 90, to my grandchildren I will still be the same *man*. Yet, since I don't remember ‘my’ past experiences *x*, I am—*qua* person—no longer identical with the *person* who underwent *x*. I—the same man—have become a different person. Taken like this, the claim that consciousness (i.e., memory) is necessary for personal identity still seems very strong.<sup>529</sup> For, claiming that my self in dementia and my younger self are different persons seems to imply a neglect of the fact that both selves are not only related to each other by forming part of the same man, but that they are also psychologically connected: Even though the older person will not remember experiences of the younger person, the experiences of the latter surely shape significantly the experiences of the former. Past experiences are still implicitly part of present experiences, even though not in an explicit

---

<sup>526</sup> Locke 2008/1689, 218.

<sup>527</sup> Locke 2008/1689, 215.

<sup>528</sup> Locke 2008/1689, 213.

<sup>529</sup> Unless one would simply consider the notion of personal identity to refer to the *sense* of identity. On such an interpretation that I cannot further discuss here, one might take personal identity to be a purely phenomenological notion.

and thematized—properly recollected—manner. Both young and old self do not wholly form discrete persons.<sup>530</sup>

However, there is an alternative way to conceive of consciousness as a necessary condition for personal identity. Instead of saying that a person must have the *actual* capacity to remember past experiences *x* in order to constitute personal identity with the one who was undergoing *x*, it is only necessary that a person has the *potential* capacity to remember *x*. Take again the old self in dementia. Although in dementia I might not be able to recall how I met my wife, provided there is a new medical treatment yet to be discovered that can make all effects of dementia undone, I still have the potential to recall it. On a much smaller scale, less dramatically, I submit that for everybody it is quite easy to see that there are many past experiences that we may not have present for years or even decades, that we cannot recall, or that when described to us, we still don't really remember—and yet might in future re-appropriate. Awakenings of such a kind, I deem, are quite common. I might have decided to go for a further career, invest 10 years into it, and then at some point come to remember that there had been moments in which I had wished for another track. Suddenly all the experiences related to that alternative track spring to my mind, experiences I hadn't recollected in years and wouldn't have been able to recollect if asked.

On this alternative way, thus, it is a necessary requirement for personal identity that I can *in principle* remember the experiences of the person that I purport to be identical with. Denying the tenability of the stronger claims doesn't affect this last claim. In fact, this claim is present in Locke's proposal and it is one that is more fundamental than the claim that actual conscious memory is a necessary or sufficient condition for personal identity, when he says: "without consciousness, there is no Person"<sup>531</sup> and "consciousness ... alone makes what we call *self*"<sup>532</sup>. Although this is a modest reading of Locke that drops the stronger claims, it is surely not a trivial claim. For it still maintains that Locke considered selfhood and personhood to be constituted originally by consciousness (and self-experience). Apart from claiming that consciousness guarantees personal identity, Locke also and more fundamentally claimed that consciousness *constrains* personal identity. On this view, parallel to self-experientialism, it is only where we have consciousness and self-experience that we might have selfhood and a person to begin with and that the question of personal identity can be posed at all. Or put differently, Locke's question of personal identity is always one that refers to the relationship between two *experiencing* selves.

---

<sup>530</sup> This fact certainly has motivated Neo-Lockean accounts that have emphasized *psychological* instead of *phenomenal* continuity as the criterion for personal identity. Cf. Noonan 2003, Parfit 1987, Perry 1976, Shoemaker 1984.

<sup>531</sup> Locke 2008/1689, 216.

<sup>532</sup> Locke 2008/1689, 216.

To conclude, self-experientialism shares one fundamental claim with Locke's concept of personal identity: the view that selfhood is defined by self-experience. Unlike Locke, however, self-experientialism distinguishes between pre-reflective and reflective experience, attributing to the former selfhood and taking reflection to be a condition for personhood but not selfhood. As to personal identity, self-experientialism doesn't commit to the view that conscious memory and the self-experience of recollection is a guarantee for personal identity.<sup>533</sup> It does share the view with Locke, however, that personal identity can only hold between experiencing selves, rendering *self-experience and selfhood a constrain for personal identity*.

By contrast, crucially, some of the recent alternative experiential approaches to selfhood precisely differ from self-experientialism with regard to the latter claim. Although they purport to present experiential approaches to selfhood, it is precisely vis-à-vis the issue of personal identity that they move away from an experiential understanding (and definition) of selfhood in order to accommodate questions of personal identity, thereby making *personal identity a priority over the question of the nature of self*. Yet, I think it is far from clear that such a prioritization of the issue of personal identity is convincing.

Before I can discuss this issue further (which I will do in Section 5.7), in the next sections, let me first discuss the main important recent alternative experiential approaches to selfhood that differ from self-experientialism. This will allow us a better overview of how the attempt to define selfhood in terms of (self-)experience and the attempt to accommodate issues of personal identity interact with each other, and whether it is worth making personal identity a priority. Or whether it is preferable, as I have suggested, to rather take selfhood as a constrain for personal identity.

### 5.3 Nida-Rümelin: Pre-reflective self-awareness as 'nature-revealing'

Let me put the issue of personal identity, which will be relevant for the sections to follow, aside for a moment. First, I want to discuss an approach that is remarkably similar to Zahavi's minimal self-view and self-experientialism, but turns out to be fundamentally different with regard to the metaphysical interpretation of self-experience: Martine Nida-Rümelin's account of pre-reflective self-awareness as "an awareness of oneself *as an experiencing subject*"<sup>534</sup> and as "nature-revealing"<sup>535</sup>. Her account shares at least four essential claims with self-experientialism:

---

<sup>533</sup> Cf. Zahavi 2014a, 76.

<sup>534</sup> Nida-Rümelin 2017, 66.

<sup>535</sup> Nida-Rümelin 2017, 76.

- (1) the view that “each of us is permanently aware of him- or herself in a pre-reflective and pre-conceptual way in any moment of his or her conscious life”<sup>536</sup>;
- (2) that pre-reflective self-awareness “of one’s own experiencing is required for phenomenal reflection as a precondition”<sup>537</sup>;
- (3) that “our understanding of what it is to be an experiencing subject originates in that omnipresent self-awareness”<sup>538</sup>;
- (4) that “ontological claims about the nature of experiencing subjects ... can only be understood and can only be convincingly motivated on the basis of phenomenological insights concerning the way in which we are in every moment of our conscious life pre-reflectively aware of ourselves as experiencing beings”<sup>539</sup>.

Briefly stated, Nida-Rümelin, like self-experientialism, proposes that it is our most basic form of self-experience through which we *can* gain insight into the nature of selfhood. In fact, as Nida-Rümelin unambiguously exclaims, “our own nature is present to us in such self-awareness in a phenomenologically manifest way”<sup>540</sup>. Already on a pre-reflective level, we are aware of our conscious being, we know what it means for a subject to be the pole of different but simultaneous experiences: for instance, “[w]hile you hear a bird’s voice, see a tree, feel the wind etc.”<sup>541</sup>. Hence, according to Nida-Rümelin, in experiencing, we are always implicitly aware and familiar with what it means for a subject to be a subject.

Yet, despite all these evident parallels with self-experientialism, there is at least one fundamental disagreement that ultimately renders Nida-Rümelin’s account opposed to—what is in fact—the *main* idea of self-experientialism, i.e., the notion that selfhood *consists* in self-experience. Nida-Rümelin explicitly denies that this is the case:

It is in virtue of the fact that the subject S is aware of itself as the one who is presented with blueness that there is a subject at all, involved in that event, who is presented with blueness. Can we make sense of that claim? I suggest that we cannot.<sup>542</sup>

---

<sup>536</sup> Nida-Rümelin 2017, 62.

<sup>537</sup> Nida-Rümelin 2017, 64.

<sup>538</sup> Nida-Rümelin 2017, 62.

<sup>539</sup> Nida-Rümelin 2017, 81.

<sup>540</sup> Nida-Rümelin 2017, 75.

<sup>541</sup> Nida-Rümelin 2017, 75.

<sup>542</sup> Nida-Rümelin 2017, 70.



The reason she denies a constitutive function of self-experience for selfhood is that she wishes to make a distinction between *basic intentionality* and the *awareness of basic intentionality*:

But basic intentionality and awareness of basic intentionality are not the same thing which means, to put it more precisely: for an event to exhibit basic intentionality does not consist in there being a subject who is aware of being in the subject position of that metaphysical structure. The metaphysical fact the subject is thereby pre-reflectively aware of does not consist in the fact that the subject is aware of it.<sup>543</sup>

Accordingly, Nida-Rümelin is skeptical about the phenomenological notion of intentionality, according to which intentionality *is* the conscious directedness at an object *x*. In fact, she explicitly mentions Zahavi and attributes to him what she takes to be a “mistake”<sup>544</sup>, namely the equation of intentionality and awareness of intentionality. While Nida-Rümelin concedes that the *metaphysical* fact of intentionality perhaps might necessarily involve the *phenomenological* fact of awareness of intentionality, she is adamant that the latter does not constitute the former. Now, to make that distinction, she emphasizes that identifying *intentionality* with *awareness of intentionality* would necessarily amount to the view that it is the self-awareness of a subject that makes that there is a subject at all. Denying that such could be the case would thus lead the identification of intentionality and awareness of intentionality ad absurdum.

What is Nida-Rümelin’s argument? Why shouldn’t selfhood be constituted by self-experience? Unfortunately, Nida-Rümelin’s refutation is sparse:

It seems obvious that the subject’s awareness of itself as the one presented with blueness does not constitute the fact that there is someone who is presented with blueness; rather, the subject’s awareness of itself as someone presented with blueness presupposes that there is someone who has that property.<sup>545</sup>

Evidently, one might hold the view that in order to be self-aware, a subject must exist before it may become self-aware. However, it is far from ‘obvious’ that this claim holds true and further reasoning would be required. Since Nida-Rümelin does not present further arguments, it is difficult to fully assess her case. Even so, what can be said is that taking the subject to exist *prior* to self-awareness has some ramifications that seem not to fit well with Nida-Rümelin’s own account of pre-reflective self-awareness.

To see what they are, let me briefly mention a classical position: Immanuel Hermann Fichte’s. In his *Wissenschaftslehre nova methodo*, Fichte addresses precisely the

---

<sup>543</sup> Nida-Rümelin 2017, 69.

<sup>544</sup> Nida-Rümelin 2017, 69 Fn.

<sup>545</sup> Nida-Rümelin 2017, 70.

objection that the subject must exist prior to it being possibly self-aware. Trying to overcome Kant's verdict that we *cannot* gain knowledge of the nature of selfhood (i.e., the pure 'I') through our self-experience, Fichte famously argued that we do possess an immediate consciousness of ourselves (*intellectual intuition*) and that our self-knowledge does precisely not stem from a higher-order reflective act that turns back on a previously existing 'I'.<sup>546</sup> To be able to launch the claim that we are *immediately* conscious of ourselves and that we are aware of our own nature, Fichte held that there is no 'I' as substance existing prior to the consciousness of the 'I' but rather that the consciousness of the 'I' *is* the 'I':

The I becomes an I only by means of an act of self-positing. It is not already a substance in advance of this act of self-positing; instead, its very essence is to posit itself as positing. These are one and the same. Consequently, *the I is immediately conscious of itself*.<sup>547</sup>

On this famous view, it is in the consciousness of *x* that the 'I' posits itself in the first place, and so only exists in the context of consciousness. The primary way of self-consciousness, on Fichte's account, is the consciousness of an activity of self-positing. And this, by the same token, presents also the being of the 'I'.<sup>548</sup> Just as for Locke person/self and consciousness are not separable, for Fichte I and consciousness are inseparable. In both thinkers, the substance that may underlie consciousness and self-consciousness matters not, and may not be taken into the definition of the 'I' or self. It is precisely *this* move that allows both thinkers to claim that we can gain insight into our nature. Fichte decisively emphasized that Kant's verdict of the elusiveness of the nature of the 'I' is unacceptable as well as is the assumption that self-awareness comes about through a higher-order reflective act.<sup>549</sup>

Now, what does all this have to do with Nida-Rümelin's position? The point is that she offers a position that involves claims of both sides—Kant vs. Locke/Fichte—that, as I want to suggest, are contradictory. On the one hand, she highlights that the subject must exist prior to it being self-aware, i.e., she re-introduces a distinction between a quasi-

---

<sup>546</sup> Cf. Lang 2020, 95.

<sup>547</sup> Fichte 1992, 114.

<sup>548</sup> Evidently, Fichte's view is very similar to self-experientialism and presents a forerunner for Husserl's phenomenology of self-consciousness and Zahavi's minimal self alike, as Green (2010) convincingly carves out. However, there are also significant differences. One central difference is that Husserl does not take the unifying self-constitution of the experiential flow and its self-giveness to be a result of activity by the 'I', the ego (cf. Zahavi 2014a, 65). See also Waibel, Breazeale and Rockmore (2010) for a collection of essays dealing with the relationship between Fichte and Husserl. The important point here is not the precise parallels and disagreements between the German idealist and the phenomenologist, but rather their shared move to undermine the doctrine that holds that the self must exist before it can become self-aware by arguing that the mode of existence of self consists in self-awareness.

<sup>549</sup> Henrich 1966, 1971, Frank 1991a, b, 2015.

Kantian subject *as it is in itself* and the subject insofar *as it appears to itself*. On the other hand, she shares the view with Fichte (and Locke) that self-experience is nature-revealing with regard to our being. However, it is far from clear how both claims are compatible with each other. In fact, I don't think they are. For, if Nida-Rümelin wishes to make the claim that the subject exists prior to it being self-aware, then the question arises as to what exactly this being prior to self-awareness ought to consist in. As I have emphasized throughout the thesis so far, self-awareness presents us with the experiential process, the stream of consciousness. And crucially, Nida-Rümelin herself holds that self-awareness is an awareness of oneself as a conscious being. Now, if she additionally wants to hold that such an awareness—which she does—is nature-revealing, she then is ultimately committed to the view that the nature consists precisely in the conscious being. The being conscious is not a contingent feature of the subject—it is its nature. This is precisely both Locke's and Fichte's point. Hence, not only does Nida-Rümelin owe an explanation of what the being of the subject consists in if it ought not to be self-awareness, she must also explain how self-awareness can be nature-revealing without itself being the nature of the subject.

Another closely related problem that emerges is that by distinguishing between being a subject and the self-awareness of the subject, Nida-Rümelin re-introduces a representational framework, which of course is a possible position that one might want to defend. Yet, when she is eager to denote pre-reflective self-awareness as *immediate*,<sup>550</sup> such a strategy doesn't seem to be available for her. For, if the self-awareness is *not* the being of the subject, how exactly can the self-awareness of experience still be an immediate awareness of oneself? If the being of the subject ought to be prior to it being self-aware, this renders self-awareness a secondary—and as it must seem—contingent feature of the subject. Ultimately, then, Nida-Rümelin seems to re-introduce some kind of a *substantial* understanding of the subject that is in itself and may become self-aware. Self-awareness, then, seems to be understood in terms of a representation of the subject. However, to stress again, Nida-Rümelin doesn't explain what the subject's being independently of it being represented is. Ultimately, thus, the subject and its nature turn out to be elusive in her account, rendering it incoherent and therefore untenable, as she advertises her notion of pre-reflective self-awareness as nature-revealing, while she at the same time maintains that that what is revealed in pre-reflective self-awareness, i.e., the stream of consciousness, is *not* what constitutes the being of the subject.

To conclude, Nida-Rümelin presents an experiential approach to selfhood insofar as she claims that it is via the phenomenology of our self-experience that we can gain knowledge about the nature of selfhood. In fact, she claims that we are in any moment of

---

<sup>550</sup> Nida-Rümelin: "One may say that the subject is immediately aware—without reflection and without conceptualization—of every specific aspect of what it is like for it to live through moment *m* while it is living through moment *m*." (2017, 65).

our conscious life implicitly and immediately aware of ourselves as consciously being and that this kind of awareness is nature-revealing. I suggest, once one subscribes to such a view, then one cannot argue any longer that behind and prior to our self-awareness there is a subject that exists prior to it being conscious and self-aware.

## 5.4 Dainton: Substrate phenomenalism

Let me now turn to a proposal that also emphasizes the stream of consciousness, phenomenal consciousness, in short, experience as the point of departure for the conceptualization of selfhood: Barry Dainton's "C-theory"<sup>551</sup>. Dainton's aim is to understand selfhood in terms of experience as much as possible, but ultimately comes to accept the view that selfhood needs to be conceptualized in non-conscious terms. Unlike Nida-Rümelin, however, he is fully aware that his notion of selfhood refers to something that is itself not experiential but rather *underlies* experience, for which reason Tim Bayne has dubbed his account "substrate phenomenalism"<sup>552</sup>. But Dainton is not only aware that his account ultimately targets a non-conscious *x* to play the role of self, he explicitly endorses this view and provides arguments in favor of it, which makes it an interesting anti-thesis to the self-experientialism I have proposed on the basis of Zahavi's minimal self-view.

Dainton's distinction between two groups of experience-based approaches to selfhood is helpful to further clarify the dialectical constellation. The first group of accounts conceives of the self as being *essentially conscious* (ECS). The second group of accounts suggests that the self is a *potentially conscious being* (PCS).<sup>553</sup> Self-experientialism would qualify as an ECS-approach, while Dainton's C-theory is advertised as a PCS-approach. Dainton, taking Zahavi's ontological commitment to be insufficiently clear, however, prefers to take Strawson's "Pearl view"<sup>554</sup> as a proxy for the ECS-side.<sup>555</sup> Let me unpack his line of argumentation against Strawson as it similarly applies to self-experientialism.<sup>556</sup>

Dainton agrees with Strawson (and self-experientialism) in that we often have self-experience: "[w]e are beings who *can* feel present and alive in our current experiences"<sup>557</sup>. Yet, he criticizes that such a form of self-experience should be essential to the being of

---

<sup>551</sup> Dainton 2008, 112, 113.

<sup>552</sup> Bayne 2010, 287.

<sup>553</sup> Dainton 2008, 79.

<sup>554</sup> Strawson 2017, 37.

<sup>555</sup> Dainton 2008, 145.

<sup>556</sup> I will discuss Strawson's *pearl view* in the next section and work out the differences to self-experientialism.

<sup>557</sup> Dainton 2008, 146.

selfhood. Even if it were the case that all experience would involve self-experience (an assumption which he rejects), he sees a wrong conclusion from self-experience in which we experience ourselves as experiencing beings to the idea that such an experience is essential to selfhood:

[H]ow plausible is it to suppose that we must actually be experiencing such feelings in order to continue to exist? Isn't it enough that we retain the *capacity* for such feelings?<sup>558</sup>

Dainton takes Strawson's argument in favor of the ESC-approach to consist mostly in the idea that it is precisely self-experience that reveals our nature as essentially conscious beings. Self-experientialism presents a similar proposal, according to which we can become aware of our nature by looking into the phenomenology of self-experience that reveals to us the nature of selfhood as conscious being. By contrast, Dainton puts forward an objection against a conclusion from self-experience:

Since a self can exist when conscious and when unconscious, experience reveals at most one aspect of what is involved in a self's existing. Our experience reveals only what it is like to be a *conscious* self, that is a self that is in the process of *having* experiences—although if the C-theory is correct, selves 'have' experiences by *producing* them. Experience can reveal no more than this, but since there is nothing that it is like to be a [sic!] unconscious self, this is only to be expected.<sup>559</sup>

Dainton's point, thus, is that self-experience not surprisingly cannot reveal the non-experiential components of self, given that they are precisely non-experiential. Accordingly, he maintains that we cannot fully rely on our self-experience when it comes to the knowledge and conceptualization of selfhood and its nature.<sup>560</sup> Rather, on his view, we must also consider the non-experiential parts of self. In fact, as his own C-theory maintains, the non-experiential parts are what is essential about selfhood: the *capacities* for consciousness. These capacities are not simply possibilities but rather that "what (metaphorically speaking) lies *immediately beneath* our actual streams of consciousness: those systems of experiential capacities which are causally responsible for these streams"<sup>561</sup>. Although Dainton considers his approach as a form of phenomenalism that construes selfhood in departure from the stream of consciousness, his emphasis lies on that what underlies the stream of consciousness. On his view, the stream of consciousness and the different experiences that figure in it are produced by what he calls "*experiential*

---

<sup>558</sup> Dainton 2008, 147.

<sup>559</sup> Dainton 2008, 147.

<sup>560</sup> That places him near those skeptical accounts that I discussed in Chapter 3.

<sup>561</sup> Dainton 2008, 132.

*powers*<sup>562</sup>. Now, crucially, on Dainton's view, selves are so-called "C-systems" that are defined as "clusters of experiential powers"<sup>563</sup>. As he further claims, what binds together different experiential powers to a cluster are their capacities to generate co-conscious experiences, i.e., a unified stream of consciousness.<sup>564</sup> It is in this sense that his position remains a form of phenomenalism: It is the phenomenal unity of the stream of consciousness that constrains the clusters of experiential powers. Rather than a bundle of experiences, then, the self is a bundle of experiential powers that present the capacities to produce unified experiences.

Evidently, the move from *actual* consciousness to *potential* consciousness renders Dainton's approach ultimately a non-experiential approach to selfhood, despite its emphasis on the stream of consciousness. Why should we take such a step? After all, though it might be true that self-experience cannot reveal non-experiential components of selfhood, this doesn't mean we have to assume the self to be really non-experiential. Or do we?

Dainton's motivation for preferring the PCS- over the ECS-approach is that construing selfhood along these lines, namely as "enduring *potentialities* for experience"<sup>565</sup>, allows us to solve the challenge of bridging the interruptions in the stream of consciousness. Put differently, Dainton sees the need to account for the challenge of personal identity across phases of non-consciousness. As he seems to believe, this challenge applies to the ECS-approach in a similar way in which Pierre Gassendi famously formulated his worries about Descartes' notion of the inseparability of self and thinking (= cogitare = consciousness):

You add that thought alone cannot be separated from you. ... Nonetheless I want to stop here and ask whether, in saying that thought cannot be separated from you, you mean that you continue to think indefinitely, so long as you exist. ... But it will hardly convince those who do not see how you are able to think during deep sleep or indeed in the womb.<sup>566</sup>

Like Gassendi, the problem Dainton sees is that while I am asleep—or under the influence of a strong and general anaesthetic—in which I am not conscious of anything, according to the ESC-approach, *I* strictly am not real, given that selfhood ought to be tied to experience. Hence, no experience, no self. Moreover, when I wake up after the effect of the anaesthetic has perished, we would expect I wake up experiencing myself in continuance of the stream of consciousness that had come to a temporary break when I got the anaesthetic. But what

---

<sup>562</sup> Dainton 2008, 87.

<sup>563</sup> Dainton 2008, 99.

<sup>564</sup> Dainton 2008, 112.

<sup>565</sup> Dainton 2008, 134.

<sup>566</sup> Gassendi 2005, 184.

is it that enables the persistence across periods of non-consciousness? Dainton gives the following answer:

I have fallen into a dreamless sleep that will last for half an hour. Simplifying only a little, according to the C-theory what keeps me in existence after my loss of consciousness is the fact that the system of experiential capacities that was causally responsible for my waking experience continues to exist. ... So long as this C-related power-system persists, I persist.<sup>567</sup>

Dainton holds that it is the experiential capacities<sup>568</sup> that ensure my persistence during non-conscious states. During the effect of the anaesthetic that what causally produces my experience and self-experience is still intact but only inhibited. Therefore, on Dainton's view, I persist—and come back to mind when the inhibition ends.

I think so far, the depiction can easily convince most of us. For, even though one might debate whether sleep or when under the influence of a strong anaesthetic one is indeed fully non-conscious,<sup>569</sup> most “have little hesitation in assuming that we could survive a temporary loss of consciousness”<sup>570</sup>. And such a survival plausibly is made possible by the uninterrupted continuance of *x*, whereby *x* is responsible for the generation of experience and self-experience.

But Dainton says more than that. His point is not only that *x*, the experiential capacities, enable the persistence of self. No, from this claim, he derives the further claim that that what enables experience and self-experience *is* the self. Accordingly, he identifies the experiential capacities with selfhood. That means, during non-conscious phases, selfhood doesn't cease to exist, it just transits from a wakeful state into a non-conscious mode. The latter is as much a valid mode of self as its wakeful, self-conscious life.

Whereas it might seem plausible to say that some *x* is responsible for the generation of experience and self-experience, *defining* selfhood in terms of *x*, is a far stronger thesis, which is not implied in the claim that *x* enables the persistence of self. Dainton's argument for this stronger thesis is somewhat pragmatic. The point is that *if* we construe selfhood in terms of experiential capacities (= *potential* experience) rather than *actual* experience, then

---

<sup>567</sup> Dainton 2008, 132.

<sup>568</sup> For Dainton, the experiential capacities are not conceptually equivalent or congruent with the brain or the body as a biological system that can generate experiences. The latter are *just an example* of the former. The collection of experiential capacities that constitute my being could, on his terms, be transferred to any conceivable *x* that is apt to host or realize the experiential capacities that *are* me. See Dainton (2008, 87): “Would your existence be threatened if your experiential capacities were to be transferred to a consciousness-generating machine, or a matrix of electromagnetic fields, or a crystalline substance, or an immaterial substance? Quite generally, *any* physical change seems survivable provided it is clear that the change in question does not put an end to one's capacity to have phenomenally continuous experience.”

<sup>569</sup> Cf. Fuchs 2017b, 299.

<sup>570</sup> Bayne 2010, 285.

the pressure stemming from the challenge of the persistence of self across non-conscious phases seems to be relieved.

Dainton might be right in saying that taking the self resp. selfhood to consist in a non-conscious  $x$  (whether this be considered as a soul-like experience-transcendent  $x$  or the brain or a collection of particular experiential powers) that generates experiences and self-experiences would amount to a possible response to the “bridge-problem”<sup>571</sup> that the interruptions of consciousness, he believes, pose. However, such a solution renders once again selfhood elusive: What exactly is the being of selfhood prior to actual consciousness? And how is this *me*, the one who I experience myself to be *in* experience? All what is said about the elusive approaches to selfhood in Chapter 3 similarly apply to Dainton’s proposal. Let me add a few further responses to be found in the literature.

First, let us hear out Strawson who—as a proponent of an ESC-approach—offers a rejoinder to Dainton:

[H]e [Dainton, P.S.] defines the self or subject of experience wholly dispositionally, as a collection of potentialities. His view accords with the idea that an entity that was a subject of experience could exist without ever actually having any experience. My focus on the thin conception of the subject leads me to say ... that no actual subject of experience ever exists in a universe in which millions of Daintonian subjects of experience exist but never have any experience.<sup>572</sup>

As Strawson remarks, according to the view Dainton defines selfhood, it would be quite thinkable that there are selves that never have been and never will be conscious but nonetheless are real. Zahavi concurs when he says:

I wouldn't recommend settling for and making do with a notion of selfhood that is also applicable and ascribable to a philosophical zombie (who, although it lacks experiences, is presumably still a living organism).<sup>573</sup>

The point is that if we were to accept Dainton’s proposal, there would be no problem in assuming that there are selves who resemble zombies: they are embodied, their organism functions and they may self-preserve on a biological level, but never have any experience whatsoever. Yet, most of us will strike such a consequence as absurd. This can further be illustrated by the following thought experiment:

Consider, for instance, you love your partner just as much as you could possibly love a person. Your partner feels the same way about you. As a consequence, you wish to spend your life together and you both hope it lasts as long as

---

<sup>571</sup> Dainton 2008, 27, 75f.

<sup>572</sup> Strawson 2009, 370.

<sup>573</sup> Zahavi 2014, 76.



possible. But life is dangerous and the risk of either you or your partner passing away is fairly high. Now, imagine, after a year of being with your partner and further assuring that the other is ‘the one’, someone offers you to plug you and your partner on a medical device that induces an artificial coma in a way that your brain and body remains unharmed, keeping your *capacities* for (self-) experience intact though without either of you being in fact phenomenally conscious during the coma. This way the risk of the end of life is minimized. But not only that, doctors also assure you that this way your brain and body receive best treatment resulting in an average life expectancy of 300 years. During the coma, you lie in bed holding hands with your partner. After the promised 300 years, your brain and body has aged and cannot be kept in maintenance despite the best medical care available. The doctors turn off the medical device and bury the corpses. Have you had a long life with your partner? Did you spend the 300 years together? Would you conceive that as an option for your love?

On Dainton’s C-theory, such a scenario wouldn’t be a case in which the self is less real than in a wakeful life. Now, one could say that selfhood consists in both conscious and unconscious modes of being and we only happen to value the wakeful life, wherefore the scenario presented strikes us as bizarre. For, we can hardly imagine anyone who would prefer having or being a healthy, long-lasting organism over enjoying a shorter period of wakeful life. We may grant that in the scenario the *possibility* for the life of the persons (and selfhood) is retained throughout, but if they turn out to never be conscious again after they have been put to dreamless sleep, most will be hesitant, I submit, to claim that after entering the sleeping phase they had a life. It’s true that they cannot be said to be dead either, because they retain the possibility of experience. Still, I think that if one were to wake them up after 300 years shortly before their organisms cannot be kept alive any longer, a feeling of having missed one’s own life will dominate. If asked how the non-conscious phase was, any person cannot but say that it was nothing ‘*to me*’, ‘*I was not there*’.

Finally, another criticism of Dainton’s approach can be found in Bayne’s discussion of different forms of experiential approaches to selfhood. As Bayne emphasizes, the problem with substrate phenomenalism is that it seems that there is no problem in thinking that one substrate—a certain collection of experiential capacities—could give rise to several and different streams of consciousness at the same time. Accordingly, Dainton’s definition of selfhood would imply that that what is essential about selfhood fails to provide the principium individuationis of self.<sup>574</sup>

All in all, Dainton’s C-theory, although offering a possible solution to the bridging problem, ultimately engenders bizarre and absurd consequences that should be avoided. Although experiential capacities might be said to be required for the being of selfhood, given that without experience there is no self in the sense of an experience of reflexivity

---

<sup>574</sup> Bayne 2010, 287ff.

(as self-experientialism has it), the experiential capacities and their disposition for experience is *not* itself the being of selfhood.

## 5.5 Strawson: The pearl view

Is there another way to address the issue of personal identity and interruptions of the stream of consciousness? Strawson's account implies a positive response to that question. Let's look how he construes selfhood and handles phenomenal gaps. As I have mentioned before, Strawson's account qualifies as an ESC-approach, and takes experience to be essential to selfhood. Like self-experientialism, on his view, not only does experience necessarily involve selfhood, selfhood consists essentially in having experience and so is not something existing independently of experience. What is special about his account is that he believes that each human person comprises of a plurality of different subjects of experience, which he calls SESMETs (Subjects of Experience that are Single Mental Things):

A sesmet, then, is a subject of experience as it is present and alive in the occurrence of an experience. It is as EEE—as embedded, embodied, and ecological—as anyone could wish. There cannot be a sesmet without an experience, and it is arguable that there cannot be an experience without a sesmet.<sup>575</sup>

Strawson takes it that each sesmet endures only as long as any period of uninterrupted consciousness is. Consequently, whenever there is an interruption of our experiential life, a new sesmet arises upon a new period of continuous consciousness. On his view, such an approach is backed by the following phenomenological observation:

I find that my fundamental experience of consciousness is one of *repeated returns into consciousness from a state of complete, if momentary, unconsciousness*. The (invariably brief) periods of true experiential continuity are usually radically disjunct from one another in this way, even when they are not radically disjunct in respect of content.<sup>576</sup>

Hence, Strawson does not only believe phenomenal gaps, interruptions in experience, are possible, he takes them to occur quite often, questioning the idea of a *stream* of consciousness.<sup>577</sup> Instead of a single continuous stream, he takes experiential life to be pervaded by recurring alternations between continuity and discontinuity, whereby distinct

---

<sup>575</sup> Strawson 2017, 62.

<sup>576</sup> Strawson 2017, 34.

<sup>577</sup> Cf. Chapter 2.

phases of continuity may share the same experiential content with each other. One ramification of such a view is that the *challenge* of personal identity seems to dissolve. For, on Strawson's view, each continuous phase of consciousness presents its own ontologically distinct *sesmet*, wherefore, strictly speaking, there seems there is no persistency across phases of non-consciousness. With Locke, thus, he shares the view that the continuity of consciousness does unify what or who oneself is. But he disagrees that personal identity can, strictly speaking, stretch across interruptions of consciousness. Accordingly, on such a line of reasoning, it would be futile to keep looking for any bridge that connects the discrete phases of continuous conscious life. His position is inspired by William James:

In essentials I agree with James ... . Using the word 'thought' in the wide Cartesian sense to cover all types of conscious episodes, he claims that each 'perishing' pulse of thought' is a self, and in a famous phrase, says that 'the thoughts themselves are the thinkers'.<sup>578</sup>

The idea, thus, is that instead of making an attempt at finding a bridging-principle that allegedly guarantees us the numerical identity with the self of our past experiences, we should rather come to acknowledge the fact that such a numerical identity is simply not available—and that there is nothing odd about it. We as human beings simply consist of a high number of ontologically distinct *sesmets*, independently of whether we develop a sense of identity or not. In fact, as Strawson emphasizes, human persons differ significantly with regard to how they experience identity. While some may lead a rather *episodic* life and don't feel strongly continuous with the self that features in their memories of past experiences, others have a stronger sense of being *diachronically* identical with their past selves.<sup>579</sup> Accordingly, Strawson suggests we all, concerning our self-experience, are located on a spectrum between Episodics and Diachronics.<sup>580</sup> Whereas there might be phenomenological differences in the self-experience of identity, Strawson holds that, ontologically speaking, Episodics as well as Diachronics are composed of a myriad of distinct *sesmets*. Strawson proposes to “call this view the *Pearl view*, because it suggests that many mental selves exist, one at a time and one after another, like pearls on a string, in the case of something like a human being”<sup>581</sup>.

Does the pearl view present an alternative to self-experientialism that can accommodate the issue of personal identity? As Zahavi highlights, the minimal self-view—and so self-experientialism—is “remarkably similar”<sup>582</sup>, but shows “one substantial

---

<sup>578</sup> Strawson 2017, 61.

<sup>579</sup> Strawson 2017, 51f.

<sup>580</sup> Strawson 2017, 54.

<sup>581</sup> Strawson 2017, 37, my emphasis.

<sup>582</sup> Zahavi 2014, 71.

difference, which concerns the question of persistency and transience”<sup>583</sup>. Unlike Strawson, Zahavi holds that “our experience of the temporally extended and unified stream of consciousness is *eo ipso* an experience of the real (and not merely illusory) diachronicity of the experiential self”<sup>584</sup>. Accordingly, he is skeptical about the need to posit an intrapersonal plurality of selves that somehow make up together the *appearance* of a diachronically unified self. In fact, as he emphasizes, such a view, he takes it, has the consequence that “we are dealing with selves that, although they might stand in a unique causal relationship to one another, are still as different from each other as I am from you”<sup>585</sup>, which, according to Zahavi, renders Strawson’s account an “absurd”<sup>586</sup> proposal.

I think Zahavi is right that such a consequence would be hard to acknowledge and must be considered a—perhaps too—high price for the possibility to accommodate the issue of personal identity (which would, strictly speaking, not be solved but rather dismissed). However, Zahavi’s particular criticism seems to me to be too strong, for the relationship between what are *my* sesmets and the relationship between myself and another person is not on a par. My sesmets are “non-overlapping”<sup>587</sup>, and so follow each other. Their relationship, accordingly, is never such that holds between distinct and commonly present conscious subjects, as it is the case when I interact with another person. Moreover, I do have a kind of first-person access to the first-person givenness of any of my sesmets (at least potentially) when I remember my past experiences carried out by a distinct sesmet, different from the sesmet I am now. By contrast, I never have such a kind of access when it comes to the experience of others. Hence, the relationship between my sesmets is more intimate and intertwined than the one I entertain with others. Crucially, then, my sesmets not only are connected to each other through a ‘unique causal relationship’ that they do not have with the sesmets of another person; they are also bound together through a form of access and sense of ownership that is characteristic for the first-person perspective.

That said, Zahavi, correctly in my view, points to the fact that Strawson’s move to subdivide a human being into ontologically distinct sesmets brings another question on the plan: How is the relationship of the distinct sesmets to be conceived of? Most importantly, I submit, although Strawson has ridded himself from the issue of the identity-preserving bridging, another bridging problem seems to emerge: How can the experiential *contents* of one conscious phase be passed over to the succeeding subject of experience? Crucially, Strawson explains the “apparent continuity of experience ... and the consistency of perspective across selves” by “the fact that sesmets ‘appropriate’—in James’s word—the

---

<sup>583</sup> Zahavi 2014, 71.

<sup>584</sup> Zahavi 2014, 74.

<sup>585</sup> Zahavi 2014, 74.

<sup>586</sup> Zahavi 2014, 74.

<sup>587</sup> Strawson 2017, 73. Strawson doesn’t exclude the possibility that sesmets could exist in a non-overlapping, parallel fashion.

experiential content of the experiences of their predecessors”<sup>588</sup>. How does such an appropriation work according to Strawson? For him, we are not faced with a challenge, as he takes the appropriation to occur “in a way that is entirely unsurprising insofar as they arise successively, like gouts of water from a *rapidly sporadic fountain*”<sup>589</sup>.

Strawson, thus, does not only offer an alternative *metaphysical* view of selfhood (compared to Zahavi’s minimal self-view), it is also important to bear in mind that his different metaphysical view is based on a different description and interpretation of the *phenomenology* of self-experience. It is his phenomenological account that ultimately must buttress his claims. Let me therefore discuss the tenability of his phenomenological views and the conclusions for the metaphysics of selfhood he draws from them. More specifically, the question is whether his alternative metaphor of a *sporadic fountain* is more convincing than the one of the *stream of consciousness*—and whether it is apt to explain the unified experiential content across succeeding but allegedly ontologically distinct *sesmets*.

First, however, I wish to highlight that I am persuaded by much of the descriptions Strawson offers us with regard to the way in which we experience ourselves. I do think that—possibly—every human person has at least once in their life felt that she is not, strictly speaking, identical with the self of remembered past experiences. Think of what one could call the ‘mirror-experience’: You look into the mirror and you recognize yourself, but you still feel somehow an awkward, non-pathological self-alienation arising in yourself. The feeling consists in a deep puzzlement about you ‘being you’, about the fact of life, and, for instance, that you have the name XY. You remember all what is typical to remember for a person of your age, and you are even aware of this, but still—perhaps only for a few seconds—you find this odd, and there is a slight sense that you—the one who is looking into the mirror—came into existence just now, thrown, as it were, into a story of a life that has already begun. I believe such experiences are common even among persons who are rather on the Diachronic’s end of the spectrum. Moreover, similarly I believe, most people sometimes recollect their past, but feel an odd, non-pathological alienation of who they were, how they took decisions, and why. Even though they ‘know’ these memories are non-deceptive, they feel distinct from rather than identical with their past selves. Crucially, such a distant feeling to one’s past selves may persist even if one is aware and recognizes that one’s current way of experiencing the world is hugely influenced by one’s past selves. I might remember exactly how it came that I decided to move to a foreign country, possibly even feel what Schechtman calls “empathic access”<sup>590</sup> with my past self, and may be aware of how the developments from there on have made me the specific person I am now—and *yet* feel somehow discrete and detached from my past selves.

---

<sup>588</sup> Strawson 2017, 61.

<sup>589</sup> Strawson 2017, 61, my emphasis.

<sup>590</sup> Schechtman 2001.

I would grant all these examples are well-explained by or at the least compatible with Strawson's distinction of episodic and diachronic forms of life, which I find phenomenologically convincing. By contrast, I am less convinced that

- (a) the recognition of these different ways of experiencing oneself commits one to accept the Pearl view, i.e., the notion that a human being comprises ontologically distinct *sesmets*;
- (b) accepting the distinction between Episodics and Diachronics is reliant on Strawson's notion of a sporadic fountain of consciousness;
- (c) the sporadic fountain of consciousness is a metaphor to be preferred over the stream of consciousness.

The focus of the following discussion, therefore, lies on the phenomenological concept of the sporadic fountain of consciousness, as it is also the basis on which Strawson's metaphysical claim of *sesmets* rests. For, Strawson's rationale runs thus: If, phenomenologically speaking, our experience is such that it is characterized by discrete phases or 'pulses' of consciousness, then it is legitimate to identify each pulse with an ontologically distinct subject: a *sesmet*. I agree with Strawson that *if* the phenomenology were to consist in a staccato-like, gappy series of islands of consciousness, as it were, *then* there would be a phenomenological motivation for the Pearl view. However, I think, as the following issues show, it is far from clear that the Strawson's interpretation of lived experience as a fountain rather than a stream of consciousness is phenomenologically accurate.

It is important to note that Strawson's proposal is somewhat different from the skeptics about the stream of consciousness that I discussed in Chapter 2. While he does find talk of a seamless stream unwarranted (given his emphasis on 'pulses' of consciousness), he does *not* deny any continuity and experience thereof whatsoever. He acknowledges there are shorter periods of time in which there is "*true* experiential continuity"<sup>591</sup>. Apart from that, the very distinction between Episodics and Diachronics implies the recognition of the possibility of a *felt* sense of diachronicity, i.e., continuity. Moreover, he speaks of an "*apparent* continuity of experience"<sup>592</sup>, which implies the acknowledgment of a further—though fallacious—kind of experience of continuity. Crucially, while he does share the belief with the skeptics that we sometimes are susceptible to the illusion of continuity, he concedes that not all experiential continuity is deceptive. Although Strawson's account is much more differentiated than those of the skeptics from

---

<sup>591</sup> Strawson 2017, 34, my emphasis.

<sup>592</sup> Strawson 2017, 61, my emphasis.

Chapter 2, his depiction of the phenomenology of our experience is equally afflicted with theoretical frictions.

### 5.5.1 The constancy of objects implies experiential continuity

A first example of these frictions, I take it, is inherent to Strawson's explanation for the genesis of the apparent continuity of experience. As Strawson emphasizes, it is rather the persistence of objects on the basis of which we develop a *belief* in experiential continuity:

If there is any support for belief in the long-term continuity of the self in the nature of moment-to-moment consciousness, it is derived indirectly from other sources—the massive constancies and developmental coherencies of *content* that often link up experiences through time, and by courtesy of short-term memory, across all the jumps and breaks of flow.<sup>593</sup>

Strawson suggests that rather than a direct and immediate experience of continuity, we come to *believe* in experiential continuity, given that at different points in time we are confronted with objects that remain the same across time. Hence, Strawson seems to follow the traditional and Kantian view according to which it is the persistence of objects in the world through which the experience of movement and time is possible to begin with. As Kant famously argues in his 'refutation of idealism' in his *Critique of Pure Reason*, the consciousness of oneself as a being in time presupposes the affection and experience of persistent spatiotemporal objects.<sup>594</sup> Kant's move is not uncontroversial, and one might ask whether the experience of a persistent object doesn't require the experience of a continuous flow to begin with. *That* is precisely the point I have made with Husserl and Zahavi in Chapter 2, and it applies here equally. For, the whole crux of Husserl's analyses of inner time-consciousness was to explain how the experience of objectivity is possible to begin with. The results of these analyses were that whenever there is an experience of a persisting object, we are dealing with a *temporal* object. Husserl's answer to the question how temporal consciousness is constituted was that it is to be described in terms of a temporal field—stretched out by the three non-independent moments of retention-impression-protection. Importantly, this temporal field implies the auto-appearance of the stream of consciousness. Any object qua persistent, i.e., temporal, object, thus, on this view, is given always only insofar as the temporally flowing character of consciousness is manifest too. It is only against the background of the manifestation of flow that something can be experienced *as persistent* to begin with.<sup>595</sup>

---

<sup>593</sup> Strawson 2017, 35.

<sup>594</sup> Kant 1996/1781, 288ff. (B 274ff.).

<sup>595</sup> Cf. Schmidt 2019.

The point, then, would be that if Strawson wants to maintain that our belief in experiential continuity is derived from the experience of persistent objects, the question arises as to what is presupposed by the experience of persistent objects. And, if Husserl and Zahavi are right, which I believe, then the experience of persistent objects intrinsically involves the manifestation of the flow, as the experience of the persistence of objects is constituted temporally, i.e., by inner time-consciousness, the stream of consciousness. Strawson, I suggest, is right to notice the connection between persistent object and our experience of the flow. However, I think he is wrong to ascribe a logical and phenomenological primacy to the object. It is not that we first experience persisting objects and only *then* come to believe (or experience) a stream of consciousness. Rather, both are phenomenologically correlative.

Hence, whenever Strawson refers to the constancies of objects, he does, by the same token, I submit, refer to the experience of continuity. The following example of Strawson is revealing in that regard:

[C]onsciousness is ‘in a perpetual flux’, and different thoughts and experiences ‘succeed each other with an inconceivable rapidity’. And yet one is experientially in touch with a great pool of constancies and steady processes of change in one’s environment including, notably, one’s body (of which one is *almost constantly aware*, however thoughtlessly, both by external sense and by proprioception.<sup>596</sup>

Strawson points to the fact that no matter how quickly our thoughts and perceptions may follow each other, no matter how disruptive our experiential life, there are always also constancies that accompany the process of changes. For instance, the constancy of one’s own body. Most importantly, to emphasize the constancy of the body, Strawson highlights that one is ‘almost constantly aware’ of one’s own body. The point is that Strawson seems to acknowledge that the constancy of the body is *per se* associated with a constant awareness of the body. It is precisely because we constantly feel our body that we experience the body as constant. It is not that we compare our experiences, say, at discrete points in time  $t_0$  and  $t_1$ , and then conclude from a higher-order pole, in reflection, as it were, that both at  $t_0$  and  $t_1$  there were each an object resembling each other: our body, which can be inferred to be the same in both cases. Rather, the experience of the constancy of the body *consists* in the constant awareness of the body, which implies experiential continuity and experience of continuity. Hence, when Strawson stresses objective constancies, he implicitly emphasizes the experience of continuity that is intrinsic to the experience of objectivity.

---

<sup>596</sup> Strawson 2017, 35f., my emphasis.



### 5.5.2 The gaps are in the content, not the stream

Similarly, I believe, this might challenge Strawson's verdict that our experience consists of disjunct moments of consciousness that follow each other, disrupted by short periods of unconsciousness. This is how Strawson depicts this view:

[C]onscious thought has the character of a (nearly continuous) series of radically disjunct irruptions into consciousness from a basic substrate of non-consciousness. It keeps banging out of nothingness.<sup>597</sup>

Let me assume that part of this statement is correct, namely insofar as there are indeed great ruptures in our experience during our wake life. I look at a cup, then turn my head away because someone called me, then attend the cup again. My perception of the cup is interrupted. Or take a longer and more extended thought, for instance, when I try to solve a mathematical equation. It might occur that my thinking is interrupted by a "thunder-clap"<sup>598</sup>. Distracted by the thunder-clap I might jump off my armchair and run to close the window. During this short period, I am not at all concerned about the mathematical task. The thunder is over for now, a strong rain sets in, but the windows are closed and I can put myself back to work. I sit at the desk and all my thoughts I had about the mathematical equation are almost blown away, so I have to start anew. Clearly, there were severe interruptions in my thought process. Now sitting at the desk again, it seems like I start from nothing, especially if my focus goes back to work, for just having closed the window is not relevant to the mathematical task.

But were there really interruptions in the experiential *process* in the sense that for a short moment there was no consciousness at all? Or did the interruptions rather concern the experiential *content*? I think at least in the example given it is clear that the breaks were in the content and that the experiential process qua process did not undergo any interruption. Though I was unable to finish my mathematical thinking, it was not disrupted by nothingness but another experiential content—the noise and my concern to prevent it by closing the window.

My contention is that this holds true even in cases where matters are less clear. Take, for instance, a meditative-kind of state in which I try to focus on the cup in front of me on the table. I intend to drop any other thought, emotion, recollection or imagination, and only attend to the cup in its concreteness. Thus, my attempt is to fix the experiential content, which will certainly require a significant attentional effort. If you are like me, then you will find this is a task that involves some sort of *flickering*. I am not an expert in meditation and my thoughts can be distracted by feelings or mental imagery that pops up.

---

<sup>597</sup> Strawson 2017, 35.

<sup>598</sup> James 1950/1890, 240.

Perhaps, while focusing on the cup, suddenly I recall something from the news that had made me sad or I remember that I still have to get *x* done today. Whenever I notice that I get distracted, I make even more attentional effort to stare at the cup and only the cup. Yet, I cannot really hold this tension for long. There are certain phases of relaxation, where I drift away, even though I do *not* think of something else than the cup—at least not in a pronounced way. For instance, there are small shifts in posture, and for a very brief moment I am eager to relax my left leg and relocate the center of muscular tension to my right leg. All changes of such a kind concern, I submit, experiential content in the sense that they are changes *in* experience rather than oscillations between experience and non-experience. Accordingly, I agree with McGilvary who criticizes James’s notion of succeeding pulses of thought, reminding the latter of his insights concerning the stream of consciousness: “The stream of thought is the fact, the pulses of thought are not facts; *the only pulses are in the objects of thought.*”<sup>599</sup> I take it, this applies equally to Strawson’s position.

But perhaps Strawson has an even more extreme phenomenon in mind, when he emphasizes that consciousness seems to ‘bang out of nothingness’. Could it be that I oversee that the short transitions between  $t_0$  and  $t_1$  at which content *c* is given are short periods in which my mind fully sinks off into experiential nothingness? Strawson seems to maintain that this is precisely the case and that we are even aware of these dark spots in experience. But the point is that even his descriptions of brief episodes of experiential nothingness between  $t_0$  and  $t_1$  remain phenomenological. That is, the short breaks, if they ought to be manifest in experience, are themselves experiential and so cannot present an example of experiential nothingness in the strict sense. I don’t see any other possibility than to describe even the experience as if one were constantly coming back from a moment of non-consciousness as a phenomenon proper that involves the experience of continuity. For that is what Strawson is ultimately saying, namely that *in* the phenomenology of experience there are gaps. For what he contends is that, before *c* was given to me at  $t_1$ , it was given to me at  $t_0$ , and in-between there was a form of experiential darkness. Such an experiential darkness, I suggest, is best understood in terms of a shift in the richness of content. In one moment, I attend the cup, for a short other moment, my mind blanks. This blankness, however, is not an absence of experience altogether. To conclude, then, I repeat James’s statement that wants to express the same idea, lending support to the notion of the stream of consciousness:

The transition between the thought of one object and the *thought* of another is no more a break in the thought than a joint in a bamboo is a break in the wood. It is a part of the *consciousness* as much as the joint is part of the *bamboo*.<sup>600</sup>

---

<sup>599</sup> McGilvary 1907, 226.

<sup>600</sup> Strawson 2017, 240

### 5.5.3 Frictions in Strawson's phenomenological proposal and metaphysical ramifications

Finally, is the phenomenological description of Strawson plausible? Although I can see the temptation and motivation to describe the fluctuations of our experiential life in terms of a series of eruptions of consciousness, ultimately, I don't think that such an interpretation goes through. First, I don't see that these fluctuations really present us with moments of complete absence of experience. Second, even if there were real breaches in the stream of consciousness (though it is unclear how they could be detected), they would still not undermine the experience of continuity, given that the blankness of the absence of any experiential content would still remain an experiential phenomenon. Third, Strawson's emphasis on the existence of both *true* experiential continuity and *apparent* or *felt* experiential continuity leads to inconsistencies. For, if he takes the most basic ground of consciousness to be fundamentally gappy, it becomes unclear what exactly true experiential continuity ought to be. Or, to consider the same issue from a different angle, "if our streams of consciousness really did have this [i.e., gappy, P.S.] structure, they would not even *seem* to possess the continuity they actually do"<sup>601</sup>. Strawson's differentiated view, which accepts

- (a) a basic gappy structure of consciousness,
- (b) short phases of true experiential continuity, and
- (c) apparent experiential continuity,

seems too ambitious in that it tries to acknowledge different claims about the phenomenology of experience that, in sum, are not compatible with each other. Of the triad of claims, one has to go in order to allow for a consistent approach. Strawson would have to get rid of (c), for instance, and claim that our experience consists of short periods of real continuity divided by breaches of the stream. But without (c), he couldn't acknowledge the experience of long-term continuity that Diachronics undergo any longer. Or, he could get rid of (b), and simply deny that there is any real experience of continuity (like the skeptics in Chapter 2). But that would undermine his claim about *sesmets* that rest on the idea that each *sesmet* is constituted by any experiential period that is truly continuous. At least, as his conception stands, he cannot drop the idea of real experiential continuity altogether. Obviously, dropping (a), for Strawson, is even less an option, given that the assumption of recurring gaps in experience is the characteristic claim of his approach.

Ultimately, I cannot see how the frictions between these three claims could be dissolved, and therefore suggest an alternative description and interpretation of the phenomenological observations that underlie Strawson's claims—and which admittedly

---

<sup>601</sup> Dainton 2008, 148.

bear a kernel of truth. Instead of ascribing gaps to the experiential *process*, I suggest to conceive of the apparent breaks in the stream as eruptions of different experiential *content*. Hence, I would drop claim (a). I suggest to keep claim (b), which I would specify by describing true experiential continuity in terms of Husserlian inner time-consciousness, i.e., the temporal field with the three moments of retention-impression-protention. Claim (c) is trickier. I think we need to distinguish a few phenomena. One is recollection, for instance, when I remember how I spent the summer last year. I take myself to *x* back then. This is phenomenologically distinct to *just* having *x*, which is still retained in consciousness. Yet, this doesn't mean that recollecting how I *x* last summer is just an appearance. It is, admittedly, less certain that I really *x* last summer, but, obviously, this doesn't mean I certainly did *not x*. All it means is that recollection is fallible. Accordingly, I suggest that when we speak of *apparent* experiential continuity, we can only mean that someone is undergoing an act of recollection that is incorrect in the same manner as one can be under the impression that one has met a certain person *y*, but it turns out one has been mistaken. From this, I emphasize, we need to distinguish the case of Diachronics. Unlike Strawson, I don't think we are dealing with a case of *apparent* experiential continuity. If I feel identical and continuous with my remembered past teenage self and his dreams and desires, recollect how they have been disappointed over the course of time, and feel how they are still present in a transformed and matured way, then there is little phenomenological motivation to denote such a sense of identity as a mere appearance. Or, to put it differently, the question is what more would be needed to acknowledge such a felt sense of identity is a true experience of continuity. To measure my affectively-laden recollection of my teenage self against the directness with which I retain experiences that have just passed seems misleading. For, if my teenage experiences were recollected with the same kind of immediacy like that what I have just experienced, it should be quite obvious that this would render the phenomenology of recollection a completely different one. There is—and, necessarily so—a phenomenological difference between the retention of an experience *e* at  $t_{-1}$  and the recollection of *e* at  $t_{-12}$  or  $t_{-100}$ . However, this difference is not one between *true* and *apparent*, but one between *direct* and *indirect* experiential continuity. I am not directly experientially continuous with my teenage experience, for in-between lie many years of experiential life. Remembering something that happened long ago unfolds in a different phenomenological manner than remembering something (even the same event) that happened recently.

All in all, then, I think Strawson's preference for the depiction of experience as a fountain of consciousness, at least as it currently stands, is confronted with some difficulties. That doesn't mean the intuition behind the idea of a fountain of conscious acts needs to be dismissed in toto. The phenomenologist Stephan Strasser, for instance, also employs the metaphor of a water fountain. However, his aim is to characterize the

dynamical structure of a basic stratum of non-intentional mood (*Stimmung*) at the center of experience out of which the numerically distinct intentional acts spurt out. The idea is that mood corresponds to the pool in which the water is located that presents the mass out of which the single water jets take their departure and find their end. Any punctual intentional act sinks back into the experiential process out of which it emerged and determines the affective quality of the on-going experiential process.<sup>602</sup> I have a lot of sympathy for this alternative metaphor. For, it accommodates both intuitions on the table, namely that our experience is of a stream-like continuity *and* that there is a myriad of singular and discrete conscious acts that spout out like a fountain.<sup>603</sup> In this picture, though there are discrete phases of pronounced consciousness of *x*, *y*, or *z*, there is no doubt that there is the one experiential process that unfolds continuously while *x*, *y*, and *z* make their entry on the stage. Such a picture, of course, however, does not any longer buttress Strawson's metaphysical view of the self as a *sesmet*, which I conclude doesn't find support by the phenomenology of our experience.

## 5.6 Bayne: Virtual phenomenalism

I will now turn to a last alternative experiential proposal, which I have mentioned already in Chapter 1: Tim Bayne's *virtual phenomenalism*. As I have argued there, although Bayne's approach partially has the appearance of a no-self doctrine, there are reasons to rubricate it among the experiential approaches to selfhood. His proposal, thus, is ambivalent. In this section, I want to show in which sense virtual phenomenalism qualifies as an experiential approach to selfhood, how it—despite some resemblance—differs from self-experientialism, and discuss whether it provides us with a better response to the issue of personal identity than self-experientialism might. It shouldn't come as a surprise that I will finally argue that virtual phenomenalism is *not* to be preferred over self-experientialism.

### 5.6.1 Virtual phenomenalism is an experiential approach to selfhood

First, let me explain the ambivalence intrinsic to Bayne's proposal and why I think we should interpret it as an experiential account rather than a no-self view.

---

<sup>602</sup> Cf. Strasser 1956, 114.

<sup>603</sup> Moreover, Strasser's emphasis on mood or feeling as a central aspect of the experiential fundament of all intentional experience is quite compatible with the view that feeling plays a crucial role in the most basic form of self-experience, which I discuss in Part III of the dissertation.

(a) On the *one* hand, Bayne has the tendency to declare selves as “the central characters in a kind of phenomenal fiction”<sup>604</sup>, the suggestion to “think of the self as a merely virtual centre of ‘phenomenal gravity’”<sup>605</sup>, which amounts to the claim that “the self is a ‘merely intentional entity’”<sup>606</sup>. As he takes it, we experience ourselves as being at the center of the phenomenal field which remains the same. Such a center, notably, is not something we only witness: “we are not merely *tracking* our selves but *creating* them”<sup>607</sup>. That makes it look like Bayne believes that selfhood is something fictitious, something *made up* by consciousness, more specifically, a by-product of the specific human way of experiencing the world, namely as an *objective* world. For, as he also emphasizes, he is “not convinced that *de se* representations must structure the experiences of any possible subject of experience”<sup>608</sup>.

(b) On the *other* hand, Bayne warns us that one could misinterpret his virtual phenomenalism as a version of the no-self doctrine:

[N]ot only is the denial of the self of dubious coherence—it is certainly not how virtual phenomenalism was advertised. I said I was going to provide an account of the self, not an account that explained the self *away*.<sup>609</sup>

To explain what he means by claiming that the self is both part of a phenomenal fiction and nonetheless real, he suggests a comparison with Hercule Poirot, a character of the literature created by Agatha Christie. Poirot, a fictitious Belgian detective, is real only in Christie’s works. “[H]is mode of existence is to be contrasted with that of real (not-fictional, actual, existent) Belgian detectives.”<sup>610</sup> Although, given the way he is described in the literature, he could show up on the list of real Belgian detectives, but he doesn’t—he is only a literary figure. The case of the self is different, as Bayne emphasizes:

[T]here is no kind of real self with which our kinds of selves could be contrasted, for it is in the very nature of selves to be virtual.<sup>611</sup>

The point, thus, is while in the case of Belgian detectives one can compare those who are only fictitious with those who are real in terms of existing among *us*, the human persons of this world. By contrast, in the case of the self no such comparison is possible. For, over and

---

<sup>604</sup> Bayne 2010, 269.

<sup>605</sup> Bayne 2010, 289.

<sup>606</sup> Bayne 2010, 292.

<sup>607</sup> Bayne 2010, 291.

<sup>608</sup> Bayne 2010, 290.

<sup>609</sup> Bayne 2010, 292f.

<sup>610</sup> Bayne 2010, 293.

<sup>611</sup> Bayne 2010, 293.

above the self as a virtual and intentional entity, there is no self that would be more real than a virtual self. Bayne's diagnosis is that whenever one were inclined to look out for "something in consciousness or its underlying substrate that might qualify as its referent"<sup>612</sup> must necessarily be disappointed. He also emphasizes that such a search is not necessary, "for it needs no such referent"<sup>613</sup>. In fact, he reassures that "[t]he kinds of selves that we possess are as real as selves get", the merely virtual reality or intentional being of selfhood "provides all the reality that we *need*"<sup>614</sup>. Even more, he also highlights that "perhaps the only thing that could play the role of the self—is a merely intentional entity"<sup>615</sup>.

The ambivalence between Bayne's description of the self as virtual and the reality of self he *does* believe in can then be dissolved in the following manner: By pointing to the self as a virtual object or merely intentional being, Bayne does not want to say selfhood is unreal, but rather wants to determine the self's *way* of being or *mode* of existence. And, when denoting the self as a virtual and intentional entity, what Bayne is ultimately stressing is that the mode of existence of self is experiential. According to virtual phenomenalism, the experiential reality of self just *is* the reality of self.

### 5.6.2 Virtual phenomenalism can address the issue of personal identity

Defining self in terms of an intentional entity, Bayne argues, has several advantages. The most important one he discusses concerns personal identity and the challenge of phenomenal gaps. He suggests that

virtual phenomenalism can handle the challenge posed by phenomenal gaps, for the virtual self ... is more than capable of reaching across interruptions of phenomenal continuity<sup>616</sup>.

To explain how such a bridging can be done, he engages again in a comparison with Hercule Poirot and literary fiction. The personal identity of Hercule Poirot, he takes, is constituted by the author along those "conventions that their readers pick up on"<sup>617</sup>. The author looks at Poirot through the lens of the readers and develops his character along those features that make readers take another story about a man to be a story about Hercule Poirot. The fact that the story and identity is constituted across discrete (and therefore interrupted) works of narration does not in any way hinder the unity of the (fictive) person.

---

<sup>612</sup> Bayne 2010, 293.

<sup>613</sup> Bayne 2010, 293.

<sup>614</sup> Bayne 2010, 293.

<sup>615</sup> Bayne 2010, 290.

<sup>616</sup> Bayne 2010, 290.

<sup>617</sup> Bayne 2010, 291.

On Bayne's view, this can be applied to the case of self in a similar manner. However, he qualifies, in the case of self, "'conventions' are not dependent on social practice—or at least not primarily so"<sup>618</sup>. The conventions that apply in the case of self are rather experiential: "We simply appropriate to ourselves those experiences that we are aware of 'from the inside'."<sup>619</sup> Put differently, the constitution of what is typically referred to as *narrative identity*<sup>620</sup> or of the identity that spans across interruptions of the experiential life is primarily—in agreement with Locke—based on first-personal experience:

We assume that the character at the centre of our current stream of consciousness is one and the same as the character at the centre of those streams of consciousness with which we are in direct contact via autobiographical memory.<sup>621</sup>

Virtual phenomenalism, thus, does not rest on the notion of an uninterrupted stream of consciousness nor does it take the unification of self across time (including phenomenal gaps) to be based on a unifying substrate of the stream of consciousness (as Dainton's substrate phenomenalism). Rather, virtual phenomenalism proposes to conceive of the self as an intentional object that is *produced by reflection, based on available first-person experience* of recollection (remembering experiences as experienced *by me*), and structured through "*narrative scaffolding*"<sup>622</sup>. Hence, the unification of self that seemed to be challenged by interruptions in the experiential life is provided by a kind of narration that is based on the story of experiential life. It is in *this* sense, to use Bayne's wording once more, that we qua selves are "the central characters in a kind of phenomenal fiction"<sup>623</sup>. Phenomenal *fiction*, however, as the comparison with Hercule Poirot showed, is not *mere* fiction. The experiential narration, instead of being just *made up by* consciousness is rather *made of* consciousness. The self, according to virtual phenomenalism, turns out to be the author, reader, and protagonist of the phenomenal fiction, i.e., of that what is made of consciousness or that what is essentially present in the mode of consciousness.

### 5.6.3 Self-experientialism is not the same

Now, how does virtual phenomenalism relate to self-experientialism, and is it superior to it? First, I wish to emphasize that Bayne's virtual phenomenalism bears high resemblance

---

<sup>618</sup> Bayne 2010, 291.

<sup>619</sup> Bayne 2010, 291.

<sup>620</sup> For a brief explanation of the concept, see Ricoeur 1991.

<sup>621</sup> Bayne 2010, 291.

<sup>622</sup> Bayne 2010, 291 (my emphasis).

<sup>623</sup> Bayne 2010, 269.



with self-experientialism insofar as it also underscores the essential experiential *mode of existence* of selfhood. Bayne agrees with self-experientialism that the stream of consciousness should not be understood as some kind of “cord”<sup>624</sup> that wouldn’t allow for any kind of interruptions. The continuity and identity of the experiential process doesn’t hinge on that it is uninterrupted, for instance by periods of dreamless sleep. As a result, virtual phenomenism also highlights that there is no problem with alleged phenomenal gaps, as these are, as a matter of fact, already bridged whenever we remember something that lies beyond the last alleged phenomenal gap, for instance, yesterday’s birthday dinner before I went to sleep. Virtual phenomenism, thus, also shares with self-experientialism the notion that phenomenal continuity consists in the *experience of continuity* rather than in the (uninterrupted) *continuity of experience*.

However, there are also significant differences. These come best to light if one brings in another—classical—account that also shows great overlap with virtual phenomenism: the position by Jean-Paul Sartre that he defends in *The Transcendence of the Ego* (1957). Characteristic for Sartre’s position is that he takes

- (1) consciousness to be *non-egological*,
- (2) the ego to be an *intentional object*,
- (3) the ego to be a *product of reflection*.

These three features of his account are well-reflected in the following quotation:

[T]he ego is an object apprehended, but also an object *constituted*, by reflective consciousness. The ego is a virtual locus of unity, and consciousness constitutes it in *a direction contrary* to that actually taken by the production: *really*, consciousnesses are first; through these are constituted states; and then, through the latter, the ego is constituted.<sup>625</sup>

Evidently, these features render Bayne’s virtual phenomenism a direct opponent to phenomenological self-experientialism that is oriented, and, as I argued, implied in Zahavi’s minimal self-view. For, Zahavi’s egological theory of self-consciousness is precisely directed against the Sartrean non-egological understanding of consciousness and self-consciousness. On Zahavi’s view, and as I have already argued in Chapter 4, for reflection to be possible it is required that the consciousness that reflects on itself must have been already aware of itself and that such a pre-reflective self-awareness is egological. Furthermore, I have tried to show that the primary way in which we are conscious of ourselves consists in the self-consciousness of the experiential flow, i.e., *not* in the consciousness of an object. Instead of repeating these discussions here, I rather wish to

---

<sup>624</sup> Zahavi 2014a, 72.

<sup>625</sup> Sartre 1991/1936, 80f.

focus on another question: Does the fact that Bayne, in agreement with Sartre, describes the self as an intentional object provide any advantages over self-experientialism (which stresses that the being of selfhood does not hinge on an objective self-experience of the self)?

#### 5.6.4 The advantages of virtual phenomenalism presuppose self-experientialism

To better understand this question, we must take a look at another component of Bayne's virtual phenomenalism that highly overlaps with another group of accounts prominent in the last decades of the philosophy of self: *narrativism about selfhood and identity*.<sup>626</sup> Although the narrativity thesis can be cashed out in quite different ways, the upshot is that, on this view, the identity of a person is constituted through the narrative someone acquires with regard to herself. Bayne also subscribes to such a view by emphasizing the comparison with Hercule Poirot and the production of a literary figure.

But his claim goes further. For him, it is not only the *identity* of a person that is constituted through narrative. It is *selfhood* to begin with that is owed to the process of narration that is intrinsic to the specific human-kind of experiencing the world.<sup>627</sup> The self, on Bayne's view, is only real insofar as it is constituted in experience as a virtual object—as well as is the person and her identity. Importantly, Bayne doesn't distinguish between self and person. The narrative generation of the self as a virtual object ought to occur in a similar way as the development of a person and her character, i.e., like Hercule Poirot in literary fiction. When reading a novel involving a character of literary fiction, we are used to accommodate gaps in the narration, not all and every moment in the life of the character are narrated, and yet, we find no difficulty in attributing different fragments of narration and narrated events to the one character a novel is about. Similarly, Bayne holds, self and person is constituted in the narration performed by a reflective act directed upon the stream of consciousness. Reflecting on our current experiences involving and extended by autobiographical memory, the self as a virtual object is generated. Just like in the case of literary fiction, the fact that the past stream of consciousness might have been interrupted by periods of dreamless sleep does not in any way undermine the possibility to constitute a virtual object around which present experiences (current perceptions in combination with recollections, anticipations and so on) are centered.

If selfhood is real merely as such a virtual object, only present in a reflective attitude towards one's own experiencing extended by autobiographical memory, then this seems to

---

<sup>626</sup> For instance, MacIntyre 2007/1981, Ricoeur 1988, 1992, Rudd 2009, 2012, Schechtman 1996, 2007, Taylor 1989.

<sup>627</sup> To recall, unlike Dennett 1992, Bayne doesn't want to say that the fact that the self is based on narrative processes would render it unreal or illusory of any sort.

fit only perfectly with a narrativistic understanding of selfhood: The self as a merely virtual object is construed on the basis of the narrative that unfolds depending on one's experiencing in combination with what is present as recollected through autobiographical memory.

Now, by contrast, self-experientialism, as I have highlighted, does not take selfhood to consist primarily in being an object of a certain reflective act. And, accordingly, self-experientialism cannot maintain, like Bayne, that it is only through a form of narration that a sort of object is constituted and which *is* the self. Does that mean that self-experientialism cannot make use of the narrative explanation for the constitution of personal identity? Does virtual phenomenism have the advantage over self-experientialism that it can account for personal identity, while self-experientialism cannot?

Let me be clear on this. Bayne's virtual phenomenism *does* present a convincing proposal as to how to accommodate phenomenal gaps. However, the kind of narrativistic understanding of selfhood and personal identity does not presuppose that one defines selfhood in terms of a virtual object that is constituted through reflection. It is also available to self-experientialism, as I will argue in a moment. In fact, such a closer examination, I further suggest, shows that self-experientialism is rather presupposed by virtual phenomenism. Hence, I submit, the benefit of virtual phenomenism, namely to be able to address the issue of personal identity, presupposes the kind of understanding of selfhood corresponding to self-experientialism. If this line of reasoning is correct, then virtual phenomenism cannot be said to be preferred over self-experientialism.

First, let me emphasize that a narrativistic understanding of personal identity is available to self-experientialism, too. For, although self-experientialism underlines that the being of selfhood does not hinge on there being an object of experience that corresponds to *the* self (a self-object), self-experientialism does not only allow for a self-objectification through different kinds of reflective acts, it is quite compatible with the view that many forms of self-experience necessarily involve some sort of self-objectification.<sup>628</sup> If I perceive someone else being angry, I experience that someone else as a form of self-object, an experiencing body, i.e., the other as being bodily (*leibhaftig*) there. And, as, for instance, Sartre himself famously argues it is precisely the gaze of the other that objectifies me: Not only do I become an object for the other, I experience myself as an object through the other.<sup>629</sup>

Accordingly, there is room in the conception of self-experientialism for a narrativistic understanding of personal identity. For, to recall, as I emphasized in Section 5.2, self-experientialism distinguishes between self and person. Thus, although self-

---

<sup>628</sup> Zahavi 1999, 142ff. Already a basic reflective consciousness implies a certain form of self-objectification insofar as a certain experience is taken as the object of a reflective act.

<sup>629</sup> Sartre 1956, 246ff., Zahavi 2014, 213.

experientialism may and does reject the view that it is through narrative that selfhood is constituted, it can acknowledge it is through narrative that the identity of self in the form of the person, who has a character and a life story, is constituted. As Zahavi suggests, both the experiential and the narrative notion of selfhood “are so different that they can easily complement each other”<sup>630</sup>. To emphasize that selfhood consists in self-experience is compatible with the idea that it is through narrative that we develop forms of self-experience that span over our lifetime and in which we experience ourselves as identical with our past self of a longer time ago.

Hence, it is not necessary to define selfhood in terms of a virtual object, as virtual phenomenism does, in order to enjoy the advantages of a narrativistic understanding of personal identity. Therefore, pointing to phenomenal gaps and the bridging problem alone doesn’t suffice to motivate the view that the self is a virtual object which is only real whenever a narrative of one’s stream of consciousness is formed based on reflection.

In fact, one might even argue that both understandings—self-experientialism and a narrativistic understanding of personal identity—are not independent from each other. As Zahavi claims, “it should be clear that the notion of self introduced by the narrative model is not only far more complex than, but also phenomenologically and ontologically dependent, on the experiential self”<sup>631</sup>. As he explains, “only a being with a first-person perspective could consider her own aims, ideals, and aspirations *as* her own and tell a story about them”<sup>632</sup>, “to begin a self-narrative, the narrator must be able to differentiate between self and nonself” and must be “in possession of a first-person perspective”<sup>633</sup>.

I think such a line of reasoning is right. It is hard to see how a narration based on reflection on one’s experiences could possibly bring about the experience of being a self in the first place. The same applies to Hercule Poirot and literary fiction. It doesn’t take the many clever and witty descriptions of the occurrences by Agatha Christie to understand Poirot is *someone* and no descriptions are delivered to make the reader understand Poirot is an experiencing being who enjoys a first-person perspective. The reader already knows. She already knows Hercule Poirot—regardless of who he is and what his story is, which will become clear over the course of the narration—is an experiencing being, someone who is undergoing experiences and aware of himself. Put differently, to have all the advantages of a narrative account of selfhood, an understanding of selfhood in terms of self-experientialism is already presupposed. In fact, the way Bayne combines a Sartrean (self as an intentional object), Lockean (continuity through autobiographical memory), and narrative account (self as a virtual object comparable with a protagonist of a novel)

---

<sup>630</sup> Zahavi 2005, 129.

<sup>631</sup> Zahavi 2005, 129.

<sup>632</sup> Zahavi 2005, 129.

<sup>633</sup> Zahavi 2005, 114.

ultimately turns out to suggest an understanding of selfhood that skips to show how an experience of oneself as an intentional and virtual object is possible in the first place. Addressing that important question of the condition of possibility, furthermore, rather points us towards an understanding along the lines of self-experientialism—or so I am convinced.

## 5.7 Self-experientialism, identity, and ontology: the primacy of self

In the last three sections, I have presented and discussed three alternative experiential accounts of selfhood that place emphasis on the issues of continuity and personal identity. Central to all of them is a specific stance towards these issues that determines a corresponding understanding of selfhood. Dainton's proposal to bridge phenomenal gaps consists in defining selfhood in terms of the capacities for consciousness, thereby avoiding the problem of interruptions in the stream of consciousness; Strawson suggests we should accept that there is no personal identity in the narrow sense to begin with (at best there is only a fallacious *sense* of identity) and that each of the many interruptions of our stream of consciousness brings forward a new subject of experience that is ontologically distinct to the subject(s) of experience existing prior to the phenomenal gap; Bayne offers the view that the self as a virtual object is constituted by the awareness of different experiences (involving current perceptions and thoughts, recollections, and anticipations of the future), wherefore it is not required that these different experiences succeed each other without any interruption.

Thus, all three approaches implicitly acknowledge the assumption that a notion of selfhood must involve a solution to the issue of personal identity. That is, they take it that personal identity is an explanandum that any notion of selfhood must have the potential to address and account for. So far, I have discussed the *specifics* of each of the alternative experiential approaches to selfhood, in the remainder of this chapter, I want to discuss that more *general* assumption that seems to underlie all of them. The question, thus, is when building our concept of self and personal identity, which of the two should be prioritized?<sup>634</sup> Should the notion of personal identity constrain the notion of selfhood? Or is it rather the notion of selfhood that should constrain the notion of personal identity?

---

<sup>634</sup> Evans emphasizes the distinction between what he calls the “person-approach” and the “self-approach”. The former addresses the question what a person is and under which circumstances two selves can be said to be identical. The latter, by contrast, focuses on the question what a self is. Throughout his *The Subject of Consciousness* (1970), he emphasizes—correctly in my view—that both question must not be conflated. Cf. Evans 1970, 19–26, et passim.

I trust everybody will agree that, ideally, no notion of either selfhood or personal identity should engender any absurdities with regard to the other notion. As Evans puts it pointedly: “[A]n answer to the question ‘What is the self?’ must not contradict the answer to the question ‘What is a person?’”<sup>635</sup> With that I agree as well as with the following statement: “Furthermore, an answer to the first question must not be thought to supply the answer to the second question and vice versa.”<sup>636</sup> That said, I do believe that absurdities that result from a notion of personal identity for our understanding of selfhood weigh heavier than the other way around.<sup>637</sup> Accordingly, I suggest we should prioritize the notion of selfhood over the notion of personal identity. Before we determine, as I have proposed in the last section, wherein the identity of *someone* lies, we should have determined what *someone* is. Before we can ask under which conditions a self can be said to be identical with a past or future self, we need to understand what self is. It is selfhood that dictates under which conditions a self can possibly be identical with another self.

As I have argued, all alternative experiential approaches to selfhood imply each specific problems with regard to the notion of selfhood. Dainton’s proposal has the consequence that selfhood is real as long as the capacity for consciousness is retained, rendering it possible that there are selves that ought to exist without ever actually undergoing experiences, which I deemed absurd. Strawson’s proposal, I have argued, doesn’t match with the phenomenology of our experience and is confronted with other bridging problems: the unification of conscious content across time. Finally, Bayne’s proposal either fails in endorsing a Sartrean non-egological theory of self-awareness (and inherits its problems discussed in Chapter 4) or at best must acknowledge that it presupposes a notion of self similar to self-experientialism.

But what exactly are the ramifications of self-experientialism and its notion of selfhood for the notion of personal identity? Do absurd consequences arise for the notion of personal identity if we understand selfhood in the way self-experientialism does? There is one general challenge that is usually attributed to an experiential account like self-

---

<sup>635</sup> Evans 1970, 21.

<sup>636</sup> Evans 1970, 21.

<sup>637</sup> Dainton explicitly opts for the opposite view: “Faced with a conflict between some general ontological preconceptions (prejudices), and a compelling account of my existence and persistence conditions, I would be inclined to allow the latter to override the former.” (2008, 342) Yet, the issue of personal identity is such a controversial subject that we cannot even be certain it is a philosophical problem that is well-formulated. As Hume famously concludes, “all the nice and subtle questions concerning personal identity can never possibly be decided, and are to be regarded rather as grammatical than as philosophical difficulties” (2007/1739, 171 (1.4.6.21)). Perhaps such a verdict goes too far in that it oversees that questions of personal identity do indeed affect much of our practical life and its understanding, as, for instance, Parfit (1971, 1987) no less famously has worked out. Hence, one should acknowledge that the issue of personal identity is a philosophical problem. But it is a problem that is primarily ‘grammatical’ in the sense that it is not even clear how to best formulate it. By contrast, the *question* what we are and what selfhood is is much easier to formulate and less controversial. Moreover, importantly, any formulation of the problem of personal identity involves a notion of selfhood already. For, with selfhood undetermined, it is quite unclear the identity of what exactly one is interested in when inquiring the nature of *personal* identity.

experientialism, which is primarily directed at a Cartesian approach. However, as I want to argue now, given the differences between self-experientialism and Cartesianism, it doesn't apply—or at least not in the same way—to self-experientialism. Let me be more specific.

The challenge concerns the *ontological status of self* during dreamless sleep. Given how Descartes understood selfhood, namely as *res cogitans*, i.e., as a substance that can be separated from the body but not from *cogitare* or thinking (which is its nature), interruptions of consciousness (= *cogitare* = thinking) posed the problem that during these phases *ex hypothesi* the ego doesn't exist. However, if the ego ought to be a substance, this is unacceptable. For a substance, according to the traditional understanding, cannot exist, then cease to exist, only to come back into existence when a phase of dreamless sleep has ended. Is self-experientialism susceptible to the same problem?

Not exactly. Bearing in mind that self-experientialism doesn't subscribe to a substantial understanding of selfhood, there is no conceptual problem with the fact that selfhood may be active during wakeful and inactive during dreamless sleep. Neither is it a problem from the first-person perspective, as Bayne's virtual phenomenalism emphasized, to bridge phases of non-consciousness. Structurally speaking, they pose no problem whatsoever. As Zahavi writes:

From the first-person perspective, it is not as if one has to reach back and establish a connection to a separate stream of consciousness.

...

There are no extended periods of unconsciousness, and linking up with an experience you had yesterday, say, an acute experience of shame or embarrassment, is no different from linking up with an experience you had earlier this morning. In both cases, we are faced with a diachronically unified consciousness.<sup>638</sup>

The point is that the structure of our recollection is such that when, at a given point in time  $t_0$ , we remember an experience at  $t_x$ , it matters not whether I also presentify and recollect the experiences that I underwent between  $t_x$  and  $t_0$ . And structurally speaking, it makes absolutely no difference whether between  $t_x$  and  $t_0$  I had been in a completely dreamless coma or lived through a rich experiential life. In fact, I never have all my experiences between the present and a remembered moment in sight and I do not need to in order to be able to remember a past moment.

Even so, some might insist, while there might be no problem from the first-person perspective, things look differently from a third-person perspective. If I see a person in a coma and the doctors assure me she is absolutely non-conscious but might wake up at some point, what exactly is the ontological status of self in this particular moment? From the first-person perspective there is no selfhood during the phase of non-consciousness. But

---

<sup>638</sup> Zahavi 2014, 73.

from the third-person perspective, wouldn't I still take that person to be a being that involves selfhood, even though dormant? And if so, doesn't this point to the fact that self-experience isn't essential to the being of selfhood as self-experientialism has it? There I see the person in the coma, she is bodily present, I see the machine for artificial nutrition, her body is warm and seems to be functioning. Yet, no signs of her being sentient in any way.

I would agree that, in such a situation, one still rightly takes the person to be real and existing even though during the phase of non-consciousness selfhood is inactive. But I wouldn't take this to be an argument against self-experientialism or a reason to define selfhood in terms of the bodily organism only.<sup>639</sup> It is true that during dreamless sleep or coma, the capacity for consciousness is retained by the functioning organism and so the possibility for selfhood is preserved. But one shouldn't overlook that when approaching the person in coma, although she is bodily present to me, something essential is missing: experience and self-experience. When the person in coma is completely non-conscious, she is not anything to herself, and so lacks selfhood.

Now, defining selfhood, contrary to self-experientialism, in terms of the bodily organism only would have the advantage that the puzzlement is reduced, which adheres to the notion that during non-conscious phases someone ought to cease to present a case of selfhood. However, defining selfhood in terms of the bodily organism only would have the problematic consequence, as described earlier, that the notion of selfhood would then equally apply to philosophical zombies that lack any phenomenal consciousness. I find that repercussion more worrisome. It is one thing to say that experiential selfhood is intrinsically bodily and involves a functioning organism, which I find plausible and compatible with self-experientialism. And, it is another thing to say that, really, the bodily organism *is* the self, which would once again objectify selfhood in a manner that self-experientialism aims to avoid, given that our most basic self-experience should be conceived of in non-objective terms.

Even though it may seem strange that during non-conscious phases, the self ought to be unreal only to become real whenever consciousness is turned on like the lamp on my table, I find that view still matches better with our experience and intuitions than the view that a functioning organism lacking any phenomenal consciousness whatsoever ought to be as much a self as a phenomenally conscious one.

---

<sup>639</sup> Olson (1998), for instance, suggests we should get rid of the notion of selfhood altogether and come to acknowledge that we are essentially organisms. See also Olson 2007, 2008. A less radical approach is defended by Fuchs (2017b) who endorses the view that there is a continuity between life (*Leben*) and mind/experience (*Erleben*). On his view, selfhood consists both in the self-related, living organism and experiential selfhood, which cannot be strictly separated. Moreover, Fuchs suggests it is the bodily organism that provides personal identity across time, even during sleep, which he doesn't consider to be fully non-conscious. Schlicht (2018) also connects the experiential self with the organism. I am sympathetic to such approaches. However, I would stress that a non-experiencing but functioning organism lacking any phenomenal consciousness is not a self.



To conclude, then, I don't think that Gassendi's objection against Descartes can be applied to self-experientialism. It is true that it is puzzling that selfhood can be inactive during non-conscious states without being said to have ceased to exist altogether, given that the possibility for consciousness is retained. Yet, I would insist that we better not conflate that what is a condition for selfhood with selfhood itself. Though a bodily organism might presumably be a necessary condition for experience and selfhood, without experience, the bodily organism lacks selfhood. For, even if a non-conscious, living organism may be said to distinguish between itself and its environment, for instance, by distinguishing between certain substances as being nutritive and benefitting a certain metabolic equilibrium and substances that need to be digested and excreted, this distinction is not one between a *de se-character* and *de re-character*. Put differently, autopoietic self-regulatory activity—often ascribed even to non-conscious organisms—might be described in terms of reflexivity, but the reflexivity qua reflexivity can only be described from an experiential point of view. It is for us, experiencing beings who make an experience of reflexivity in the self-awareness of our experience, that we can recognize the reflexive behavior *as reflexive* (i.e., self-related) found in non-conscious living organisms. But for the non-conscious, living organism reflexivity means nothing, even if it can distinguish between  $p$  and  $q$ , whereby  $p$  corresponds to the organism itself. That is, for the non-conscious organism the distinction is simply between  $p$  and  $q$ , even though  $p$  happens to be the organism itself. In this sense, I think it is accurate to say a non-conscious, living organism lacks selfhood. Reflexivity and selfhood, I contend in accordance with self-experientialism, are best understood as experiential notions.

## 5.8 Conclusion: Self before person

My aim in this chapter was, on the one hand, to further specify phenomenological self-experientialism by distinguishing it from Cartesianism and Lockeanism, and, on the other hand, to distinguish it from other prominent experiential approaches to selfhood that have been recently offered. A great deal of my attempts was negative in that I wanted to show that these alternatives each face several problems and therefore are *not* to be preferred over self-experientialism. Moreover, given that most of these alternatives motivated their claims against the background of the issues of personal identity, I addressed the question whether we should prioritize the notion of personal identity when it comes to the task of determining what selfhood is, which seems to be a silent assumption in the alternative experiential approaches to selfhood different from self-experientialism. Two claims resulted from my discussion. First, the more specific claim that the issue of personal identity, and, particularly, the bridging problem, alone are not sufficient to make a case for the alternative

experiential approaches to selfhood, as their solutions to the bridging problem engender severe complications with regard to the notion of selfhood. Second, and more generally, I argued that the notion of selfhood has priority over the issue of personal identity. On this view, it is only after we have determined what selfhood is that we can approach the task of determining wherein personal identity lies. That is, the notion of selfhood constrains personal identity, as I tried to emphasize by offering a modest interpretation of Locke's account of personal identity.

It should be noted that, beyond these claims, my own analysis of personal identity is admittedly very limited. Although I have defended the view that typical objections against Cartesianism and Lockeanism do not apply to phenomenological self-experientialism, given the relevant differences of the latter from the former, I haven't provided a comprehensive theory of personal identity myself. Doing so would require a thorough investigation, one which I cannot deliver here within the confines of this work. My modest goal here with regard to the issue of personal identity has only been to reject the view that one could launch a notion of selfhood based on a certain understanding of personal identity.

What further conclusions can be drawn from this? I believe of the presented accounts phenomenological self-experientialism is to be preferred. It presents an alternative to the three skeptical views about self-experience and selfhood as well as to the Cartesian and Lockean proposals. Since it is also less problematic than other recent experiential approaches to selfhood, I believe it should be the standard account to be beat by any new account or alternative theory. That said, evidently, there are other challenges to phenomenological self-experientialism that will need attention in future research. In the following, last part of the dissertation, I want to focus on one of the possible challenges that have been formulated against the minimal self-view and which, by implication, also apply to self-experientialism. The objection consists in the notion that the mere experienced reflexivity intrinsic to experience, i.e., pre-reflective self-awareness, is too formal a feature to present a convincing notion of selfhood, particularly if we measure it against the richness of the human self and its multidimensional nature.

## Part III



## Chapter 6

# Minimal Self and the Feeling of Being Oneself

The peculiarity of Feeling, therefore, is that there is nothing but what is subjectively subjective; there is no object different from self—no objectification of any mode of self.<sup>640</sup>

If we, without being distracted by school doxa, ask our immediate consciousness itself, there can be *no doubt that our true, inner being and life is constituted through the feelings*, in which we feel our state and its meaning immediately, and through the strivings, through which we effectively determine ourselves and give ourselves the directions from one moment to the other. This is the true core of our existence, as it comes to consciousness for us.<sup>641</sup>

The point at which the reality of the I becomes the most palpable, are the emotional states, the simple feelings, moods, and the affects.<sup>642</sup>

The preceding chapters introduced phenomenological self-experientialism as an ontological commitment of the minimal self-view and reconstructed the theoretical motivations for attaining it. Moreover, they compared it with recent alternative experiential approaches to selfhood, resulting in the claim that self-experientialism is to be preferred over existing competing theories. However, in the debate, there exist also other criticisms of the minimal self-view that impact on phenomenological self-experientialism. These criticisms concern the notion of selfhood implied in the minimal self-view and its rather thin character: Is the mere experience of reflexivity involved in all experience really enough to constitute selfhood? For some, this question must be answered negatively. Therefore, they suggest to enrich the concept of self with further aspects that renders the self less thin. My aim in this chapter is to show that we do not have to give up the idea of a thin subject, but that we need to account for an experience of being as part and parcel of the minimal self. That is, on the view I propose, only if the experience of reflexivity involves some

---

<sup>640</sup> Hamilton 1870, 432.

<sup>641</sup> Sigwart 1881, 226, my translation and emphasis: “Wenn wir, unbeirrt von Schulmeinungen, unser unmittelbares Bewußtsein selbst fragen, so kann darüber *kein Zweifel sein, daß unser eigentliches, inneres Sein und Leben durch die Gefühle konstituiert wird*, in denen wir unseren Zustand und seine Bedeutung unmittelbar empfinden, und durch die Strebungen, durch welche wir wirksam uns selbst bestimmen und uns die Richtungen von einem Moment zum anderen geben. Das ist der wahre Kern unserer Existenz, wie sie uns zum Bewußtsein kommt. Darin erscheinen wir uns als wirkliche, im Zusammenhang des Erleidens und Wirkens mit der übrigen Welt stehende Wesen.”

<sup>642</sup> Oesterreich 1910, 7, my translation: “Der Punkt, an dem das Vorhandensein des Ich sozusagen am handgreiflichsten wird, sind die emotionalen Zustände, die einfachen Gefühle, die Stimmungen und die Affekte.”

sense of being can it legitimately be called a sense of self in the first place. As some have put it: “[T]he property of *being the very subject that one is* enters essentially into the phenomenological character of all one’s experiences.”<sup>643</sup> My suggestion, in this chapter, is that such a sense of being is involved in the minimal self in terms of feeling. To formulate this claim, I introduce the notion of the feeling of being oneself, which I develop based on a discussion of Rudolf Hermann Lotze’s notion of self-feeling and Matthew Ratcliffe’s concept of existential feelings. Crucially, I will propose that the structure of the minimal self, as Zahavi describes it in terms of Husserl’s inner time-consciousness, incorporates a feeling of being oneself. Accordingly, my understanding of the minimal self doesn’t present a completely alternative approach but rather builds on Zahavi’s. The rationale behind the whole point of the chapter is: If we want to accept his minimal self-view as a convincing conception of selfhood, which is not too meagre in that it does involve a ‘who’ (i.e., an individual experiencer), then we need an account for how a sense of being is involved even in an experience that only features a minimal self. To equip the minimal self-view with such an account, we must show how feeling could be considered an intrinsic aspect of the minimal self. Why feeling? Because feeling is typically taken to convey a sense of being.<sup>644</sup> Let me try to illustrate the challenge of the sense of being by the following example:

Our everyday life is characterized by quite different ways of experiencing ourselves. When I take the car to visit my parents who live near the woods outside of the city and drive silently on the small highway that makes its way through serpentine, my self-experience, lost in thought, is a very specific one. It is different from the self-experience that I undergo when I engage in an intense conversation with my parents. For instance, *when* they question my life decisions or when I find them fighting about what I take to be negligible issues and try to mediate between them. Presumably, my self-experience is also different in situations such as engaging in sports activities with my friends, going to the theater with my partner, helping a passenger in the park who is being molested by strangers, or trying to accomplish the tasks my boss has put on my shoulders.

All these experiences and self-experiences, regardless of their heterogeneity otherwise, arguably involve an awareness of one’s own existence. Whether I am driving to my parents, talking to them, being at the theater or wandering through the park, I always experience myself as there, being, in existence. Likely, most will find it also plausible to

---

<sup>643</sup> Turausky 2014, 249.

<sup>644</sup> This was eminently the case in the post-Kantian German Romanticism, as Manfred Frank (2002, 36f., 77–92, 190–206) shows in his excellent study. Other examples include Rudolf Hermann Lotze and those he influenced, such as the proponents of “Munich phenomenology” (Salice 2020/2015, Smid 1982, Vendrell-Ferran 2008, 2015) or other philosophers in the German academic landscape like Sigwart 1881, Oesterreich 1910 or Ziegler 1893. Martin Heidegger prominently defended the view that our understanding of being essentially involves feeling, an approach that has recently been further developed by Matthew Ratcliffe under the label of “existential feelings” (2005, 2008).

attribute some sense of being to other possible experiences that are common in our everyday lives.

Matters, however, become more complicated when we start pondering about *the most basic* form of self-experience, i.e., pre-reflective self-awareness or the temporal self-givenness of the stream of consciousness. Following Zahavi, I argued that the experience of reflexivity involved in the experience of an object *x* amounts to a sense of self. The minimal self is an abstract and formal notion insofar as it applies not only to some specific and concrete experiences but—qua formal aspect of any phenomenal givenness whatsoever—to all experiences.<sup>645</sup> Yet, not all experience *per se* seems to involve an experience of a world that is comparable to the world we experience in our everyday life as human beings. As, for instance, Husserl emphasizes in his famous example of the possible annihilation of the spatio-temporal world (“Vernichtung der Dingwelt”<sup>646</sup>), for consciousness to be consciousness, it is not required that it is an experience of the spatio-temporal world that unfolds law-like in which thing-like objects appear that persist through time. Experiences can be chaotic, the unfolding of their contents unmotivated, without ceasing to present us with some conscious content. Yet, even without establishing what would be a stable reality, a temporally unfolding stream of consciousness is conceivable and as such would involve a minimal self. In fact, the claim behind the minimal self-view is that even very basic and isolated experiences, such as dim feelings of warmth or a flickering of a light against a dark background involving an otherwise absolute absence of any other thought, perception or will, would involve a minimal self.

Consider my experiential field as it is in the example mentioned earlier, when driving on the highway. Easily one can imagine how it is characterized by perceptual contents such as the street, the steering wheel or other cars, feelings (e.g., I feel afraid to tell my parents a drastic life decision), desires (e.g., perhaps I just want to get past this traffic jam) or thoughts (e.g., perhaps I am enjoying the ride and while monitoring the driving process several memories, phantasies, judgments, concerns cross my mind). Now, all these experiences that form part of my experiential field occur in a context: the concrete reality of my personal life.

Now, crucially, in reflection, I can abstract from this context and pull out from the experiential field the single experiences that form it. Some of them will require that one considers the context of the experiential field in order to understand the isolated experience in reflection. Take, for instance, the fear of telling my parents my life decisions. To understand even this isolated experience properly, it must involve the background of my life, the decisions I take and the motivation for fearing the response of my parents. In short,

---

<sup>645</sup> Though a theoretical and admittedly technical notion that is built upon abstraction, the minimal self should not be considered as a mere abstraction. Cf. Zahavi 2005, 130, Zahavi 2014a, 89.

<sup>646</sup> Husserl 1976 (Hua 3/1), 104.

isolating the experience of fear will require to include a certain sense of being that corresponds to the concrete fear and context of telling my parents the life decisions. For instance, only focusing on the bodily feelings that might be involved in the concrete experience of fear will not suffice to understand the concrete experience of fear at stake.

But how about the bodily feelings? If I were, reflecting and abstracting, interested in the bodily feelings only, what would be required to grasp them? Certainly, not my personal life, nor my interpersonal history with my parents, let alone, any interpersonal world. To understand the bodily feelings, it will suffice to refer to a very basic sense of reality, one that concerns the “localization”<sup>647</sup> of the feelings that makes them feelings of the body and, presumably, descriptions of what constitutes kinesthetic experience<sup>648</sup>. This will also include a basic sense of reality, a spatial world in which the body is located, some rudimentary ‘outside’ without which a body could not be experienced to begin with.

But how about single sensations considered, notably, still remaining in abstraction, each isolated from the others that together feature in what has traditionally been described as one’s “body schema” or “body image”<sup>649</sup>? Do these basic sensory experiences bear some sense of reality, any experience of being? And if so, how is the sense of being at stake to be characterized?

*That* is the challenge I will address in this chapter. Interestingly, the challenge has not—at least not explicitly—come up in the debate on the minimal self. The most important reason, I conject, is the fact that the debate on the minimal self is at least threefold and concerns different though interrelated questions<sup>650</sup>, which correspond to the different conceptual roles of the minimal self as described in Chapter 4:

(1) <i>Phenomenal consciousness:</i>	
What is structurally required for phenomenal consciousness to be possible?	<b>Answer:</b> The temporal self-givenness of the stream of consciousness.
(2) <i>Self-consciousness:</i>	
What constitutes the most basic structure of self-consciousness?	<b>Answer:</b> Pre-reflective self-awareness in terms of the temporal self-givenness of the stream of consciousness.
(3) <i>Selfhood:</i>	
What structural conditions must be met to be able to speak of ‘someone’? What is it that constitutes the being of selfhood?	<b>Answer:</b> The most basic form of self-consciousness: pre-reflective self-awareness.

<sup>647</sup> Husserl 1952 (Hua 4), § 36. See also Husserl 1973b (Hua 16), § 48 et passim.

<sup>648</sup> For the notion of kinesthetic experience, see Section 5.1.

<sup>649</sup> Head 1920, Schilder 1923, Gallagher 1986a, b.

<sup>650</sup> Cf. Zahavi 2014a, Ch. 2.



Now, the first two issues explicitly relate to predominantly (some might even say *merely*) phenomenological questions: the basic structure of consciousness and self-consciousness. Hence it seems that the gravitation towards the first two issues in the debate has impacted upon the way in which the minimal self-view has been seen, assessed, and discussed. Yet, whereas the question of a sense of being can be easily put into brackets when dealing with the first two issues, once the attention is turned to the third issue, it seems remarkably less evident that it could be neglected. For, the minimal self-view (and the self-experientialism implied) consists precisely in the notion that the minimal self is both the most basic form of *self-experience* and the primary way of *being* of selfhood. Naturally, then, the question arises as to how the being qua existence of selfhood figures in self-experience. Leaving this question unanswered, I suggest, makes the minimal self-view seemingly vulnerable to a number of different objections that have been voiced against Zahavi's approach in the past years, and which, I will argue, can be dissolved by clarifying how an experience of being can be involved even in the most basic form of self-experience. This, I further propose, will also require to bring *feeling* into the picture.

The chapter proceeds as follows: First, I will discuss the accusations that the minimal self is too thin and present the responses we can find in Zahavi's work. I then go on to discuss Lotze's critique of a conception of selfhood that is based on reflexivity only and find it might apply to Zahavi's minimal-self view. I conclude from that that Zahavi's responses with regard to the *too-thin-charge* might not suffice to dissolve all doubts about a thin notion of self. I then attempt to develop the notion of the *feeling of being oneself* in order to make a thin notion of self, one that involves self-feeling, plausible. To this aim, by drawing from Lotze, Husserl and Ratcliffe, I argue that feeling could be said to be an integral aspect of the temporal self-givenness of the stream of consciousness. On this view, no experience is fully neutral, because it always involves a sense of the possibility of pleasure and pain. This sense, it will be argued, presents a feeling that amounts to a *self-feeling* and conveys a sense of being: a feeling of being oneself.

## 6.1 Is Zahavi's thin self too thin?

"Thin, thinner, thinnest"<sup>651</sup> Zahavi emphasizes, responding to some criticisms, should his notion of minimal self be understood. In fact, the minimal self, pre-reflective self-awareness considered as the temporal self-givenness of the stream of consciousness, is such a thin understanding of self that it is compatible with any kind of experience whatsoever.

---

<sup>651</sup> Zahavi 2017b.

Or, to be more precise, it is so thin because only then the claim that it is a “*ubiquitous*”<sup>652</sup> feature of phenomenal consciousness is possible to begin with. After all, one central claim of the minimal self-view is that the minimal self is a feature part and parcel of any experiential state. But, as just mentioned, the minimal self plays a threefold conceptual role and, apart from elucidating the necessary structure of phenomenal consciousness and self-consciousness respectively, also is meant to account for the being of selfhood, the experience of being an individual experiencer. One important ramification of this entanglement of phenomenal consciousness, self-consciousness, and selfhood is that the minimal self applies to all creatures that have phenomenally conscious states, including (even non-cognitive) animals.

For some, however, conceptualizing the self in such a thin way so as to enable the claim that all phenomenal consciousness involves a sense of self, goes too far. Joel Krueger highlights that conceiving of the minimal self as a structural feature of consciousness renders it impotent to account for the self’s individuality: “But if the minimal self is merely an empty structural feature of consciousness, how is the phenomenal character of experience individuated? How does subjectivity become *my* subjectivity?”<sup>653</sup> A similar point is voiced by Jonardon Ganeri: “The worry about Zahavi’s ‘minimal self’ is that it is too minimal to display what makes one self distinct from another . . . .”<sup>654</sup> Dionysis Christias emphasizes “that Zahavi’s ‘experiential self’ is so thin that its connection with the pre-reflective dimension of selfhood at the distinctively human, conceptual, ‘space of reasons’ level becomes problematic”<sup>655</sup>. A somewhat similar criticism can be identified in B. Scott Rousse’s paper who claims—based on a reading of Heidegger—that the minimal self-view underestimates the role of practical self-understanding as indispensable part of the most basic form of self-experience: “Insofar as pre-reflective self-awareness involves awareness of affordances and solicitations, and insofar as awareness of affordances and solicitations depends upon pre-reflective self-understanding, *pre-reflective self-awareness depends upon pre-reflective self-understanding*.”<sup>656</sup>

Following Zahavi, in Chapter 4, I argued that pre-reflective self-awareness is to be characterized as egological, and in doing so, the aim was to show that even on the most basic level of phenomenal consciousness we are dealing with a form of selfhood that is constituted by the temporal self-givenness of the stream of consciousness. That is, an experience of reflexivity that amounts to a sense of self and thereby presents the primary form of self-being. The criticisms just mentioned, however, though not always explicitly

---

<sup>652</sup> Zahavi 2011b, 60.

<sup>653</sup> Krueger 2011, 51.

<sup>654</sup> Ganeri 2017, 1718.

<sup>655</sup> Christias 2019, 1.

<sup>656</sup> Rousse 2019, 16.

denying the egological character of pre-reflective self-awareness, seem to maintain that even if the experience of reflexivity were to amount to a sense of self, the sense of self at stake would be too formal to present the most basic form of *real* selfhood, the real existence of someone. Rather than the relationship between the *experience of reflexivity* and the *sense of self*, what is now put into question then is the relationship and association between a primordial *sense of self* and the *being of selfhood*, the reality or existence of an individual experiencer. The dispute thus now concerns whether it is sufficient to carve out a minimal sense of self in the phenomenology of experience in order to demonstrate that experience always involves a real *someone*<sup>657</sup> who is undergoing the experience.

After all, it seems possible to admit that yes, experience features a rudimentary sense of self, but at the same time to maintain that such a rudimentary sense of self is far from presenting a convincing account of selfhood. For, to some it might simply seem not enough to be aware of one's experiences to constitute selfhood. That is, it is possible to accept the notion of the minimal self plays some of its purported conceptual roles but not all of them. Hence, in this vein, it is possible to hold that the minimal self might be useful to elucidate the structure of phenomenal consciousness (insofar as phenomenal consciousness is to be characterized as something that, formally speaking, always is experienced as involving self-awareness) or the structure of self-consciousness (insofar as the association of self-consciousness and time-consciousness delivers a positive description of self-consciousness and its structure), but it might prove insufficient when it comes to the third conceptual role of the minimal self: the notion of selfhood.

A particularly pessimistic assessment of the explanatory power of the minimal self for our understanding of selfhood can be found in a recent paper by Robert W. Clowes and Klaus Gärtner who “question the necessary interrelation of the different conceptual roles played by pre-reflective self-awareness”<sup>658</sup> though they acknowledge “that these three theses are coherent”<sup>659</sup>. First, they doubt that appealing to for-me-ness can explain phenomenal consciousness, i.e., it cannot show “why marking a mental state with mineness should render that mental state conscious”<sup>660</sup>. Thus, they believe that the minimal self is not informative when it comes to the explanation of phenomenal consciousness. Though they are right that the minimal self purports to play a role in the understanding of phenomenal consciousness, I think they misunderstand the claim being made by the minimal self-view. The claim is *not* that painting an otherwise unconscious mental state in

---

<sup>657</sup> With ‘someone’, I am not necessarily referring to a person but only to an individual experiencer, a real who. ‘Someone’ in this sense presents a notion that ought to present a contrast to the ‘no one’ of allegedly anonymous experience.

<sup>658</sup> Clowes & Gärtner 2020, 625.

<sup>659</sup> Clowes & Gärtner 2020, 625.

<sup>660</sup> Clowes & Gärtner 2020, 627.

me-ish color renders the mental state conscious.<sup>661</sup> The claim, as I described it in Chapter 2, is that experience always involves for-me-ness as a necessary phenomenological ingredient. The explanatory claim that is being made refers to the notion of *causa formalis* rather than to the notion of *causa efficiens*, i.e., that for-me-ness presents a necessary formal aspect of phenomenal consciousness without which it would cease to be phenomenal consciousness. No claim is being made about an alleged combination of an unconscious mental state and for-me-ness turning the former into a conscious mental state.

The more interesting point of their critique though is another. For, according to Clowes and Gärtner, “we need a thicker notion of the pre-reflective self”<sup>662</sup>, because “Zahavi’s notion of mineness (and the minimal self) is so thin ... that there is a danger of cutting it off from anything it is like to be a self in everyday parlance”<sup>663</sup>. This is interesting insofar as they explicitly reply to Zahavi’s response to the charge that his notion of minimal self is too formal and thin. Most importantly, namely, the *too-thin-charge*<sup>664</sup> did not pass Zahavi unnoticed and he offers an answer. Let us see whether it suffices to convince the critics. As the example of Clowes and Gärtner shows, this doesn't seem to be the case. Accordingly, more might be needed. But first we need to have a look at what Zahavi has got to say.

## 6.2 Mick and Mack: What is missing in the minimal self?

To once more convince his critics, Zahavi offers a thought experiment that ought to explain not only why we should accept that for-me-ness is in-built into the phenomenology of experience but also that for-me-ness has an individualizing character so as to present a case of the minimal self:

Imagine two perfect twins, Mick and Mack, who are type-identical when it comes to their physical and psychological properties. Currently, both of them are gazing at a white wall. In terms of content, the two streams of consciousness

---

<sup>661</sup> In fact, such an approach would resemble the *Higher-Order Theories of Consciousness* (HOT) in a relevant sense. In HOT, the idea is that an otherwise unconscious mental state *p* becomes a conscious mental state *q* by virtue of the fact that reflection is directed at *p*. Crucially, the minimal self-view is opposed to any attempt that would explain phenomenal consciousness in terms of an appropriation through the ego—however construed—of an otherwise anonymous and unconscious state. Cf. Zahavi 1999, 2007.

<sup>662</sup> Clowes & Gärtner 2020, 623.

<sup>663</sup> Clowes & Gärtner 2020, 629.

<sup>664</sup> The *too-thin-charge* is to be distinguished from what Zahavi dubs the “*anonymity objection*”, the view that “there is no for-me-ness or mineness on the pre-reflective level, nor ... any I or self” (Zahavi 2014a, 26).

are type-identical. From a third-person perspective, there is no relevant qualitative difference between the two.<sup>665</sup>

From a third-personal perspective, Mick and Mack are not to be distinguished, both not only look perfectly the same, they also share the same experiential content, a white wall. Matters look differently though if we turn to the first-personal perspective:

Let us assume that I am Mick. Although my mental and physical characteristics continue to be type-identical with those of Mack, there would be, for me, a crucial difference between our respective tokens of experience, a difference that would prevent any kind of conflation.<sup>666</sup>

Whereas from a third-person perspective Mick and Mack are indistinguishable, Zahavi emphasizes, the same does not apply when considering the way the situation is presented *for* Mick. Even if Mick were to experience the white wall just as Mack does, and even if he were aware of the fact that Mack enjoys just the same experiential *content* (=the white wall), this would not change the fact that his own experiencing the white wall is, phenomenologically speaking, fundamentally different from the way in which Mick experiences Mack to be experiencing the white wall. That is, even acknowledging the type-identity of experience of his own experience and Mack's experience, both tokens of experience are given to Mick differently. *For* Mick, there is no risk of confusing his own experience with Mack's.

Crucially, according to Zahavi's proposal, what explains that fact must lie *in* the experience though not in the experiential *content*. Instead of being a matter of content, as it should not come at surprise at this stage, on the minimal self-view, it is the *how* of the givenness of the experiential content that distinguishes for Mick his own token of experience from that of Mack:

The latter [the first-personal character of experience, P.S.] amounts to a distinct but formal kind of experiential individuation. This, of course, is why (*pace* various critics) I think that the very subjectivity of experience, its first-personal character, although being quite formal, amounts to a kind of self.<sup>667</sup>

The idea, though confusing to the ears of some it might be, thus is that the *individuation* occurs through what is a *general* feature: the for-me-ness of experience which is common to the experience of all conscious beings. It should be noted though that to see a contradiction between individuation and the universal character of for-me-ness betrays a premise of dubious validity: the belief that the *individuating* phenomenology of for-me-

---

<sup>665</sup> Zahavi 2014a, 22.

<sup>666</sup> Zahavi 2014a, 22.

<sup>667</sup> Zahavi 2014a, 23.

ness would have to present itself an *individuate* phenomenology (as if the phenomenology of for-me-ness of each experiential self were qualitatively unique). That, however, would only be required if Mick's and Mack's respective token of experience would be—*per impossibile*—both given to Mick from a first-personal perspective. Only in this case would it be required that on top of being given from a first-personal perspective Mick's token of experience had an additional ingredient that helps to tell it apart from Mack's token of experience. But such a problem does not emerge, as Mick's and Mack's token of experience each are given distinctively to Mick and Mack respectively. That from a point of nowhere, from some kind of a metaphysical perspective, as it were, both Mick's and Mack's experiences are qualitatively absolutely identical does not amount to a problem of individuation at all. For, the two tokens of experience could not possibly be given both from the first-person perspective at the same time without collapsing into one—quite inconceivable—experiential state that would be identical to neither Mick's or Mack's token of experience. The problem that both tokens of experience are lived through in one experiential field is impossible and therefore the problem of distinguishing between the two does not emerge. Or, to put it differently, they are always already distinguished in virtue of their distinctive givenness: having a first-personal character (self-awareness or for-me-ness) or second-personal character (empathy).<sup>668</sup>

I take it then that Zahavi does have a reply ready at hand when it comes to countenancing worries about individuation. Unfortunately, though, it seems, accounting for individuation in terms of for-me-ness still does not suffice to convince the critics. However, I want to propose that phenomenologically describing the sense of being inherent to the minimal self further might persuade those who still believe the minimal self is too thin. To develop my proposal, which I suggest might release the pressure on the minimal self-view stemming from the too-thin-charge, let me hear out another critical voice.

### 6.3 Caring about oneself—a Lotzean idea on *Fürsichsein*<sup>669</sup>

“Not in the fact that thinking corresponds to the thought lies the meaning of self-consciousness; for this aspect is not unique to *our* I alone, but is the general nature of *every*

---

<sup>668</sup> Zahavi 2011a, Zahavi 2014b.

<sup>669</sup> Frechette (2013, 659) and Textor (2018) translate *Fürsichsein* with *for-me-ness*, whereas the original English translation has “immediate self-existence” (1885, 2:679). Textor highlights that the latter “does not get close to the literal meaning” (2018, 175). But I think that is too strong a rejection, because the advantage of ‘immediate self-existence’ is that it shows that we are dealing with existence and not only an alleged essence or merely form of selfhood. The literal meaning would be best expressed by *being-for-oneself*. Yet, since I believe *for-me-ness* cannot be thought independently of being, that essence and existence of selfhood are inseparable, and given that the notion of *for-me-ness* is now well-established in the current debates, I will translate *Fürsichsein* with *for-me-ness*. When I want to stress the difference between

I”<sup>670</sup>, Lotze states in his epochal *Mikrokosmos*, seemingly committing what I just suggested is the error of a confusion between the individuate and individuating phenomenology of for-me-ness. However, matters are more complicated. For, Lotze explains just in the next sentence what he believes plays the individuating role in self-consciousness and through which we distinguish ourselves from others: “Obviously through the fact that it is the thinking of *our* thoughts.”<sup>671</sup> Apparently, then, he seems to concur with Zahavi’s view that it is the for-me-ness or mineness that does the individuating. That, however, would be too hasty a conclusion. In fact, Lotze both shares some insights with Zahavi’s view and disagrees on some other aspects. For, although he would agree that it is the mineness of experience that is individuating, he would nonetheless propose a different theory of how the mineness of experience comes about and how it is to be characterized—namely *not* through the sheer “reflexivity”<sup>672</sup> of experience and *not* as a general let alone invariant feature.

The tension that stems from the comparison of Lotze’s and Zahavi’s view on mineness, as I now want to elaborate, will prove fruitful for the contemporaneous discussion of the minimal self-view, notably the experience of individuation and the experience of being. For, what I want to suggest, Lotze’s position is to be found in the middle between Zahavi and those who criticize the thinness of the minimal self, thereby providing, as I propose, some new options. Interestingly enough, Lotze himself conceives of his own position as some kind of “middle way”<sup>673</sup>. Let me briefly illustrate the two lines of thinking that he criticizes and which require on his view a third position that lies between them: his own proposal.

### 1) *Pervasive reflection (= reflection theory of self-consciousness)*

The first line of thinking consists in the view that our conscious life is equipped with a form of self-consciousness that distinguishes between I and Not-I and that consists in an explicit intentional awareness of oneself as the center of experience—which ought to persist across the different states of mind.<sup>674</sup> In other words, the view Lotze has in mind is a variant of a

---

*Fürsichsein* (=for-me-ness involving an experience of being) and mere *for-me-ness* independently of the question of the experience of being, I speak of *Fürsichsein* without translating it.

<sup>670</sup> Lotze 1884/1856, 279, my translation: “Nicht darin liegt die Bedeutung des Selbstbewußtseins, daß Denkendes und Gedachtes zusammenfallen, denn dieser Zug bezeichnet nicht *unser* Ich allein, sondern die allgemeine Natur *jedes* Ich . . . .”

<sup>671</sup> Lotze 1884/1856, 279, my translation: “Dadurch gewiß, daß es das Denkende *unserer* Gedanken ist.”

<sup>672</sup> See Textor 2018, 171.

<sup>673</sup> Lotze 1884/1856, 278, my translation: “mittlerer Weg”.

<sup>674</sup> Lotze 1884/1856, 278: “Ich meine das Selbstbewußtsein, in welchem wir uns als Ich von dem Nicht-Ich der übrigen Welt unterscheiden und die Mannigfaltigkeit der inneren Zustände auf dies Ich, als den zusammenhaltenden Mittelpunkt aus- und eingehender Wirkungen, beziehen.”

reflective theory of (self-)consciousness, i.e., the idea that it is egological consciousness based on reflection that ought to be intrinsic to consciousness, even from birth on.

2) *Ego by education (= non-egological theory of self-consciousness)*

The second line of thinking consists in the notion that self-consciousness is something that emerges only through education and the process of development over the course of time, “as a fortunate by-product of the mechanism of the stream of thought”<sup>675</sup>. In other words, what Lotze has in mind here is, roughly speaking, the idea of a non-egological theory of (self-)consciousness, the idea that consciousness per se does not feature self-consciousness *in sensu strictu* but only may give rise to self-consciousness depending on how consciousness unfolds over time.

Evidently, both lines of thinking are also positions the minimal self-view is diametrically opposed to and which have even played a crucial role—qua contrast foil—in Zahavi’s development of the minimal self.<sup>676</sup> That pushes both Lotze’s and Zahavi’s position somewhere in the same area, but as it will turn out, they don’t land on the exact same spot. The reason for that lies in the different argumentative routes to be found in each respectively. Most importantly, while they both share some criticism of the reflective theory of (self-)consciousness and non-egological theory of (self-)consciousness, some of the criticism directed by Lotze against the reflective theory of self—at least in *some* regard, as I will argue—apply also to Zahavi’s pre-reflective though reflexive notion of minimal self. Moreover, crucially, the very same argument that is partly applicable to Zahavi’s minimal self undermines his critics (mentioned in Section 6.1) alike—if not, as I believe, to a much greater extent. That constellation makes Lotze’s position particularly interesting. Let me now explain it.

First, it is important to note that Lotze was in possession of “Fichte’s original insight”<sup>677</sup> that reflection cannot bring about self-consciousness out of an otherwise anonymous and selfless consciousness. As he makes abundantly clear, he finds such a notion problematic:

We can skip silently all those attempts, which, following ill-chosen analogies of the world of sensory perception (*Sinnenwelt*), purport to show how an activity of the selfless being originally directed at the external ought to turn into

---

<sup>675</sup> Lotze 1884/1856, 278, my translation: “als ein glückliches Nebenerzeugniß aus dem Mechanismus des Vorstellungsverlaufes”.

<sup>676</sup> Zahavi 1999, 2007.

<sup>677</sup> Henrich 1966.



a self-illuminating self-consciousness mediated by the resistance of the world of the Not-I as if it were a reflection of a ray of light hitting a surface.<sup>678</sup>

For Lotze, it is simply “throughout inconceivable how it [the activity, P.S.] could change its nature by reflection and turn what is a blind activity into the selfhood of for-me-ness (*Fürsichsein*)”<sup>679</sup>. Importantly, this rejection is directed at both lines of thinking that Lotze takes to be wrong-headed. According to the first line, self-consciousness is present from birth on, but involves reflection that ought to be *logically* primordial to or the ground of self-consciousness. According to the second line, self-consciousness only develops over time through reflection, i.e., reflection ought to be *temporally* and *logically* primordial to self-consciousness. In both cases, on Lotze’s view, it is however principally unclear how reflection could transform an otherwise anonymous consciousness into the self-consciousness of someone. Put differently, that argument, which he shares of course with Zahavi,<sup>680</sup> alone could be enough to undermine the two lines.

But Lotze introduces and focuses on another argument. The problem that he sees with defining self-consciousness and for-me-ness (*Fürsichsein*) in terms of a reflective self-relatedness is that, even though in being self-aware of my experience when experiencing some object *x*, my experience might be given differently compared to the experience of *x*, this alone does not suffice to make a—notably—real distinction between I and Not-I. Or to be more precise, Lotze thinks the fact of reflective self-awareness as such does not elucidate the essential distinction between I and Not-I as compared to the distinction between different objects (on the side of the Not-I). For Lotze, what makes selfhood unique and individual is that in being self-conscious or for-me-ness (*Fürmichsein*) the distinction between I and Not-I is absolute. It is not on a par with the distinctions to be found between objects on the side of the Not-I. But this asymmetry between the distinction ‘I vs. Not-I’ and the distinction between ‘Not-I (object 1) vs. Not-I (object 2)’ cannot be cashed out, according to Lotze, if we rely only on reflection as a cognitive or epistemic activity (*Erkenntnis*).

To illustrate that thought, Lotze mentions the example of an angel, a being that enjoys perfect reflective knowledge of itself, i.e., including its own inner experiential states and soul, but which would always lack a real understanding of the distinction between itself and the world, for the simple reason that in knowing and cognizing all these fine-grained

---

<sup>678</sup> Lotze 1884/1856, 572, my translation: “Mit völligem Schweigen können wir daher alle jene Versuche übergehen, welche nach übelgewählten Analogien der Sinnenwelt zu zeigen denken, wie eine ursprünglich nur nach außen gerichtete Tätigkeit des noch selbstlosen Wesens durch den Widerstand, den ihr die Welt des Nicht-Ich wie eine Fläche dem Lichtstrahl leiste, in sich zurückgelenkt und dadurch in das sich selbst erfassende Licht des Selbstbewusstseins verwandelt werde.”

<sup>679</sup> Lotze 1884/1856, 572f., my translation: “ganz unbegreiflich, wie sie durch diese Reflexion ihre Natur verwandeln und aus blinder Thätigkeit zur Selbstheit des Fürsichseins werden könnte”.

<sup>680</sup> Zahavi 1999, 2007.

details about itself it would still run short of an understanding why it should care more about itself than about the world and the things to be found in it. Put in other words, the angel would be *indifferent* to its own being—and in *this* sense the angel would fail to understand the distinction between itself and the world.<sup>681</sup> For Lotze, it is not enough to be able to discriminate *between* (a) what are the own experiences and the own experiencing being on the one hand *and* (b) the objects of experience on the other hand. Such a *discrimination* does not overcome a fundamental *indifference* intrinsic to cognition (*Erkenntnis*) that does not yet amount to the *distinction* between I and Not-I. Put differently, the phenomenology of the distinction between I and Not-I is not exhausted by a *cognitive* distinction, it must contain more.

But what is that ‘more’? For Lotze, it is the value that one sees in one’s own life, the self-interest or self-concern that presents part and parcel of for-me-ness (*Fürsichsein*) and that literally makes all the difference. As Lotze further argues, in order to be able to experience one’s own value, in fact, no reflection is required, i.e., no higher-order states of mind are necessary that are only available to well-developed creatures like the human being. To show that, he refers to the example of a crushed worm writhing in pain. Though the worm has no cognition *in sensu strictu*, i.e., lacks conceptual thinking and the capability to reflect, Lotze emphasizes, it nonetheless makes a distinction between its own being and its environment. It experiences its own being by taking an interest in it. And now the question arises: How does even the worm, whose experience lacks reflection, experience for-me-ness (*Fürsichsein*)? The answer, for Lotze, is obvious: The phenomenological way to experience the own being and of self-concern is *feeling*:

Not through the fact that we think the coincidence of the thinking with the thought, but by *feeling the immediate value* that it [the coincidence, P.S.] has for us does it ground our self-consciousness and elevates irrevocably the distinction between ourselves and the world over those oppositions through which one object is distinguished from another.<sup>682</sup>

The “genuine interest”<sup>683</sup> that we take in the distinction between ourselves and the world is what (at least partially) constitutes the distinction to begin with, and crucially this interest, just like any other value, on Lotze’s view, is given “through feelings of pleasure and pain”<sup>684</sup>. It is only because of the feelings that the distinct givenness of my experience as

---

<sup>681</sup> Lotze 1884/1856, 281.

<sup>682</sup> Lotze 1884/1856, 280f., my translation and emphasis: “Nicht indem jenes Zusammenfallen des Denkenden mit dem Gedachten von uns gedacht, sondern indem es in dem unmittelbaren Werthe, den es für uns hat, gefühlt wird, begründet es unser Selbstbewußtsein und hebt unwiderruflich den Unterschied zwischen uns und der Welt über alle Verleihung mit den Gegensätzen hinaus, durch die ein Gegenstand sich vom andern sondert.”

<sup>683</sup> Lotze 1884/1856, 279, my translation: “eigenthümliches Interesse”.

<sup>684</sup> Lotze 1884/1856, 280, my translation: “durch Gefühle der Lust und Unlust”.

compared to how the world and the objects in it are experienced by me matters in any relevant sense. That is, even though, as exemplified by Lotze’s angel as a “perfect intelligence”<sup>685</sup>, without any feeling, i.e., pleasure or pain,<sup>686</sup> whatsoever one might be able to distinguish between one’s own experience and the world, but it would not matter. For, although without feeling there might be some *epistemic* distinction between self and world, there would not be a *real* distinction, there would not be *for-me-ness* proper. Lotze’s worry thus is that if the *de se* remains only cognitive or epistemic it collapses into some kind of *de re*—despite the asymmetrical givenness between the self-awareness of my experience and the experience of some object *x*. Lotze’s alternative suggestion then is the following:

Hence *self-consciousness*, on our account, *is only the explication of a self-feeling* the primordial and original vitality of which is not directly augmented by the development of knowledge ...<sup>687</sup>

Preceding to any self-consciousness, with which Lotze refers to *reflective* self-consciousness, we have to acknowledge a pre-reflective self-feeling that grounds the possibility of reflection and the real distinction between oneself and the world:

Self-consciousness is only an additional effort to analyze the facts that are lived through with the measures of knowledge (*Erkenntniß*) ...<sup>688</sup>

[W]e would admit that the ‘I’ can only be *thought* in relation to the ‘Not-I’, but we would add that it [the ‘I’, P.S.] is already *capable of being lived through* (*erlebbar*) prior to any such relationship, and that this primarily grounds the possibility that it can be thought in form of such a relationship in the first place.<sup>689</sup>

Lotze’s second argument against the reflection theory of self thus consists in the notion that reflective self-consciousness and the distinction between I and Not-I is parasitic on, i.e., presupposes, a self-feeling of one’s own absolute value—*absolute*, because the self-

---

<sup>685</sup> Lotze 1884/1856, 281, my translation: “vollendete Intelligenz”.

<sup>686</sup> Importantly, Lotze distinguishes between sensation (*Empfindung*) and feeling (*Gefühl*). Cf. Lotze 1846, 190, Lotze 1852, Ch. 2 § 20, Lotze 1889/1881, 48. Hence, Lotze’s ‘perfect intelligence’ would have sensations that are involved in perception, but no feelings of pleasure and pain.

<sup>687</sup> Lotze 1884/1856, 281, my translation and emphasis: “So gilt uns das *Selbstbewußtsein* nur für die *Ausdeutung eines Selbstgefühls*, dessen vorangehende und ursprüngliche Lebendigkeit durch die Ausbildung unserer Erkenntniß nicht unmittelbar gesteigert wird ...”

<sup>688</sup> Lotze 1888/1864, 572, my translation: “Das Selbstbewußtsein ist nur die später kommende Bemühung, mit den Mitteln der Erkenntniß diese erlebte Tatsache zu zergliedern ...”

<sup>689</sup> Lotze 1888/1864, 572, my translation and emphasis: “würden wir ... unsere Stellung so nehmen, daß wir zugäben, Ich sei *denkbar* nur in Beziehung auf Nicht-Ich, aber hinzufügten, es sei vorher außer jeder solchen Beziehung *erlebbar*, und hierin eben liege die Möglichkeit, daß es später in jener Form denkbar werde”.

concern does not hinge on a certain feature or experiential content but adheres to the for-me-ness (*Fürsichsein*) as such:

On closer examination does this fact [the coincidence between the reflecting and the reflected consciousness, P.S.] explain very little of the puzzle of the genuine interest that we have with regard to that distinction [between I and Not-I, P.S.], and which has very little in common with the sheer embeddedness in a specific case and the genuine character of interest that is attached to it.<sup>690</sup>

Lotze thus highlights that the self-interest or self-concern that we all, at least at some point in our life, are explicitly aware of cannot be explained either by reflection on ourselves (and our experience) nor by some specific circumstance of our concrete life. However, since we do generally care about ourselves and since, on Lotze's view, it is that kind of self-concern that precedes and is required by any real distinction between ourselves and the world, we must understand the primary way in which we are for us (*Fürsichsein*) as a way that contains an experience of our self-value, which only can be characterized as feeling.

If Lotze is right about this, then this has immediate ramifications for the two lines of thinking in the middle of which he finds his own position. Let me now determine where his position lies between the two lines with more precision and show how it also allows for a new perspective on Zahavi's account and his critics.

### 6.3.1 From Lotze to a new perspective on Zahavi's notion of minimal self and its critics

So far, I have tried to demonstrate that Lotze's notion of pre-reflective self-feeling that precedes any reflective self-consciousness is directed against both (1) the view that we possess reflective self-consciousness in any moment of consciousness and invariantly even from birth on and (2) the view that we develop self-consciousness only over the course of time in virtue of a reflective attitude towards our otherwise anonymous consciousness motivated and fueled by education and conceptual knowledge. But what exactly does Lotze mean when he says we need a middle way between the two and in which sense does his own approach present such?

First, it is important to note that he shares with (1) the idea that experience as such is endowed with for-me-ness (*Fürsichsein*), though he disagrees that it ought to be some kind of reflection built-in to experience doing the job of providing for-me-ness. In contrast to (1) and in line with (2), Lotze emphasizes that reflection is something that comes about only over the course of time, shaped and informed by our encounter with the world. But in

---

<sup>690</sup> Lotze 1884/1856, 279, my translation: "Aber genauer betrachtet erklärt dieser Umstand sehr wenig das Rätsel des eigenthümlichen Interesses, das wir an diesem Unterschiede nehmen, und das sehr wenig mit der bloßen Theilnahme an der interessanten Eingenthümlichkeit eines besonderen Falles gemein hat."

contrast to (2), he still believes that in the very first moment of experience, even prior to any explicit discrimination between oneself and the world, our experience is already characterized by for-me-ness (*Fürsichsein*) that consists of a pre-reflective self-feeling, a feeling of one's own absolute value. The crucial point is that the feeling of absolute value of oneself cannot be learned or gained through education nor through the most detailed knowledge about oneself provided by meticulous reflection on the individual constitution of one's own concrete and personal inner life.<sup>691</sup> Lotze's position, then, is a middle way insofar as it also defends the view that for-me-ness (*Fürsichsein*) is not something that must be developed through education but is built-in to the architecture of experience; while it holds that it is not as a product of a reflective cognition but in terms of a pre-reflective feeling that for-me-ness (*Fürsichsein*) is present.

Most importantly, Lotze's argument against (2), the view that enriching the knowledge about oneself is what ultimately enables and constitutes for-me-ness (*Fürsichsein*), also plays a role in some of those who are skeptical about the thinness of the minimal self. Let me therefore now come back to those criticisms of the minimal self that place emphasis on the need for a thicker notion of self, and which suggest to enrich the concept of the minimal self in one or another way. For instance, Christias's who sees a need to account for the human-specific, normatively shaped "space of reasons"<sup>692</sup> which ought to be integrated into the minimal self; Rouse's claim that "pre-reflective self-awareness depends upon pre-reflective self-understanding"<sup>693</sup> of one's own practical concerns and possibilities; or Clowes and Gärtner's call for a "pre-reflective situational self"<sup>694</sup>. What these criticisms and alternative suggestions share is the notion that, even on the pre-reflective level of the minimal self, selfhood must include an awareness of one's own concrete situation constituted by specific possibilities, norms, and practical concerns. For an experience to have the character of for-me-ness, these authors suggest there must be an experience of the specific situation that makes *my* experience the experience of a concrete individual and thereby *mine* to begin with. Like (2), these authors believe that a better understanding of one's own individual condition is required to constitute for-me-ness to begin with. Unlike (2), though, they do not think it is reflection that brings about this better understanding. Rather, they take it to be a constituent even on the pre-reflective level already.

Does this difference in experiential locus, the emphasis on the pre-reflective level rather than the reflective level, change anything with regard to Lotze's argument concerning the need for an immediate awareness of one's own value as the decisive

---

<sup>691</sup> Cf. Lotze 1884/1856, 280.

<sup>692</sup> Christias 2019, 5.

<sup>693</sup> Rouse 2019, 16.

<sup>694</sup> Clowes & Gärtner 2020.

grounding aspect of for-me-ness (*Fürsichsein*)? As I have already indicated, I don't think that is the case. In fact, I believe it does not matter whether one conceives of the experience of one's individual situation to be a product of reflection or to be something that already characterizes one's pre-reflective experience. The crucial point is another, namely that no awareness, experience or knowledge about what is my situation or individual constitution can bring about the genuine self-concern that Lotze claims to be an essential aspect of selfhood, but in fact rather remains reliant on that very self-concern.

Let me develop that point by discussing one of the criticism directed at Zahavi in more detail: Rouse's appeal to Heidegger.<sup>695</sup> Rouse points to the "practical self-relation" that is to be contrasted with a merely "theoretical self-relation" because "it has *not* just to do with issues of epistemic access to or awareness of my own experience or 'states'"<sup>696</sup>. Rather, Rouse emphasizes, for Heidegger *mineness* (*Jemeinigkeit*) is something that refers to the fact that "I relate to my life or my own being as something that is on-goingly *at issue* or *at stake* for me"<sup>697</sup>. Importantly, then, for Heidegger *mineness* (*Jemeinigkeit*) involves and is grounded on concrete practical concerns that refer to my very own possibilities constituting my situation in the world. Being situated in the world, we experience objects that present themselves as useful for a certain goal or that present a threat to what we take our possibilities to be. What we experience in the world is imbued with significance that refers to our practical concerns, our goals and projects. The experience of the world thus always reflects how we understand ourselves, what we take our possibilities to be. Rouse emphasizes that it is such a pre-reflective self-understanding implied in our experience of the significance of our environment on which pre-reflective self-awareness rests, and without which *mineness* is not possible.

However, and that would be the point I would like to make drawing from Lotze, it is not the experience of significance that conveys the self-interest with or the self-concern about my own being. Quite to the opposite, it is rather the intrinsic and primordial self-concern that enables the experience of significance to begin with. Heidegger is very explicit about that when he says:

Existing being-in-the-world as such is disclosed in the for-the-sake-of-which, and we called this disclosedness understanding. In understanding the for-the-sake-of-which, the significance grounded therein is also disclosed.

...

---

<sup>695</sup> Rouse 2019.

<sup>696</sup> Rouse 2019, 14.

<sup>697</sup> Rouse 2019, 14.

The statement that the for-the-sake-of-which and significance are disclosed in Da-sein means that Da-sein is a being which, as being-in-the-world, is concerned about itself.<sup>698</sup>

For Heidegger, the for-the-sake-of-which does not refer to a specific goal we might pursue but concerns the being of *Dasein*, which is itself not yet just another ‘something being useful for *x*’ but precisely the ground of all significance: my own being as such insofar as it is of absolute value, that which my whole being is about, allows me to experience something as significant in the first place.<sup>699</sup>

Without mentioning Lotze, who has been undoubtedly of great influence on him, Heidegger acknowledges that it is out of a fundamental self-concern that the experience of significance and world unfolds for us. It is only because *Dasein* intrinsically cares about itself, about its own being, that something can matter and shape what is *my* situation.

What does that mean for the criticism of Zahavi’s thin notion of self? Rouse suggests that “pre-reflective self-awareness is not indifferent to one’s practical-existential immersion in a situation where, in light of one’s pre-reflective self-understanding, things and possibilities already matter and solicit one to act”<sup>700</sup>. And from that, he concludes that “Zahavi’s account loses the right to claim a primal foundational status for the experiential self, but it gains a more expansive taxonomy and phenomenology of the pre-reflective modes of human agency and selfhood”<sup>701</sup>.

I think Rouse is right to emphasize the role of significance in the structure of our experience of the world, and more so if he refers to the pre-reflective modes of *human* selfhood. Even so, the problem with enriching the notion of the minimal self in such a way is that it implicitly amounts to the view that only humans and non-human animals that experience world in a human way present cases of selfhood. Or he could claim like Christias that “the minimal experiential self of humans (and its empathic powers) would have to be conceived as being *different in kind or form* from the minimal self of non-human animals or human infants”<sup>702</sup>. However, I suggest this is a deep confusion about what the notion of minimal self is to provide. Of course, there is a difference between non-human selfhood and human selfhood, but it makes no sense to distinguish between different forms of *minimal* self. There are the minimal criteria for something to qualify as selfhood, and once they are met the concrete selfhood at stake can be very different in kind as compared to another form of selfhood. What is it that makes both cases examples of selfhood? *This*

---

<sup>698</sup> Heidegger 1996, 134 (2006/1927, 143), see also 78 (84).

<sup>699</sup> As Heidegger writes: “There we found that the referential totality of significance (which is constitutive for worldliness) is ‘anchored’ in a for-the-sake-of-which.” (1996, 179 (2006/1927, 192))

<sup>700</sup> Rouse 2019, 16.

<sup>701</sup> Rouse 2019, 17.

<sup>702</sup> Christias 2019, 6.

is the question the minimal self gives an answer to. Hence, if Rouse, Christias, Clowes and Gärtner, among others, suggest the minimal self is too thin and needs to be enriched by aspects of human agency or experience of the world, it should be noted that doing so one runs the risk of conceptualizing selfhood in such terms that it only applies to human beings.<sup>703</sup> But what if we were to grant that some non-human animals do have phenomenal consciousness and self-experience, should we refrain from speaking of selfhood? I don't think the critics of Zahavi's minimal self have a good point in claiming that in these cases we should not speak of selfhood because these animals do not present the kind of human existence Heidegger conceptualized as *Dasein*. If Lotze's worm writhing in pain has phenomenal consciousness, self-experience and in feeling pain cares about itself, I don't see why we should not attribute selfhood to it only because it has no understanding of objects distinct from it as useful for the purpose of *x*. Put differently, it is one thing to point to the role of the experience of significance for selfhood and another to claim that a human-like experience of the world with all its complexity is necessary to understand the experience of significance (which then ought to ground selfhood). Again, I think Rouse is right when he, just like Lotze, emphasizes the experience of significance as an essential feature of selfhood. But I think he is wrong when he suggests that the experience of significance hinges on the complexity of the human world and the fact that human beings can understand *x* as useful for *y*. As I think Lotze's argument makes clear, and as I have tried to show even in Heidegger, it is primarily self-concern that grounds any experience of significance and which can even be attributed to non-human animals. We don't need to assume a pre-reflective understanding of the *human* situation in order to account for the experience of significance which is derived from the more fundamental self-concern. And it is therefore that I believe Rouse errs when he thinks highlighting the role of the experience of significance requires to enrich the concept of the minimal self by the "phenomenology of the pre-reflective modes of *human* agency"<sup>704</sup>.

Accordingly, I am neither—if not even less—persuaded by his claim that Zahavi "loses the right to claim a primal foundational status for the *experiential* self"<sup>705</sup>. First, even if human-like experience of significance were decisive for selfhood, and even if human-like agency were to be taken to be decisive for selfhood, in both cases we would still deal with an experiential self. For, trivially, not only is an *experience of significance* still an *experience* but also everything that is distinctively human action and agency involves

---

<sup>703</sup> That, of course, is a possible theoretical option, which, as I have already stressed in Section 5.2, is a delicate matter. Here, I only want to emphasize that adopting a stance that ties selfhood with a human-kind of experience has the implication that conscious animals that are not endowed with a human-kind of experience would then be deprived of their status of being a who or someone. I find that is unacceptable a position and against the intuition of many, though I cannot further go into these important discussions in this dissertation.

<sup>704</sup> Rouse 2019, 17, my emphasis.

<sup>705</sup> Rouse 2019, 17, my emphasis.



experience in one way or another. Second, I do not see why the self-concern that grounds the experience of significance should not refer to a concern about oneself qua experiencing being. If the worm in pain cares about itself, it cares about itself as experiencing, the pain is a striving towards relief—which is itself essentially an experience and prefigured as such *in* the experience of pain.

To sum up, I think making the notion of minimal self thicker by loading it with aspects specific to human experience is not only unnecessary and unwarranted, given the reasons for it provided, it also comes with the problem that it implies that selfhood is only to be found in human or human-like experience. Moreover, such an approach would also have to give reasons why self-awareness if it were to be attributed to non-human experience would *not* amount to a more basic form of selfhood. I cannot see any reason for that other than a motivation to privilege human being over other animals. In any case, appeal to self-concern and significance is not sufficient to make selfhood specific to human *Dasein*, because self-concern and significance might in fact also be integral to a much lower developed experiencing being that is not equipped with conceptual thinking or similarly construed understanding (e.g., pre-conceptually and performatively taking the hammer to be useful for hammering).

What I want to suggest in the following is that if one were to enrich the notion of the minimal self, then one would have to preserve its thinness. Crucially, as I want to elaborate now, this does not preclude that we conceptualize the minimal self as described by Zahavi in a way that involves self-concern. Before I turn to the question as to how I think self-concern and minimal self might in fact be compatible and intertwined, let me first deal with the question why we should make any attempt to integrate self-concern into the minimal self at all.

### 6.3.2 Lotze's *Fürsichsein*, self-concern, and Zahavi's minimal self

To anticipate, my main point why I think self-concern and feeling should be integrated into the minimal self is that Lotze's central argument against the reflective theory of self also applies to Zahavi's notion of pre-reflective though reflexive self-awareness. The problem is that the mere reflexivity, the mere self-awareness of my experience, cannot explain nor bring about that we care about the distinction between ourselves and the world. The question whether self-awareness is pre-reflective or only to be found on the reflective level is orthogonal to that issue. If the minimal self is conceptualized in such a way that it is wholly *indifferent* to itself, then it simply becomes a puzzle how or what exactly it is on a higher level that ought to make the *difference*. If Lotze and Heidegger are right that the experience of significance is dependent on a primordial self-concern, which I believe they are, then there is no content whatsoever that could initiate any self-concern if it isn't yet

already at stake on the pre-reflective level. For, that was precisely Lotze's point. Neither reflection and the shift in experiential structure that ensues from it nor any reflective knowledge about my own inner experiential constitution can introduce self-concern about myself. If I haven't been pre-reflectively caring about myself already, whether I reflect or gain knowledge about myself doesn't change anything in that regard.

Now, there seems to be a rejoinder available on Zahavi's end. For, he could simply claim that the minimal self is only an abstract notion that covers a moment of a more complex life—and which only ought to show what is required for phenomenal consciousness and what the general structure of the most basic form of self-consciousness (i.e., pre-reflective self-awareness) consists of. Yet, such a rejoinder would only cover the first two conceptual roles of the minimal self—and not the one that refers to the minimal self as it features in self-experientialism, the minimal self as giving rise to the being of someone. Thus, if we want to lay out—as I have done in the previous chapters—the minimal self as the primary mode of being of selfhood, then one cannot maintain that with selfhood one is only referring to the mere formal *selfhood* or *for-me-ness* independently of any being. As one can say with Husserl, the essence and existence of self are inseparable.<sup>706</sup> Moreover, “I exist not only, I am all the time also *being for myself (für mich seiend)* and as that always and necessarily conscious of myself in the most primordial originality.”<sup>707</sup> Any attempt, and I am not saying Zahavi is engaging in it,<sup>708</sup> to extract some concept of selfhood while abstracting from all sense of being is futile. When there *is* someone (as an individual experiencer), that particular someone is taking herself as being, i.e., has a sense of being.

Thus, my suggestion is, when we are dealing with the third conceptual role of the minimal self, we are talking about the minimal self insofar as it gives rise to the being of someone. And then we have to account for a primordial way of self-concern involved. For, if Lotze is right, it is the self-concern that constitutes any sense of being, *Fürsichsein*—or literally *being-for-itself*—in the first place. Put differently, while Zahavi's *for-me-ness* could still—though wrongly—be interpreted as a merely formal description of the essential structure of selfhood (without bearing in mind the being of selfhood), *Fürsichsein* emphasizes the being, the existence of selfhood. Crucially, the difference between mere reflexive *for-me-ness* in terms of Zahavi and Lotze's notion of *Fürsichsein* is that the latter explicitly tries to account for the immediate sense of one's own being, wherefore the mere reflexivity is not sufficient. And importantly, Lotze tries to account for the sense of being in terms of self-concern. *This* is the positive reason why I suggest Lotze's account is informative for the discussion of the minimal self. Lotze provides a different perspective

---

<sup>706</sup> Husserl 1973a (Hua 14), 159.

<sup>707</sup> Husserl 2006 (Hua Mat. 8), 364, my translation and emphasis: “Ich bin nicht nur, ich bin ständig auch für mich seiend und als das allzeit und notwendig mir in ursprünglichster Originalität bewusst.”

<sup>708</sup> The only thing I would like to note is that he has not dealt with the question how a sense of being could be involved on the pre-reflective level of experience.

on the challenge of individuation that Zahavi's critics find the minimal self to be failing at. His proposal with regard to self-concern presents a different perspective, because he suggests a solution of the challenge of individuation without appealing to a more enriched experiential structure or some content that ought to constitute self-concern and individuation. Accordingly, his proposal is still 'formal'<sup>709</sup> and seems to be compatible with a thin notion as Zahavi's minimal self.

Let me now explore how such a compatibility of Lotze's notion of self-concern and Zahavi's minimal self could look like, and how feeling—the way self-concern as the *ratio essendi* is said to manifest itself experientially—could be integrated into the minimal self-account.

## 6.4 Integrating feeling into the minimal self account

Before engaging in the attempt to such an integration, one immediate worry needs to be addressed: If the minimal self qua pre-reflective self-awareness is what is intrinsic to any experience *x*, then doesn't feeling which is an experience also involve once again the minimal self? For, if feeling is an experience then doesn't that require that there is some awareness of that feeling? And if so, doesn't that precisely show that the minimal self is already presupposed by any feeling to the effect that feeling itself cannot possibly form part of the minimal self that it seems to require?

That, however, would only be the case if feeling were yet another content, something—to speak with Husserl—that presents an additional *impressionale* datum alongside other sensory contents that determine the experiential content of a given moment. But as I have argued in Chapter 2, notably drawing from Zahavi himself, experiential givenness is not exhausted by experiential content. And crucially, the feeling phenomenon I am targeting now is not concerned with a specific object that is felt but rather with the experiential process—in which objects can be given to begin with—itsself: the being of experiencing, that which in reflection we objectify when we consider ourselves experiencing beings. The feeling phenomenon at stake is a self-feeling of the experiential process. And the experiential process is manifest in a non-objective way, something that is not some second-order content, but rather in-built to the way in which any content whatsoever is given. Hence, the idea is that the self-feeling of self-concern is part of the phenomenological structure of the minimal self, rather than some addition to it. Crucially, while pre-reflective self-awareness might *conceptually* be distinguished from the pre-

---

<sup>709</sup> With 'formal' here I mean only that, on Lotze's view or at least according to my reading of it, self-feeling doesn't require a specific experiential content but adheres to the form of experience as such. I will further specify this claim in Chapter 7.

reflective self-feeling, this doesn't mean they couldn't be *phenomenologically* intertwined, inextricable even—and form together the most basic form of self-experience. The task to which I want to dedicate the following sections, then, is to investigate whether there is a way in which feeling could be integral to Husserl's notion of inner time-consciousness, the very phenomenological structure of Zahavi's minimal self.<sup>710</sup> Interestingly, we find a description of such a task in Husserl, though, as far as I am concerned, he never fully, nor for that matter satisfactorily (neither to himself), accomplished it. Even so, his deliberations not only provide some first steps, it is also precisely the problems he is confronted with that gesture towards the direction in which to look for an alternative solution. As such, then, his analyses present a good point of departure for the development of my own approach.

#### 6.4.1 Husserl's *lebendige Gegenwart*: the 'I' as feeling in affection and action

In his later texts on time-consciousness, dating from the years 1929-1934, the so-called *C-Manuscripts*, Husserl focuses on an analysis of what he calls “lived presence” (*lebendige Gegenwart*), experience as it is lived through without it being reflected on. In this context, he also addresses the primal process of the temporal givenness of experiential content prior to not only any reflection whatsoever, but even prior to the intentional activity (in sensu strictu) of the 'I' attending the object given. The guiding question for Husserl is how to describe the phenomenological structure of consciousness in which something presents itself in experience the most original way: What is the phenomenological structure of the consciousness of being affected by *x*?

For Husserl, that question is important because it is the primary mode of any lived presence whatsoever (“*urmodale Gegenwart*”<sup>711</sup>) and therefore the most basic ground out of which our world-experience emerges. To pin down the primary mode of lived presence of *affection*, one would think, it is necessary to distinguish and contrast it from the consciousness in which there is an 'I' directing its attentional gaze at some object *x*, i.e., a consciousness where the 'I' is in *action*. Husserl indeed formulates a twofold task along these lines:

Departing from that [i.e., asking back in which kind of consciousness our experience of the world as valid is grounded, P.S.] the system of tasks is prefigured. 1) We have to engage in an analysis of the primary mode of lived presence and have to learn to understand it in its process of transformation (*Wandlung*) ... , the primordial temporalization, in which the I-alien hyletic

---

<sup>710</sup> Zahavi acknowledges that Husserl considers feelings to be involved even on the most basic levels of experience but does not further elaborate on that (1999, 125 Fn. 91). Generally, the role of feeling in the most basic processes of experience in Husserl is only sparsely discussed in Husserlian scholarship. But see, for instance, Lee 1993, Lotz 2007, Bower 2014.

<sup>711</sup> Husserl 2006 (Hua Mat. 8), 130, 297, 346, 350, 351 et passim.

quasi-world has its pre-being; 2) then the I, for which this pre-world is and through or in whose functioning in affection and action world in the strict sense (*eigentlich*) comes into being in the first place (*Schöpfung*), in a plurality of levels of constitution (*Schöpfung*), to each of which relative worlds correspond.<sup>712</sup>

The idea thus is that two experiential spheres can—though abstractively—be distinguished. The sphere of *affection* (the ‘I’ is affected by  $x$ ) and the sphere of *action*<sup>713</sup> (the ‘I’ attends  $x$  in an intentional act). The former corresponds to the primary way of experiential content insofar as it is given in the temporally unfolding stream of consciousness, and which is presupposed by the latter insofar as the ‘I’ can only attend something that has previously affected the ‘I’. In the same manuscript, however, Husserl immediately concedes that such a distinction runs the risk of committing “the error of being *too* abstract”<sup>714</sup>.

For, if one were to think affection and action as two separate spheres, it is throughout impossible to conceive of how the intentional act of the ‘I’ ought to reach the intentional content. But even describing it like that would already pose the problem that the intentional activity of the ‘I’ as such cannot possibly be isolated from the content. The intentional action of the ‘I’ without the content amounts to nothing nor is the ‘I’ anything for itself without experiential content.<sup>715</sup> Moreover, the kind of givenness at stake in affection does not completely lie outside of the intentional structure, even though it is itself not yet the attentional directedness of the ‘I’ towards an object  $x$  in terms of an intentional act. If it were to lie outside and to be non-intentional throughout, it could never give rise to the motivation of the ‘I’ to attend it.

Somehow, and that is the important point, which makes Husserl’s analysis so interesting, we have to think of the consciousness of affection of one that becomes *modified* when it turns into an intentional act of the ‘I’. Instead of conceiving of affection and action of two spheres that need to be bridged by the ‘I’, the latter—*action*—is a modification of the former—*affection*. For Husserl, that means that affection already involves the form of “the whole intentionality and the being-for-us of the object”<sup>716</sup>. And that, most importantly, implies that even on the level of affection we have a participation of the ‘I’, though not yet

---

<sup>712</sup> Husserl 2006 (Hua Mat. 8), 350, my translation: “Von da aus ist das Aufgabensystem vorgezeichnet. 1) Wir müssen in der urmodalen Gegenwart uns zunächst umtun und sie in ... Wandlung verstehen lernen, der urmodalen nicht-ichlichen Wandlung, der Urzeitigung, in der eine ichfremde hyletische Quasi-Welt ihr Vor-Sein hat, 2) dann das Ich, für das diese Vor-Welt ist und durch dessen oder in dessen Funktionieren in Affektion und Aktion die eigentliche Welt zur Schöpfung kommt, in einer Vielheit von Schöpfungsstufen, denen relative Welten entsprechen.”

<sup>713</sup> Here, “action” only refers to the activity of the ‘I’ insofar as it actively directs its attentional gaze at  $x$ .

<sup>714</sup> Husserl 2006 (Hua Mat. 8), 350, my translation and emphasis: “[Man gerät so leicht in] den Fehler, *zu* abstrakt zu sein.”

<sup>715</sup> Husserl 2006 (Hua Mat. 8), 351f.

<sup>716</sup> Husserl 2006 (Hua Mat. 8), 344f., my translation: “in der Weise der ganzen Intentionalität und des Für-uns-Seins des Gegenstandes”.

active in terms of an attentional and intentional directedness at  $x$ . For the point is, according to Husserl, if something affects consciousness, there is an instinctive striving towards more givenness of  $x$ , that what grounds the motivation for any attentional turn towards  $x$ . The question, for Husserl, then arises as to whether the striving that precedes attention could possibly be non-egological:

Is it not that all striving is already egological? But how can striving be egological, if the 'I', as it is the case in the primary mode of hyletic presence, is not involved? Or are we dealing with different modes of 'being involved'?<sup>717</sup>

Husserl recognizes the problem that if he considers the 'I' to show up in the strict sense only when there is a pronounced attentional activity of the 'I', then it seems prior to attention, i.e., in the primary mode of hyletic presence of affection, the 'I' does not participate. But if that was the case, then it would be unclear how affection could eventually trigger the attention of the 'I'. And that is the reason why Husserl starts to consider whether there couldn't be different modes in which the 'I' can be involved in consciousness.

How could the 'I' be involved in affection, if it is not by attentional and intentional activity (which would already be an action)? Wherein lies the egological character of pre-attentional as well as pre-reflective experience? When Husserl tries to answer that question, *that* is the moment in which he brings in the notion of feeling:

To this belongs the question of salience and the role of feeling, the feeling in the unity of primal impression (the total) as mood, but in the way of the constitution through streaming with its forms of invariation and variation.<sup>718</sup>

Feeling, Husserl considers here, could be just the way in which the 'I' is involved even in affection before any attentional action of the 'I' is undertaken. But how? Crucially, Husserl rules out that feeling is some experiential *content* alongside that what is affecting consciousness, i.e., what is given in what Husserl calls primal impression: "But are these feelings only contents of higher order, feeling only founded contents? ... Doesn't this go into the infinite? By contrast, let us take a closer look on the issue."<sup>719</sup> The problem Husserl sees with understanding the feelings at stake as part of the content of experience is that it seems that the founding content (=hyletic affection) and the founded content (=feeling)

---

<sup>717</sup> Husserl 2006 (Hua Mat. 8), 351, my translation: "Ist nicht alles Streben schon ichlich? Aber wie kann ein Streben ichlich sein, wenn das Ich, wie bei den Hintergründen der urmodalen hyletischen Gegenwart, gar nicht dabei ist? Oder handelt es sich um verschiedene Modi des 'Dabei'?"

<sup>718</sup> Husserl 2006 (Hua Mat. 8), 351, my translation: "Die hierhergehörige Frage der Abhebung und der Bedeutung der Gefühle und das Gefühl in der Einheit der Urimpression (der totalen) als Stimmung, aber in der Weise der strömenden Konstitution mit ihren Gestalten der Unveränderung und Veränderung."

<sup>719</sup> Husserl 2006 (Hua Mat. 8), 351, my translation: "Sind diese Gefühle aber nur Inhalte höherer Stufe, Gefühle bloß fundierte Inhalte? ... Geht das nicht gar ins Unendliche fort? Indessen, sehen wir uns die Sache näher an."

ought to find yet another configuration that should present a further content—et ad infinitum. If the content of affection  $p$  determines feeling, and if the feeling so-determined is itself a sort of content  $q$ , then  $q$  qua content should determine a further feeling  $q + l$ . Such an infinite regress of content cannot be tolerated and would overstretch the limits of lived presence. It would render it impossible.

Most importantly, however, for Husserl this does *not* mean that feeling could not form part of the primary mode of lived presence. Rather, he concludes from that, that experiential content, the Not-I, has to be distinguished from feeling which is egological: “Content is the Not-I, feeling is already egological.”<sup>720</sup> It is by feeling that the ‘I’ is involved in the affection, a “feeling being-involved of the I”<sup>721</sup> prior to attention: “[I]nseparable is the I and its I-alien, in every content in the unity of content and in the whole unity *is* the I *as feeling*. Feeling is the state of the I prior to all activity, and when it is active, in the activity.”<sup>722</sup> Hence, feeling provides for Husserl a way to think how the ‘I’ can be said to already be involved and participating in the consciousness of affection, even though it is not yet active in the sense of a pronounced intentional act. Crucially, that means, that for Husserl already in affection we need to distinguish between hyletic content (‘Not-I’) and feeling character of the hyletic content (‘I’), which means that the ‘I’ as a pole of experience is already manifest in affection, even though it has not yet become explicit through a thematizing reflection.<sup>723</sup> Although in affection, in the primary mode of lived presence, I and Not-I can only be distinguished through abstraction, Husserl emphasizes that “the unity of hyle”, i.e., the unity of the content, and “the unity of feeling”, i.e., of the ‘I’, need to be distinguished.<sup>724</sup>

It is important to note though that in spite of highlighting that “to the streaming presence ... belongs the abiding structure of hyle and hyle in characters of feeling”<sup>725</sup>,

---

<sup>720</sup> Husserl 2006 (Hua Mat. 8), 351, my translation: “Das Inhaltliche ist das Ichfremde, das Gefühl ist schon ichlich.”

<sup>721</sup> Husserl 2006 (Hua Mat. 8), 351, my translation: “fühlendes Dabei-Sein des Ich”. The notion of feeling as the way in which the ‘I’ participates (*Dabei-Sein*) in the experience of affection can also be found in Lotze who speaks of “participation” (“*Antheilnahme*”): “Until now, we have considered sensations only as plays in which the mind remains an indifferent witness. But they are not always like that and rather elicit often feelings, that is, those sensations, in which the witnessing mind at the same time takes a part that is characterized by pleasure or pain.”, my translation: “Wir haben die Empfindungen bisher nur als Schauspiele betrachtet, für welche der Geist ein gleichgültiger Zuschauer bleibt. Sie sind dies nicht immer, sondern veranlassen häufig Gefühle, d.h. solche Empfindungen, an denen der zuschauende Geist zugleich einen durch Lust oder Unlust charakterisierten Antheil nimmt.”

<sup>722</sup> Husserl 2006 (Hua Mat. 8), 352, my translation and emphasis: “[U]ntrennbar ist Ich und sein Ichfremdes, bei jedem Inhalt im Inhaltzusammenhang und bei dem ganzen Zusammenhang *ist das Ich fühlendes*. Fühlen ist die Zuständigkeit des Ich vor aller Aktivität und, wenn es aktiv ist, in der Aktivität.”

<sup>723</sup> Cf. Husserl 2006 (Hua Mat. 8), 352.

<sup>724</sup> Cf. Husserl 2006 (Hua Mat. 8), 352, my translation: “Wir unterscheiden Einheit der Hyle und Einheit des Gefühls.”

<sup>725</sup> Husserl 2006 (Hua Mat. 8), 352, my translation: “Zur strömenden Gegenwart ... gehört nun die beständige Struktur Hyle und Hyle in Gefühlscharakteren.”

Husserl is concerned about the possibility to sufficiently show that feeling is indeed part of the primary mode of lived presence. Husserl mentions the problem that the primary mode of lived presence cannot be recalled and investigated in reflection as it is the case with intentional consciousness, wherefore the analysis of the primary mode of lived presence always has the character of a construction that remains methodologically problematic.<sup>726</sup> Despite this general problem, Husserl formulates the task of showing how exactly feeling could be thought of as a structural moment of the primary mode of lived presence—a task he has not accomplished though.

Yet, even his deliberations as I have presented them provide us with some insights, which can be summed up as follows: (1) Somehow the ‘I’ must be involved in the primary mode of lived presence which is egological already, even though the ego is not yet explicit. (2) Feeling is egological and presents the way in which the ‘I’ is already non-explicitly manifest in affection. (3) Feeling is not a specific content of consciousness or complex of contents but rather the character in which content is given.

Husserl’s analysis reflects the common distinction, already propagated by Lotze and wide-spread in German philosophy and psychology in the late nineteenth and early twentieth century: the distinction between sensation (*Empfindung*) and feeling (*Gefühl*). Crucially, this distinction corresponds to the notion that sensation refers to that what is given in perception as the *object*, whereas feeling refers to that what is *subjective* in experience, i.e., the ‘I’.<sup>727</sup> In his essay “Das Selbstbewusstsein. Empfindung und Gefühl” (1901), Theodor Lipps defends the view that it is feeling that primarily provides self-consciousness and to which all kinds of self-consciousness trace back. One major task for Lipps is to show that feelings cannot be reduced to sensations nor a complex of sensations, as surely, that would amount to the view that the ‘I’ qua feeling is some object in consciousness on a par with world objects. In his *Phänomenologie des Ich in ihren Grundproblemen* (1910), Konstantin Oesterreich picks up on this task and develops a defense of the view that feeling—insofar as it concerns the ‘I’—needs to be distinguished from any sensation and sensory phenomenology.<sup>728</sup>

But in all these accounts a more positive answer to the question what feeling phenomenologically consists of is wanting. That applies similarly to Husserl’s account

---

<sup>726</sup> Cf. Husserl 2006 (Hua Mat. 8), 352f.

<sup>727</sup> Cf. Lotze 1846, 190, 1852, Ch. 2 § 20, 1889/1881, 48, Wundt 1874, Sigwart 1881, Ziegler 1893, Dilthey 1894, Lipps 1901, 1902, Maier 1908, Oesterreich 1910.

<sup>728</sup> A related but not congruent debate is the discussion on the nature of emotions. In the 20<sup>th</sup> century and the last decades, two general approaches (and mix-forms of them) have dominated: (1) the feeling or sensualistic theory of emotion, (2) the cognitive theory of emotion. More recently, attempts have been made to show that emotions can neither be reduced to sensations nor cognition, and that, accordingly, emotional phenomenology is *sui generis* (e.g., Montague 2016, Ch. 9, Dewalque 2017, Slaby 2017). Without having to make any final judgement on the alleged *sui generis* phenomenological nature of feeling (considered as *Gefühl* distinguished from *Empfindung*) here, what I do want to suggest is that feeling needs to be distinguished from sensation (*Empfindung*), Husserl’s affection, and cognition.



which leaves open the question how the feeling character of affection can be described in a way so as to make it compatible with the consciousness of affection and sensation. In contrast to his predecessors, however, when distinguishing feeling from sensation (or hyle/impression), Husserl would have more options to positively describe feeling as a specific form of givenness. For, we have to bear in mind that the primary mode of lived presence, according to Husserl, consists in the experiential field of retention-impression-protention. Thus, Husserl's transformation of the claim (a) that feeling is to be distinguished from sensation into the claim (b) that feeling is not a hyletic or impressional datum, *ex negativo*, motivates the idea that feeling might be better described in terms of retentional or protentional consciousness, the other two moments of the Husserlian experiential field. That means, by—wrongly—attempting to consider feeling as an impression or impressional moment, Husserl has not yet exceeded all possibilities to integrate feeling in the temporal field of experience. As far as I know, Husserl did not further address this task of which he was nonetheless implicitly and vaguely aware.<sup>729</sup> However, we do find insights in that regard in other works of phenomenological philosophy, most recently in Matthew Ratcliffe's concept of "existential feelings"<sup>730</sup>. Or so I want to suggest.

#### 6.4.2 Ratcliffe's existential feelings: Protention, possibility, and being

Imagine you hear about a small competition in the company you work for. Those who score in a certain range in a sort of game will receive an award or even a promotion. Now consider someone being exhausted after weeks of overwork, lack of sleep and wanting gratification through his colleagues and supervisors. Compare her situation with her colleague who just came back from holidays. Her supervisor had allowed her the week off after the closure of a deal that was the product of a lot of work done in the context of fruitful collaboration involving mutual respect. Although both persons might have the very same capabilities, likely they will perceive the competition in a fairly different way. The former might perceive the challenge as just yet another occasion for frustration and hard time. By contrast, the latter with the embodied self-confidence stemming from a combination of a successful performance, treat, and relaxation, most likely will find herself developing an intrinsic motivation to enter in the competition and perceive it as a welcome challenge, an invitation to unfold and reconfirm one's own vitality.

Clearly, both employees hear about the competition in a different individual context as their existential conditions diverge significantly, wherefore the experience of the competition is completely different too. According to Matthew Ratcliffe, inspired by

---

<sup>729</sup> Cf. Husserl 2006 (Hua Mat. 8), 354.

<sup>730</sup> Ratcliffe 2005, 2008, 2012, 2016, 2020.

Martin Heidegger, one's personal existential condition that shapes one's experience of  $x$  is not simply something *underlying* the experience of  $x$ , but rather something that is disclosed to us. It forms part of the experience of  $x$  and is thus experiential itself. On Ratcliffe's view, it is through a certain class of feeling that we primordially become aware of our existential condition: *existential feelings*. Ratcliffe defines existential feelings as "constituting ways of finding oneself in a world that shape more specific experiences"<sup>731</sup>. Neither being at the center of one's attention nor directed at a certain aspect of the world, they "constitute the structure of one's relationship with the world as a whole"<sup>732</sup>. As such they form "ordinarily part of the background structure of experience"<sup>733</sup>.

Hence, Ratcliffe ascribes a fundamental role to these feelings, which makes them candidates for a closer examination that might provide some insights as to how feeling could be involved even in the primary mode of lived presence. In fact, there are more specific reasons why existential feelings lend themselves to such an analysis, which becomes evident when one considers the key features of existential feelings.

*1) Existential feelings are neither mere affect nor directed at a certain object.*

One important characteristic of existential feelings is that, unlike emotions such as anger, they do not refer to specific objects. While anger is usually directed at a certain circumstance in terms of 'feeling angry that  $p$ ', existential feelings are not intentional in the same way. However, neither are they non-intentional throughout, as Ratcliffe considers them "part of the structure of intentionality"<sup>734</sup>. As such, they do not present 'mere affect' either, as if they were only an effect the world has on a subject. Instead, Ratcliffe emphasizes, existential feelings disclose the world as a whole and *thereby* allow for any intentional directedness at specific objects in the world to begin with. Rather than simply calling them non-intentional, Ratcliffe suggests existential feelings are "pre-intentional"<sup>735</sup>.

*2) Existential feelings do not present a certain class of acts.*

Given that they are described as pre-intentional, unlike intentional emotional experience, existential feelings are not understood in terms of a specific class of intentional acts to be contrasted with, for instance, perception, recollection or imagination. As pre-intentional, they do not form a *sui generis* act-class and rather precede the intentional activity of the

---

<sup>731</sup> Ratcliffe 2005, 53.

<sup>732</sup> Ratcliffe 2005, 61.

<sup>733</sup> Ratcliffe 2005, 53.

<sup>734</sup> Ratcliffe 2008, 36.

<sup>735</sup> Ratcliffe 2010, 604, 608, 2012, 10. Cf. Strasser 1977, Ch. 1. Some defend existential feelings are intentional states though. Cf. McLaughlin 2009.

ego. Hence, neither do existential feelings have a specific object nor do they necessarily involve the ego qua intentionally active.

3) *Existential feelings are experiential though not perceptions.*

To say existential feelings are not a sui generis act-class, does not mean they are not experiential. But that poses the question of their givenness. In which sense are they felt? Are they perceptions then? Not quite. Ratcliffe is adamant that existential feelings are bodily feelings but he highlights they are not feelings *of* the body.<sup>736</sup> They are not simply sensations the sum of which amounts to some sort of mental image of our body that we have when we were to reflect on how exactly our body feels at a given point in time. They are also to be distinguished from physiological affect, as Ratcliffe aims to clarify by describing the experience of a numb hand. Rather than feeling simply nothing, feeling a numb hand is experiencing the absence. Hence, bodily feeling is not restricted to feeling the body insofar as it is a material object in touch with other extended objects. The feeling body is more than the experience of a physiological body.<sup>737</sup> Moreover, as he argues on another occasion, what we experience through our senses, i.e., what we experience physiologically, can remain unaltered, while our feeling may well change significantly. A well-known example is “delusional ‘atmosphere’ or ‘mood’ that sometimes precedes full-blown schizophrenia”<sup>738</sup>. In these experiences, a sudden and dramatic alteration of one’s over-all experience may occur without any change in one’s sensory perception. Such an existential change, nonetheless, is clearly felt. Hence, Ratcliffe suggests, existential feelings are not to be equated with sensory perception.

4) *Existential feelings are experiences of possibility and part of perceptual experience.*

However, refraining from identifying existential feeling with sensory perception does not mean to construe existential feeling as something detached from perceptual experience either. Quite the contrary, Ratcliffe highlights that existential feelings are intrinsic to perceptual experience, given that they are precisely the experiential background for any one thing to be a possible object for our perception to begin with. But how are they part of perceptual experience? Crucially, Ratcliffe points to the experience of possibility. On his view, and appealing to Husserl and other phenomenologists, whenever we perceive something, our experience is not exhausted by that what is given in terms of that what is sensorically given in actuality. Referring to the notorious cup, he gives an example of an experience of possibility: “I experience the cup as something I could see from another angle

---

<sup>736</sup> Ratcliffe 2008, 38.

<sup>737</sup> Cf. Ratcliffe 2008, 56.

<sup>738</sup> Ratcliffe 2014, 41.

by *doing x* or as something that I could touch *by doing y*.”<sup>739</sup> The hidden side, though not given to me in a strict sense, is still experienced as something that I *could* see if I turned around the table. Another example he provides is a sharp-looking knife in the kitchen the experience of which includes the experience of the possibility that I cut myself.<sup>740</sup> Crucially, pointing to Husserl’s notion of *horizon*, Ratcliffe emphasizes these possibilities are not something “inferred from what is actually perceived”, but “integral to the experience”<sup>741</sup>. Furthermore, the possibilities that center around this or that object are usually not perceived in a neutral way, they are enticing or repellent, something to long for or to avoid.<sup>742</sup> In other words, we feel about them in a certain way. It is in *this* sense that Ratcliffe understands existential feelings as experiences of possibility.

5) *Existential feelings convey a sense of reality.*

Hence, Ratcliffe can say that existential feelings open up a “space of experiential possibilities”<sup>743</sup> to us, they disclose different potential outcomes of a given situation and thereby provide a sense of our reality. What we experience as possible shapes how we feel and what we take our situation to be. Et vice versa—depending on how we feel ourselves this or that appears as possible to us. Surely, as Ratcliffe further elaborates, these possibilities come in different modes, as that what is possible can be *likely*, *unlikely*, *merely theoretically possible* or even *certain* and so forth.<sup>744</sup> Now, importantly, it makes a difference in which mode a possibility is given for how it feels. Consider the above-mentioned example of the competition. Both employees are aware of the brute fact that they could win the challenge, but the mode in which this possibility is experienced differs. This difference is a difference in existential feeling.

6) *Existential feelings convey a sense of belonging to the world.*

However, existential feelings do not only disclose the world to us—as if it were something completely independent and detached from us. Rather, they equally and inextricably reveal how *we* are situated in the world. They are similarly feelings about *our* existence and not about the general *conditio humana*, as it were. As such, they tell us who we take ourselves to be, prior to any reflection. It is through existential feelings then that we develop a “sense

---

<sup>739</sup> Ratcliffe 2014, 43. See also Ratcliffe 2008, 130.

<sup>740</sup> Ratcliffe 2014, 43.

<sup>741</sup> Ratcliffe 2008, 130.

<sup>742</sup> Ratcliffe 2008, 132.

<sup>743</sup> Ratcliffe 2005, 56.

<sup>744</sup> Cf. Ratcliffe 2008, 132ff., 2014, 43ff.

of belonging to the world<sup>745</sup> and of being involved in its events.<sup>746</sup> It is therefore that Ratcliffe can say: “Any sense of self that we have is grounded in existential feeling, even though it might not be exhausted by it.”<sup>747</sup>

#### 7) *Existential feelings are pervasive.*

Ratcliffe highlights that existential feelings can vary to a high degree and that they can change quickly or persist even over a life time when habitualized. In any case, however, as the general background of experience some existential feeling always forms part of the experience: “all experience is structured by some variant of existential feeling”<sup>748</sup>.

### 6.4.3 Existential feelings and the stream of consciousness

As these seven key characteristics show, the conception of ‘existential feeling’ presents a proposal regarding the question how we can conceive of feeling to be involved even in the pre-intentional, pre-reflective primary mode of lived presence that Husserl envisaged in his later manuscripts and which correspond to the phenomenological structure of the minimal self: inner time-consciousness. Following Ratcliffe, I want to suggest it is the *experience of possibility* that conveys not only a sense of being presupposed by any world to emerge in the first place but also presents the way in which the ego, though still throughout unthematically, participates even in the most basic experience of affection. Thus, my proposal is to understand the feeling participation of the ‘I’ (*fühlendes Dabeisein des Ich*) in affection in terms of protentional consciousness. For, what is protended, that is, protentionally experienced to possibly happen just *is* the way in which the concrete world as a whole appears and unfolds. If our experience was limited to and exhausted by what is actually present in impressional consciousness, no world could ever be experienced. And, what our world is like depends on what is possible in it. To experience world is not to witness the mere fact of a number of objects standing around, in terms of Heideggerian *Vorhandenheit*, as it were, but to experience possible events and happenings. But crucially, these events are not merely events of a detached and separate world. Experiencing possibilities always also means to experience one’s *own* possibilities, as any protended event is the protention of the same conscious process to proceed. That, I suggest, is what Husserl means when he says:

---

<sup>745</sup> Ratcliffe 2008, 39.

<sup>746</sup> Ratcliffe 2005, 53f., Ratcliffe 2008, 39, 121.

<sup>747</sup> Ratcliffe 2008, 121.

<sup>748</sup> Ratcliffe 2008, 40.

Originally living [in the permanent streaming] I am ‘conscious’ of myself as present, past and future I-live ... . But in this life, I am conscious of the world ... .<sup>749</sup>

In every presence taken as a phase, and in permanent on-going presence, I am in such a way that I transcend my present being. I am real and concrete as a permanent presence; that is my concrete being. But it is concrete streaming, as that it contains the continuity of the intentional modifications of the present primary mode of now.<sup>750</sup>

The self-experience that grounds any explicit egological consciousness of the ‘I’ is originally and always an experience of a streaming that implies a transcendence of what is actually given. When I see the front side of the cup on the table, I simultaneously experience the possibility of seeing its back side. Crucially, it is experienced as a possibility for *me*, the consciousness that is actually seeing the front side. Insofar as consciousness involves an experience of the possible, it transcends itself towards a possible future. In the primary mode of lived presence, such a possible future is not merely experienced as possible, rather, motivated by what is actually given, consciousness protentionally anticipates such a possible future *for itself*. Hearing the tone of a melody there is a protentional anticipation of the resounding of a further tone—be it a new phase of the currently given tone or another tone. But as such, consciousness transcends itself towards its *own* future. As Nicholas de Warren puts the point:

[T]he consciousness of any possible future involves an implicit consciousness of myself as ‘there’ in the future ... . [M]y consciousness anticipates its own temporal self-constitution and anticipates *itself* as the source of any possible future constitution of experience, including its own.<sup>751</sup>

Protentional consciousness, then, is an indispensable component of what forms the horizon, which involves not only the possibilities concerning the givenness of this or that object, but always and inextricably connected also the possibilities of myself, the consciousness undergoing the phase of experience at stake. Taking this notion seriously, we could then say, it is protentional consciousness through which not only the horizon of the world emerges, but correlatively, though not yet explicit and thematized, the egological horizon, the anticipation of my own continuity and persistence. The ‘I’, thus, prior to any

---

<sup>749</sup> Husserl 2006 (Hua Mat. 8), 120, my translation: “Urtümlich lebend [im stehend Strömen] bin ich meiner ‘bewußt’ als gegenwärtiges, vergangenes und künftiges Ich-lebe ... . In diesem Leben bin ich aber bewußt der Welt ... .”

<sup>750</sup> Husserl 2006 (Hua Mat. 8), 129, my translation: “In jeder Gegenwart als Phase genommen, und so in stehender, fortwährender Gegenwart, bin ich so, dass ich mein gegenwärtiges Sein transzendiere. Wirklich und konkret bin ich als ständige Gegenwart, das ist mein konkretes Sein. Aber es ist konkretes Strömen, als das enthält es die Kontinuität der intentionalen Modifikation des momentanen Urmodus Jetzt.” See also Husserl 2006 (Hua Mat. 8), 130, 149, 150, 152, 254.

<sup>751</sup> De Warren 2009, 197f.

pronounced intentional activity, already participates even in the consciousness of affection by way of a protentional self-anticipation, which is at the same time the protentional anticipation of possible world events. World-horizon and egological horizon, when considered in terms of the primary mode of lived presence, are one and the same. It is also in this sense that Husserl can distinguish: “The subjective as affection, as actus, as occupation with the ontic” and “the subjective as feeling, as mood, as universal horizon-like ‘feeling of life’”<sup>752</sup>.

Hence, inspired by Ratcliffe’s concept of existential feelings, I want to suggest we can find a way to understand how an egological feeling is involved in the consciousness of affection, the primary mode of lived presence: as an experience of possibility in terms of protentional consciousness.

This, I propose, can further be illustrated by Husserl’s notion of *interest* that he takes to be the essential and grounding basis of experience inherent to the most original form of affection. This basis consists in a general and “all-founding interest”<sup>753</sup>, a “general interest in objects as such”<sup>754</sup>. More precisely, Husserl sees in the phenomenon of affection, the being presented of something, a fundamental, i.e., instinctive, experience of “curiosity (*Neugier*)”<sup>755</sup>. Being presented with something, the stream of consciousness unfolds temporally and an object is passively constituted, though, as Husserl emphasizes, this involves a kinesthetic ‘I follow’ a bodily orientation in that what is affecting consciousness.<sup>756</sup> This, according to Husserl, already presents an experience of interest, which is not constituted in the specific characteristics of the object given, but in the affection of something as such: “The constitution of beings, of objects (that exist or are in modes of being consciously given) is nothing else than the constitution of interests. World-pregivenness means: A universal interest is founded and thereby the individual form of all interests.”<sup>757</sup> Prior to any evaluation of an object as pleasurable or painful, as worthy, interesting (in the strict sense) or neutral, according to Husserl, the mere being affected-by is imbued with a genuine interest, which presents the general structure of interest that is modified in any specific interest in *x*: “The most basic, all-founding interest, thus, is that of the original and abiding

---

<sup>752</sup> Husserl 2006 (Hua Mat. 8), 363, my translation: “Das Subjektive als Affektion, als Aktus, als Beschäftigung mit Ontischem, das Subjektive als Gefühl, als Stimmung, als universal horizonhaftes ‘Lebensgefühl’.”

<sup>753</sup> Husserl 2006 (Hua Mat. 8), 325, my translation: “allfundierendes Interesse”.

<sup>754</sup> Husserl 2006 (Hua Mat. 8), 331, my translation: “allgemeines Interesse für Gegenstände als solche”.

<sup>755</sup> Husserl 2006 (Hua Mat. 8), 323.

<sup>756</sup> Husserl 2006 (Hua Mat. 8), 323.

<sup>757</sup> Husserl 2006 (Hua Mat. 8), 74, my translation: “Die Konstitution von Seiendem, von Gegenständen (die sind oder in Seinsmodalitäten bewusst sind) ist nichts anderes als Konstitution von Interessen. Weltvorgegebenheit besagt: Ein universales Interesse ist gestiftet und hinfert individuelle Form aller Interessen.”

functioning curiosity (*fungierende Neugier*), or better said, the experiencing and, indeed taken as most basic, the sensorical experiencing interest.”<sup>758</sup>

Crucially, although Husserl highlights that it is the affection through something that triggers a fundamental curiosity and striving towards the temporal unfolding constitution of the experienced *object*, he does not restrict the interest to the interest in objectivity: “The being-consciously-given-to-me in the most general sense is affecting-me, and pronounced ‘consciousness’ is being-as-participating, being-actively-interested.”<sup>759</sup> Interest in objectivity is a ‘being-actively-interested’ and therefore involves egological participation. It is the interest in affection that drags me in and lets me participate. And, therefore, Husserl also says the curiosity taken as “a feeling of pleasure” is a “pleasure of being-as-participating”<sup>760</sup>. Hence, the interest in the experience of the object through which one is affected is not only a pleasure of ‘being affected by *x*’, but also a pleasure of ‘being-as-participating in being affected by *x*’. The all-founding interest, thus, is also an interest in oneself, the experiential process, as such. Accordingly, Husserl can say that the “instinctive pleasure of fulfilment does not concern a final state, but the whole process”<sup>761</sup>. It is the “unity of the process of intention-fulfilment” which is the “telos” that is being fulfilled in the “constant activity”<sup>762</sup> of intentional consciousness. In having consciousness of *x*, a general interest is constantly fulfilled, namely the on-going of experience, the ‘unity of the process of intention-fulfilment’—involving the self-experience of the experiential process. For, the interest in objectivity as such is, by the same token, precisely an interest in the experiential process or subjectivity, given that it is the subjective experiential process in which any object is given. Striving towards object-giveness is striving towards givenness, i.e., experience. “Life is always already a life of interest (*Interessenleben*)”<sup>763</sup> which translates into an “interest in life (*Lebensinteresse*) as the horizon in which the current activities are held together”<sup>764</sup>. Husserl emphasizes that “the horizon of being is a horizon

---

<sup>758</sup> Husserl 2006 (Hua Mat. 8), 325, my translation: “Das unterste, allfundierende Interesse ist also das der ursprünglichen und immer weiter fungierenden Neugier, oder wir sagen besser, das erfahrende und, in der Tat zuunterst genommen, das sinnliche erfahrende Interesse.”

<sup>759</sup> Husserl 2006 (Hua Mat. 8), 323f., my translation: “Das Mir-bewusst-Sein im weitesten Sinne ist Mich-Affizieren, und ausgezeichnetes ‘Bewusstsein’ ist Dabei-Sein, Aktiv-interessiert-Sein.”

<sup>760</sup> Husserl 2006 (Hua Mat. 8), 324, my translation: “Nennen wir ‘Neugier’ ein Lustgefühl, so ist diese Lust am Dabei-Sein sozusagen im Stand des Genusses, hier aber in sich die fortgehende Affektion.”

<sup>761</sup> Husserl 2006 (Hua Mat. 8), 328, my translation: “Die instinktive Intention und instinktive Lust der Erfüllung betrifft nicht einen Endzustand, sondern den ganzen Prozess ... .”

<sup>762</sup> Husserl 2006 (Hua Mat. 8), 328, my translation of parts of the following sentence: “also die Einheit des Prozesses der Intention-Erfüllung, das ist selbst das Telos, das ist, dass sich die instinktive Intention, die einheitlich von vornherein auf dieses Ineinander der Intentionalität und ihrer Entspannung geht, und sich als einheitliche nicht in einer Phase, sondern im ständigen Tun erfüllt... .”

<sup>763</sup> Husserl 2006 (Hua Mat. 8), 252, my translation: “Das Leben ist immer schon Interessenleben ... .”

<sup>764</sup> Husserl 2006 (Hua Mat. 8), 75, my translation: “ein Lebensinteresse als Horizont, in dem sich die momentanen Tätigkeiten halten”.



of interest”, and that “the single acts are not separated”, but “are connected and hinge together in the unity of a striving life”<sup>765</sup>.

Although Husserl does not explicitly link the notion of the *feeling participation of the ‘I’* in affection with the notion of *being-actively-interested* in affection, and neither with the notion of *protentional consciousness*, I think the concept of *interest* as discussed harbors their intrinsic connection. For, interest implies a certain reference to value that, as experience, is best described as feeling rather than in terms of an alleged neutral experience. Second, interest also implies the participation of the ‘I’ as caring about *x* in terms of being affected by *x*. And crucially, in caring about *x* consciousness cares about itself as the consciousness in which *x* is given. Third, interest is not something that is consumed in affection but rather triggered in the first place. That means, interest is immediately future-directed striving towards new phases of *x*. The all-founding interest, which Husserl calls an instinctive curiosity, in German, *Neugier*, literally means *desiring* or *striving towards the new*. Interest, qua *inter-esse*, thus, means the participating being of the ‘I’ in being affected by *x*, which is intrinsically a being that is future-directed. As Husserl puts it: “The being and life, in which I-being unfolds, comes to its being, always already having being, is being that is directed towards being ... (in terms of a form of correlation broadly construed: living is directed towards living-to-come).”<sup>766</sup> The inter-esse of interest, then, expresses the egological being as being affected by *x*, which is a being that is essentially temporally structured and stretches out towards the future that is not yet being realized. And crucially, the interest qua inter-esse is not a neutral experience but comes with a primordial feeling of value for oneself as the locus of all possible givenness, a value that is not based on certain characteristics of oneself or the objects given but is in this sense “immediate”<sup>767</sup> and absolute as Lotze emphasized. As the following quote of his *Medizinische Psychologie* further illustrates, Lotze’s proposal of self-concern and self-feeling fits well with Husserl’s phenomenology of the primordial way of lived presence (as I have depicted it):

The feelings of which we expect the foundation and fixation of self-consciousness are not singular and salient feelings of high intensity. Rather, it is the uninterrupted and steady stream of participation in our own experiential being (*Teilnahme an uns selbst*), in which the subtlest promptings of feeling

---

<sup>765</sup> Husserl 2006 (Hua Mat. 8), 35, my translation: “Der Seinshorizont ist Interessenhorizont, und die Einzelakte sind nicht zusammenhangslos, ... , sondern hängen verkettet zusammen zur Einheit eines strebenden Lebens ... .”

<sup>766</sup> Husserl 2006 (Hua Mat. 8), 18f., my translation: “Das Sein und Leben, worin Ich-Sein statthat, zu seinem Sein kommt und immer schon Sein habend, ist Sein, das auf Sein vorgerichtet ist ... (in einem uneigentlichen Korrelativsein: Leben ist auf Leben-werden gerichtet).”

<sup>767</sup> Lotze 1884/1856, 280.

that are always associated with sensations (*Empfindungen*) and thoughts (*Vorstellungen*) drag us in.<sup>768</sup>

This sounds all good and solemn, but, wait a minute, some might wonder, wouldn't such an account of basic felt self-interest mean that just any consciousness of *x*, any experiential state of ours, would come with some kind of feeling of pleasure or pain, as Lotze also seems to maintain?<sup>769</sup> And is that not a bit far-fetched, given that we surely have also quite neutral experiences that although having a certain what-it-is-likeness do not carry a strong aspect of felt value? Is it true that it is the stream of consciousness as such that we care about? Or is James right when he emphasizes that it is not the “stream of thought”<sup>770</sup> that would give rise to the “*direct feeling of regard for*” one’s “*own pure principle of individual existence*”<sup>771</sup>? As he further says, “[t]o have a self that I can *care for*, nature must first present me with some *object* interesting enough to make me instinctively wish to appropriate it for its own sake . . . .”<sup>772</sup> And doesn't that imply that only specific objects with certain characteristics entice us in a relevant way so as to elicit felt interest? Does such a view not fit better with the idea that “affect-free mental episodes”<sup>773</sup> exist, as Textor objects against Lotze’s theory of immediate self-concern and self-feeling? Let me explore the phenomenology of allegedly neutral experiences further and assess to what extent and how exactly they might present an obstacle for the theory of feeling as part of the primordial way of lived presence offered here.

#### 6.4.4 On the phenomenology of ‘neutral’ experiences

I discuss allegedly ‘neutral’ experiences from two perspectives: (a) through the lens of the Heideggerian notion of *Angst* and (b) through the lens of Lotze’s notion of feeling as an experience of pleasure or pain. Both views suggest that speech of allegedly neutral experience doesn’t go through as easily as many might think. For, whether an experience can be called neutral depends on whether we consider the context in which experiences occur, their experiential background. While some tokens of experience might be described as neutral and affect-free (if we abstract from the context), the experiential process in which they occur isn’t. This, I suggest, can be argued if we combine the two perspectives.

---

<sup>768</sup> Lotze 1852, § 37 art. 422, my translation: “Nicht einzelne hervortretende Gefühle von großer Intensität sind es, von denen wir die Begründung und Festhaltung des Selbstbewußtseins erwarten, sondern eben dieser ununterbrochene gleichmäßige Strom der Teilnahme an uns selbst, in welchen uns die stets mit Empfindungen und Vorstellungen sich verknüpfenden leisesten Regungen des Gefühls hineinziehen.”

<sup>769</sup> Cf. Lotze 1884/1856, 272, 1888/1864, 571.

<sup>770</sup> James 1950/1890, 319.

<sup>771</sup> James 1950/1890, 318.

<sup>772</sup> James 1950/1890, 319.

<sup>773</sup> Textor 2018, 168.

1) Heidegger's Angst: The 'neutrality' of 'neutral' experiences is a mode of feeling

A first possible way to respond to the challenge of alleged neutral experiences consists in discussing one particular possibility that characterizes the being of all of us: the possibility of not-being, i.e., the possibility, anticipation even, of our death. Famously, Scheler<sup>774</sup> and Heidegger<sup>775</sup> argue we do not learn, over the course of time, as it were, that the possibility of death is a feature that is characteristic of our being but that was unknown to us before we had heard about or witnessed the death of another person. Quite the contrary, they argue that the possibility of not-being is precisely built into the temporal structure of experience, something that is disclosed to us prior to any explicit knowledge and reflective experience and thus part of the primordial way of lived presence. Accordingly, they don't take it merely to be a possibility that, as a matter of fact, adheres to our being, but one that is also an *experienced* possibility. Therefore, they describe it as part of our *sense of being*.

In fact, as Heidegger emphasizes, the experience of the possibility of not-being is the experience of one's individual existence par excellence. First, it is precisely the possibility of not-being that brings afore one's own being *as a whole*, given that the possibility of one's not-being means that nothing remains possible for one. Hence, Heidegger can say that death "signifies a peculiar possibility of being in which it is absolutely a matter of the being of my own Da-sein"<sup>776</sup>. Being aware of my own possible death, I am aware of myself as a whole or totality. The wholeness as experienced relies on the completion of my being, which it finds in its end. And therefore, the consciousness of my possible end is the consciousness of my whole being. For Heidegger, this is the true *Seinsgewissheit* that lingers on the ground of our being—and not the Cartesian *cogito ergo sum*.<sup>777</sup> Second, the experienced possibility of not-being, of death, according to Heidegger, conveys a sense of self-individualization. For, one's own death is something that no one can do instead of me; no one can substitute me in the event of death. It is truly *mine* and in this sense the paradigmatic case of my very own individual possibility. The possibility of my not-being is at the same time the possibility of my being in terms of the being of my possibilities. Once the possibility of death becomes realized, I cease to exist and with it all my possibilities. The consciousness that these latter possibilities could (and will) vanish is the primordial way in which these possibilities are given to me to begin with. Whatever

---

<sup>774</sup> Scheler 1979.

<sup>775</sup> Heidegger 2006/1927, §§ 26–29, 1979, § 34.

<sup>776</sup> Heidegger 1996, 223 (2006/1927, 240).

<sup>777</sup> As he writes: "This certainty that it is myself who is in the state of coming-to-an-end is *the most basic certainty of Dasein itself* and is a real proposition of existence, whereas the cogito sum only appears to be such. ... *sum moribundus* ...—*the moribundus grounds the meaning of the sum in the first place*." My translation: "Diese Gewißheit, daß ich es selbst bin in meinen Sterbenwerden, ist *die Grundgewißheit des Daseins selbst* und ist eine echte Daseinsaussage, während das cogito sum nur der Schein einer solchen ist. ... *sum moribundus*—*das moribundus gibt dem sum allererst seinen Sinn*." (Heidegger 1979, 437)

particular possibility becomes fulfilled or disappointed, some other possibilities will remain intact, my being remains possible and with it my possibilities. But in death, my possibilities dissolve all at once. Accordingly, the possibility of all possibilities, the possibility of impossibility, unifies the set of my possibilities, because I experience them tied to and depending on that one peculiar possibility of death. *These* are the possibilities (in the world) that might remain after my passing away. My grandchildren will have children on their own long after I will have been gone. It is possible. But *here* are *my* possibilities, when I am gone, I won't be able to climb the Himalaya.

Crucially, this particular possibility of not-being that is present in any moment of *Dasein*, in any moment of possibility-consciousness, and which discloses the meaning of *my* possibilities, is arguably not experienced in a *neutral* manner. Scheler calls *Angst* the “emotional apriori of the proud ‘cogito ergo sum’”<sup>778</sup>. The same notion can be found in Heidegger:

*But the attunement which is able to hold open the constant and absolute threat to itself arising from the ownmost individualized being of Dasein is Angst. In Angst, Da-sein finds itself faced with the nothingness of the possible impossibility of its existence. ... Because the anticipation of Da-sein absolutely individualizes and lets it, in this individualizing of itself, become certain of the wholeness of its potentiality-of-being, the fundamental attunement of Angst belongs to this self-understanding of Da-sein in terms of its ground.*<sup>779</sup>

The point then is that I am not experiencing the possibility of my own impossibility and death in an indifferent manner. Some of my experienced possibilities might be given as neutral, and even as my experience unfolds I might not have specific, let alone intensive or pronounced, feelings about a particular experience. However, on Heidegger's view, there is no indifference concerning one's experiential being as the ground for all possibilities. My being is under constant threat, I can cease any moment, and I am aware of that, pre-reflectively, pre-conceptually, in the feeling of *Angst* which not only betrays the fundamental self-concern, which Heidegger acknowledges just like Lotze, but also reveals my being in the first place.

It is the consciousness of my possible death that makes me sensitive to particular possibilities and their meaning for my own being. It is against the background of *Angst*, the disclosure of my being as always already and constantly under threat, that something can appear threatening, harmless and so forth. This does not mean that all my experience is characterized by a pervasive feeling of acute fear that would be the manifestation of constant self-concern. Neither does it imply that, ontogenetically, as it were, consciousness

---

<sup>778</sup> Scheler 1979, 31, my translation of parts of the sentence: “Angst gebiert die Rechenhaftigkeit der Lebensführung und ist das emotionale Apriori des stolzen ‘cogito ergo sum’.”

<sup>779</sup> Heidegger 1996, 245 (2006/1927, 265f.).

would depart from a consciousness of the experiencer's possibility of not-being. It doesn't even necessarily imply that consciousness must involve an *explicit* awareness of the experiencer's possibility of not-being. It simply means that my experienced possibilities are always such that they have an affective or feeling valence that is determined by how they relate to the possibility of not-being. If, say in primitive and rudimentary forms of consciousness, there is an absence of an explicit awareness of the possibility of not-being, then this means that these forms of consciousness have a specific feeling character that is precisely determined by the fact that the possibility of not-being is not explicitly experienced. Now let such a consciousness 'learn' that it could cease to exist by somehow making the possibility of not-being explicit. Arguably, according to the thesis defended by Scheler and Heidegger, making the possibility of not-being explicit will ensue a change in phenomenology. Plausibly, the modification is one concerning feeling. For, being now explicitly aware of the possibility of not-being comes with a certain feeling. And now we may ask: Did making explicit the possibility of not-being introduce feeling to the sphere of consciousness in the first place? Is the modification in phenomenology best conceived of in terms of a change from non-feeling (= lack of all feeling) to feeling? Or, by contrast, and that is what I take Scheler's and Heidegger's account to be, should we regard the modification that occurs as one *of feeling*? On that view, it is a change in feeling rather than the initiation of feeling that is brought about by making explicit the possibility of not-being. And, crucially, the idea is that the state prior to making that central possibility explicit is one that is nonetheless determined by the possibility of not-being. It feels somehow not to feel in danger. Not feeling under threat doesn't amount to a lack of feeling but, on Heidegger's view, is a specific mode of *Angst*.

Neutral experiences are, against this background, misunderstood if taken as affect-free. Rather, the supposedly *neutral* experience is a mode of *Angst* insofar as it is the experience of a situation in which no acute threat is present. Yet, this does not mean that the general possibility of threat were absent. Quite the contrary, experiencing a situation in a neutral tone is relative to the excitement and upheaval of situations in which I find myself or some of my particular projects under threat. Its meaning as neutral is derived from the general possibility of threat that is, in the neutral experience, absent. This absence is not necessarily a deficient mode, but can (perhaps should) be seen as something positive. Experiences that appear neutral, lacking excitement, threat, fear, but also pronounced joy or pain, still feel like something. That the neutral experience is something positive means that it is not a lack of feeling but a feeling of a lack—of threat, of pronounced pleasure etc.<sup>780</sup>

Hence, a first response to the challenge of neutral experiences consists in the indication that neutral experiences need to be understood against the experiential

---

<sup>780</sup> Cf. Ratcliffe 2012, 72.

background of *Angst*, i.e., as a mode of feeling rather than an absence of affective meaning. Such a hermeneutic response, thus, doesn't consist in the claim that any consciousness necessarily involves a tiny bit of vague fear, as it were. The claim is that the description of allegedly neutral experiences as *neutral* is derivative of our – the thinkers who are engaging in the description – human understanding as it is determined by the sense of the possibility of not-being. The suggestion, therefore, is that we shouldn't be too quick in associating neutral experiences with an allegedly feeling-free kind of consciousness but should interpret whatever we think is best described as a 'neutral' experience in terms of a certain feeling. That is, the point is not that there are no experiences that we could refer to as 'neutral' experiences. The point is that the neutrality of these experiences should be interpreted as a certain mode of feeling rather than in terms of an absence of feeling. Let me now try to further corroborate the thesis that whatever we take to be 'neutral' experiences does in fact involve aspects that plausibly possess feeling character.

## 2) Lotze, the feeling of pleasure and pain, and the variation of 'neutral' experiences

A second response can be launched by linking Lotze's understanding of feelings of pleasure and pain, Husserl's notion of protentional consciousness and horizon, and Ratcliffe's concept of existential feeling. The suggestion I want to make in bringing these approaches together is to develop a descriptive model of the phenomenology of experience according to which all experiences are characterized by a space of possibilities that involves possibilities of pleasure and pain. Crucially, building on Lotze, what I want to argue is that, although not all experiences might present experiences of pleasure and pain, all experiences involve an experience of the *possibility* of pleasure and pain. Thus, since the experienced space of possibilities always involves pleasure and pain – no experience is fully neutral in the sense of a complete absence of feeling. This is so, I will suggest, because also the experienced *possibility* of pleasure and pain has a feeling character, not only *actual* pleasure and pain. I conclude from that that what we might characterize as 'neutral' experiences does not undermine the view that all experience involves feeling.

My strategy to further carve out the feeling character of 'neutral' experiences consists in a sort of *phenomenal contrast argument*<sup>781</sup>. I offer an investigation of a certain 'neutral' experience *x* and compare different experiential contexts in which *x* might occur, showing that *x*, depending on each experiential context, might feel different respectively. Let me develop that thought.

Lotze claims that the feeling character that lets us experience our absolute self-value consists in the experience of pleasure and pain. Construing *x* in terms of an alleged 'neutral' experience, by contrast, means to consider *x* in terms of an experience that itself is not an

---

<sup>781</sup> For a discussion of the methodological idea behind phenomenal contrast arguments, see Chudnoff 2015, Ch. 2 and Siegel 2010. See also Strawson 2010, Siewert 1998, 2011, and Kriegel 2015.

experience of pleasure or pain. Take, as a case of  $x$ , for instance, my waiting at a bus station. Waiting for the bus can easily be understood as a ‘neutral’ experience. The waiting does not entice me to do anything. I don't enjoy it, nor does it bother me in a way that would be felt. I don't even feel a longing for the bus to come. I simply wait.

Now, on a first interpretation of Lotze's thesis, such experience, given its ‘neutrality’ concerning pleasure and pain, ought to be not only bereft of feeling, it should also lack the character of mineness, because it lacks feeling.<sup>782</sup> Lotze offers an explanation for these ‘neutral’ cases when he says that in the “matured mind” it might be that some experiences do not seem to have the character of pleasure or pain attached to them, but that this is the case because the feelings of pleasure and pain are like “washed-out colors”<sup>783</sup>. Once used to certain impressions, the original pleasure or pain associated with them can lose their force to the effect that impressions or aspects of experience seem to be bereft of any character of pleasure or pain, but really, their felt intensity has simply decreased due to habituation. As Textor remarks, this, however, is not a satisfying answer, given that it would amount to an approach of feelings of pleasure and pain that are precisely not felt.<sup>784</sup>

I want to offer a second, alternative interpretation of the case of ‘neutrally’ waiting for the bus, which includes the following specifications or modifications of Lotze's proposal:

- Instead of saying that each aspect of experience (Lotze speaks of “stimulations of the soul”<sup>785</sup>) must elicit pleasure or pain to be experienced as mine, one could simply say that in each experiential episode (as a whole) to be experienced as mine, some feeling of pleasure or pain needs to be involved. That means, a total experiential episode comprising many different aspects is experienced as mine if at least *one* of them has the character of feeling. Acknowledging neutral aspects of experience does not mean that the experiential process as such is experienced as neutral. In fact, when Lotze says that some aspects of experience may seem neutral this is so because they are not relevant to a certain goal, which means that pursuing that goal comes with a certain experienced pull of interest, something I feel and care about. The neutrality of the orange trees is relative to my goal to catch the bus, their perception yet occurs within a context of interest and feeling (I dislike the idea of not catching the bus, and I hope to catch it, whereby catching it *will* make me happy).

---

<sup>782</sup> Cf. Textor 2018, 169.

<sup>783</sup> Lotze 1884/1856, 272

<sup>784</sup> Textor 2018, 169.

<sup>785</sup> Lotze 1884/1856, 272

- Rather than claiming that all contents of experiences and all partial conscious acts that together form a total experiential episode need to elicit pleasure or pain in themselves, my alternative reading of Lotze’s account suggests that all my experience is such that it *involves* aspects of pleasure and pain. It means that even experiential *episodes* that seem to be free of pleasure and pain occur in the context of the experiential *process* that involves pleasure and pain.
- Crucially, to form part of the experiential process, I suggest, pleasure or pain do not have to refer to the experiential *content* of a given point in time. On one occasion, Husserl emphasizes that “not only hyletic data themselves in their impressionality can affect, but also the hyletic anticipation of data, namely as ‘values’, as positively affecting”<sup>786</sup>. I think this is correct. Anticipating pleasurable experience may itself have already feeling character of pleasure, although it is phenomenologically distinct from the fulfilled anticipated experience of pleasure. Looking forward to eating ice cream soon, involves a feeling of pleasure although it is not yet the pleasure of eating ice cream. Put more generally, instead of claiming that the *experience of pleasure and pain* necessarily constitutes feeling, combining Lotze’s account with Ratcliffe’s notion of existential feeling, I wish to suggest *experiences of the possibility of pleasure and pain* are non-neutral experiences either and so have equally feeling character. Pleasure and pain can be involved in the experiential process when they are experienced as *possible*.

To further illustrate how feelings of anticipated pleasure and pain are involved in the experiential process, and even in ‘neutral’ experiences, I now want to present a case of phenomenal contrasts. The upshot of these contrasts is that the very same ‘neutral’ experience *x*, such as waiting for the bus, can yet vary depending on the concrete experience of possibilities of pleasure and pain that are involved in the experiential process in which *x* takes place. Thus, the idea is that if varying the experiential situation in which *x* takes place by altering the experience of possibility of pleasure and pain plausibly modifies the overall experiential character of *x*, then we have good reason to believe that even in a neutral experience *x* feeling plays a crucial role and so forms part of *x*—despite the alleged ‘neutral’ character that is superficially ascribed to *x*. Compare the five following cases:

- (1) Neutrally waiting for the bus, whereby my experience involves an *anticipation of pleasure*.
- (2) Neutrally waiting for the bus, whereby my experience involves an *anticipation of pain*.

---

<sup>786</sup> Husserl 2006 (Hua Mat. 8), 323, my translation: “dass nicht bloß hyletische Daten selbst in ihrer Impressionalität affizieren, sondern auch schon hyletische Antizipationen von Daten, und zwar als ‘Werten’, als positiv affizierenden”.



(3) Neutrally waiting for the bus, whereby my experience involves *no anticipation* of neither pleasure nor pain but the experience of *a general possibility for me to experience pleasure or pain*.

(4) Neutrally waiting for the bus, whereby my experience involves the *experience of a general possibility of pleasure or pain* but the *anticipation of the impossibility of the experience of pleasure and pain for myself*.

(5) Neutrally waiting for the bus, whereby my experience involves *no experience of possibility of pleasure or pain whatsoever*, as if I were 'blind' to pleasure or pain and completely indifferent concerning what or how my experience further unfolds.

Let us now examine the phenomenal contrasts between the five cases. In all of them, the experience of waiting is taken to elicit neither pleasure nor pain. But in each case, the experiential background is different. To better see this, we can consider a few examples. Looking forward to meeting my friend, I might anticipate a nice evening and don't mind the waiting for the bus (1). By contrast, it feels differently not to mind the waiting for the bus, if I anticipate to meet my girlfriend and I think she will leave me (2). Perhaps I am in no particular anticipation of any pleasure and pain, while I wait for the bus though I am aware that if I step forward on the bike lane I might get hurt (3). It is also possible to imagine a case in which one—possibly under the influence of some drug—experiences one's situation in a detached way with the following characteristics: one remembers and knows that normal experience involves pleasure and pain; one experiences a form of emotional numbness that motivates the anticipation that one's own case is such that one is incapable of feeling anything in terms of pleasure or pain; one knows experiences of pleasure and pain are *generally possible*, but 'feels' they are *impossible for oneself* (4). Finally, there is also the rather experimentally construed case in which there is a lack of a sense of pleasure and pain altogether, a total ignorance (= not-knowing) of pleasure and pain (5).

*What can we learn from phenomenal contrasts between the different cases?* First, the contrast between the first and second case shows that what seem to be similar neutral experiences can still differ depending on the concrete experiential process in which they unfold. It matters whether an experience *p* occurs within in the horizon of prospective pleasure or prospective pain. And crucially, the difference at stake is one that concerns feeling. For, even if one were to maintain that the difference concerns primarily thought, such a difference would precisely affect feeling. For instance, when in case (1) the waiting is accompanied by positive thoughts (fantasies about the meeting with my friend), whereas

in (2) negative thoughts (fantasies about the moment my girlfriend will say that she will leave me) predominate, then this difference is manifest in feeling.

Some might acknowledge the different feelings involved in (1) and (2) but maintain that neither (1) nor (2) are truly neutral experiences in the strict sense and that case (3) should be focused on when talking about neutral experiencing. On this line, (1) and (2) might be said to evidently involve feeling, whereas (3) is bereft of feeling, also given that no concrete anticipation of pleasure or pain is at the horizon. In response, I wish to propose that it certainly makes a difference whether one waits for the bus in anticipation of pleasure or pain (1,2) on the one hand. Or, on the other hand, whether one waits for the bus without any anticipation of pleasure or pain (3). For instance, because one is in a meditation-like state of mind. However, I think it would be wrong to misunderstand this difference as a difference between feeling experiences and experiences without *any* feeling. Again, I want to suggest the difference between (1,2) and (3) is also a difference *in* feeling. This, I believe, can be made plausible by considering the phenomenal contrast between (3) vs. (4) and (5) respectively.

Does it matter whether I take it to be generally possible for me to experience pleasure and pain? Comparing (3) and (4), I think we must admit both cases are very different. It matters whether I feel it possible that something could happen to me that either elicits pleasure or pain. It matters whether my world is experienced in such a way that certain joyful events are generally possible for me or whether there is no hope for me to make experiences that fill my mind with joy. Accordingly, Ratcliffe argues, rightly in my view, we must distinguish between *losing all hope* and *losing all hopes*.<sup>787</sup> I can feel all possible events that I have hoped for become unlikely to the effect that—as a matter of fact—I don't hope for them to happen any longer. All my concrete hopes, thus, might vanish, without me losing a general sense of the possibility that something might come up that will change everything for the better. Accordingly, even after having lost all hopes, I might retain what Ratcliffe calls “pre-intentional”<sup>788</sup> hope. It consists in the general hope that new specific hopes arise provided something still unknown changes my situation. However, in some cases, a general hopelessness might accrue that renders any specific hope impossible. Consider, for instance, states of severe depression, in which a person feels that no matter what happens nothing will be able to elicit an experience of fulfilment, not the slightest feeling of satisfaction or happiness. Evidently, the general lack of hope for happiness is something that features in the predicament of depression, characterizing precisely what it feels like to be depressed. And, importantly, such a condition is to be distinguished from (3). It matters whether pleasure is experienced as possible for me or not.

---

<sup>787</sup> Ratcliffe 2013a.

<sup>788</sup> Ratcliffe 2013a, 60.

That (3) and (4) are arguably phenomenologically very different also comes afore, if we now contrast depression-like experiences with (4). While depression often involves the absence of experienced possibility of pleasure for oneself, this absence is usually accompanied with severe experiences of pain. In depression, pain is an experienced possibility and the general hopelessness is painful. However, if we were to strip such an experienced possibility of pain away (thereby construing a state like (4)), I think it is fair to say that we would see a quite different experience.

Yet, some might still not be persuaded and argue that while (3) is a case in which there is no feeling involved, (4) is a case in which feeling is involved, given that the absence of the possibility for oneself to feel a certain pleasure is anticipated. Hence, one might seemingly still be able to maintain that only those experiences in which more or less concrete anticipations of pleasure and pain occur involve feeling and that, when those anticipations are absent, feeling is so too.

In response to this kind of objection, I suggest we need to compare (3) and (5). If it were really true that the possibility of pleasure and pain must feature in an experiential episode *in terms of* more or less *concrete anticipations* so as to elicit any feeling, then there shouldn't be a significant difference between (3) and (5). Just like in (5), an experiential episode along the lines of (3), in which no concrete anticipation of pleasure or pain is on the horizon, should have the character as if pleasure and pain were not possible. But, unlike in (4), not in terms of an experienced impossibility, but rather in terms of a total absence of the issue of the possibility of pleasure and pain. That is, as if the possibility of pleasure and pain did not feature in any way in the experiential episode. I take cases like (3) to be quite common in the life of mature human persons and I don't think that these experiential episodes are in any way comparable to those rather experimentally construed cases of (5). Sometimes we might have experiences that seem relatively neutral and that do not comprise pronounced anticipations of states of pleasure or pain to come. However, when these experiences transform into experiences of pleasure or pain, it is not as if we were surprised by a new class of experience: feeling. When I await the bus in terms of (3) and suddenly I smell a nice perfume that I enjoy, I don't learn the category 'pleasure and pain' anew, as if during neutrally waiting for the bus I have become totally oblivious to the possible experience of pleasure and pain. Hence, I suggest cases like (3) are fundamentally different to those like (5). By contrast, those who deny that in (3) feeling features in the experiential process would have to defend that (3) and (5) are experientially congruent. Consequently, I propose that it makes an experiential difference whether I take pleasure or pain to be possible (case 3) or whether pleasure and pain are unknown forms of experiences to me (case 5), and that this is a difference that concerns feeling. If we compare (3) with (1,2), the former might seem, relatively speaking, neutral and free of a pronounced feeling, but if we compare it with (5), most will agree, I believe, that it *feels* differently. And differently

in the sense that in contrast to (5), (3) can plausibly be said to involve feeling, whereas in (5) it becomes much harder to see the feeling character.

To conclude, what I suggest we can learn from the phenomenal contrasts between the different cases is that, though subtle, there are significant differences in feeling, depending on which concrete experiential process a given experience  $x$  occurs in. Moreover, even if  $x$  is an allegedly 'neutral' experience, feeling still plays a role, it is the way in which the I participates in an experience structured and determined through the experience of future possibilities. Put like this, however, what seem to be 'neutral' experiences are not neutral in the strict sense but only *relatively* neutral experiences, that is, 'neutral' measured against more pronounced feelings of pleasure, of acute joy, hate, or pain. The only true 'neutral' experience might be identified in case (5), a form of experience that is construed in such a way that it incorporates no experienced possibility of pleasure or pain. *This*, I propose, is the true anti-thesis to the feeling thesis, developed throughout the chapter, that feeling is intrinsic even to the most basic form of experience and self-experience. Whereas all the allegedly 'neutral' experiences we all know of in our lives, such as waiting for the bus, do not really present cases that would count against the view that experience involves feeling, only an abstractive and artificial construal of an experience remains as a candidate of allegedly affect-free experiential states. This has, as I have already announced, significant ramifications for the dialectical constellation, which I discuss in the remainder of the chapter.

## 6.5 Minimal self and the feeling of being oneself

What does the fact that most 'neutral' experiences are not fully neutral tell us about the relationship between feeling and minimal self? I suggest that the cases discussed in the last section require us to reassess the dialectical constellation of the thesis that feeling is involved in the minimal self. Most of what we would describe superficially as 'neutral' experiences have turned out to still involve non-neutral aspects: experienced possibilities of pleasure and pain. Accordingly, these experiences seemed to present clear counterexamples against the feeling interpretation of the minimal self. If our lives comprised recurring neutral and affect-free episodes, why should we believe the most basic form of experience and self-experience includes feeling? But on a closer examination, these experiences proved less 'neutral', and 'neutral' only insofar as they are compared to pronounced feelings and abstractively taken out of the experiential process in which they occur. Obviously, that makes it much harder to deny the feeling character of the experiential process. Yet, it remains, of course, possible to formulate doubt about it. I depicted the case (5) as a case of experience in which the possibility of pleasure or pain is completely absent

and suggested this kind of experience remains as a candidate for affect-free experiential states. However, it is important to acknowledge that this kind of experience has itself a *hypothetical* and, in this sense, problematic status. As long as we don't want to identify case (3) with (5), which I deem most won't have any propensity for, we must come to the conclusion that we simply don't have the kind of experience described in (5). No matter how 'grey' or affectively flat an experience in our life might seem, always, as a matter of empirical fact, we do have an experience of the possibility of pleasure and pain and therefore somehow feel ourselves in this or that way. Crucially, if that is correct, then this rather motivates the thesis that it is precisely the experiential process as a whole—and not necessarily some present particular object—in which feeling is rooted, and that feeling is a pervasive aspect of experience. Perhaps there are experiential states like (5), bereft of any experienced possibility of pleasure and pain. But surely, our experience is not like that, and so we don't know. And *that* is the shift in the dialectical constellation. The possible form of experience that *does* present an anti-thesis to the feeling theory is of dubious status. We don't know whether it is *really* possible or whether it is only conceivable, which it is, taken in abstraction from our real and concrete experience.

But doesn't that suffice to show that the minimal self doesn't require feeling? If an experience is conceivable that lacks feeling, but qua experience includes the minimal self, doesn't that mean that minimal self and feeling are separable? In my view, that would be too hasty. We need to bear in mind that the minimal self, on Zahavi's view, ought to fulfil three conceptual roles: (a) to explain the phenomenological structure of phenomenal consciousness, (b) to explain the phenomenological structure of self-consciousness, and (c) to explain the phenomenological structure of the most basic form of selfhood.

Now, if we consider the first two roles, and if we thus simply conceive of an experience *x* (a) and as *x* given to me (b), then the being of someone, myself the experiencer, is abstracted from. However, if we want to conceive of the minimal self in terms of (c), and when we accordingly try to focus on the experiencer while the givenness of *x*, in order to really capture the experience of *being* the experiencer to whom *x* is given, we need to—trivially—also consider the *experience of being* involved. Crucially, this distinguishes (c) from (a) and (b). To imagine and focus (when reflecting and in abstraction) on the phenomenal state *x* (a) and its first-person character of its givenness (b), we can easily and even must abstract from the sense of being that is involved in the being of selfhood (c). But if we want to focus on the being of selfhood we must precisely account for the sense of being involved.

Most importantly, however, if we were to argue that cases like (5) were possible, i.e., experiences in which no experienced possibility of pleasure or pain partakes, an experiential state of total emotional indifference and affective ignorance, as it were, then precisely, I suggest, following Ratcliffe's notion of existential feeling and Lotze's felt self-

concern, we will find it hard to capture the sense of being. Husserl's 'feeling participation of the I' (*fühlendes Dabei-Sein des Ich*) that he envisages in his later manuscripts, it becomes elusive when we try to cut off feeling from the primordial lived presence. But with feeling, so the thesis runs, goes the sense of being.

So, let me be clear: We can imagine cases like (5), but when we do so, we only imagine a phenomenal state and its first-personal character, and so abstract from the being of concrete selfhood. We then only imagine an experiential process, a stream of consciousness, in which phenomenal contents run off without any self-concern, no prospect of pleasure or pain, no awareness of the possibility of pleasure or pain either, no care—and no feeling. Hence, we shift from imagining the minimal self in terms of (3) to imagining the minimal self in terms of (1) and (2). For the conceptual roles the minimal self ought to play with regard to (1) and (2), no sense of being and feeling is required. But for the role of (3), somehow we need to make plausible how the sense of being features in the minimal self. I can see no other way of doing so than by acknowledging feeling as part of basic self-experience, following a great number of thinkers, in the philosophical tradition but also more recently.<sup>789</sup>

As I have tried to show, emphasizing feeling as *integral* part of the minimal self does not undermine the original account of Zahavi, but rather complements it. That means, I do not want to say that the minimal self and its phenomenological characterization through Husserlian time-consciousness is too formal and too minimal. But what I *do* want to say is that Zahavi has focused predominantly on the first two conceptual roles of the minimal self. He is right in acclaiming the third conceptual role for the minimal self too, but its phenomenological characterization is, I would suggest, incomplete, as Zahavi's account hasn't addressed the sense of being involved. It needs to be complemented by what I would like to suggest to call the *feeling of being oneself*<sup>790</sup>. Such a feeling, Husserl's 'feeling participation of the I', is, notably, integral to the temporal structure of the stream of consciousness. Hence, we do not need to complement inner time-consciousness by an additional component or any further self-*content* that would render the minimal self materialized and would objectify it. The feeling of being oneself, in its most basic *form*, only refers to the—temporal—*unity* of the experiential process *as mine*.

Let me finally illustrate this by once again turning to the phenomenological structure of inner time-consciousness and the kinds of 'intentionalities' involved. Husserl, and so does Zahavi, distinguishes between

---

<sup>789</sup> I will discuss some of the recent examples in the next chapter and contrast them with my own proposal.

<sup>790</sup> I am grateful to Matthew Ratcliffe for suggesting this term to me when discussing my work.

(a) transverse-intentionality (*Querintentionalität*):

This ‘intentionality’ is directed at the impressional moment and the congruence of meaning between the impressionally given and that what is retained in retention: for instance, the meaning of the tone ‘c’ of a melody as lasting.

and

(b) longitudinal-intentionality (*Längsintentionalität*):

This ‘intentionality’ is directed at that what is new in impressional consciousness as compared to what is retained, intending thus the streaming of the flow of consciousness.

Now, these ‘intentionalities’ are possibly distinguished in reflection, but pre-reflectively they are, of course, entwined. The same applies to the conceptual roles of the minimal self, which, I want to propose, correspond each to an ‘intentionality’ or phenomenological moment of the primordial way of lived presence.

(a) If we consider the first conceptual role of the minimal self to explain the phenomenological structure of phenomenal consciousness, we may focus on transverse-intentionality. Then we can see how inner time-consciousness (the minimal self) enables the givenness of *x* in the first place. It is the temporal streaming character that makes possible that something is phenomenally present and given now, and it is transversal intentionality in which something can be constituted as a persisting object of experience to begin with.

(b) If we consider the second conceptual role of the minimal self to explain the phenomenological structure of self-consciousness, we may focus on longitudinal intentionality. Then we can see how inner time-consciousness, the retentional self-givenness of the flow of consciousness, provides pre-reflective self-awareness, which, qua reflexivity of experience, motivates a minimal sense of self.

But how about the third conceptual role of the minimal self to explain the phenomenological structure of the being of selfhood? Surely, as I have argued in the previous chapters with Zahavi, the being of selfhood consists in having self-experience. But how about the experience of being? Does it equally correspond to longitudinal intentionality? I would like to offer a different option and suggest that there is a third kind of ‘intentionality’ involved in inner time-consciousness that can be distinguished in reflection:

(c) pole-intentionality<sup>791</sup>:

This ‘intentionality’ is directed at that what is new in impressional consciousness as compared to what is protended. Focusing on the constant change of impressional consciousness and the covariation of what is protended, we intend the invariant pole of experience, the fact that any anticipated experience-yet-to-come is the experience-yet-to-come of the present consciousness.

Just like in the case of transverse-intentionality and longitudinal-intentionality, pole-intentionality, as I suggest to call it, is an *intentionality* insofar as it is a possible way of directedness in reflection with regard to my streaming experience. Accordingly, I wish to avoid the following misunderstanding: I do not wish to imply that the pole of experience was already explicit in pre-reflective experience. But what I do want to claim is that it is implicitly *felt*, as possibility-consciousness is essentially, provided the analysis of this chapter is correct, a non-neutral experience. And taken as an ‘intentionality’, it is a non-neutral attitude towards my own being, though pre-reflectively only in a felt and not yet thematized manner. Just like Zahavi, following Husserlian scholarship, can say that there is “a *protoreflection* in the core of the prereflective self-awareness”<sup>792</sup>, we may say, so I suggest, such a protoreflection is by the same token a *protoidentification* with oneself, the future experiencer. Just like I can ponder on my past experience in reflection and thereby make myself explicit as the ego-pole that has undergone different experiences, I can, in reflection, ponder on my future and make myself explicit as the ego-pole that will undergo different experiences. But the crux is that such an *anterograde* reflection and identification is already prepared on the pre-reflective level. Protentional consciousness is already the awareness that there is possibly more experience to come, that there might be given something else *to me*, and thus, that I, the present consciousness am more than is currently given in actuality. The protentional openness towards the world is the openness towards my own future. This future might not be explicit on the pre-reflective level, let alone explicit as *my* future qua ego. But, and this is the view I propose, the future of myself and *so* my being, is always already felt.

---

<sup>791</sup> I chose to speak of *pole*-intentionality because Husserl was eager to conceptualize the participation of the ‘I’ as the pole in affection and action, prior to its reflective thematization: “[A]nd we can, in abstract attitude ..., consider how a genetic beginning would have to be construed that already comes with hyle and feeling, but also with the I-pole as pole of affection and action.” My translation: “[U]nd wir können in abstraktiver Betrachtung ... überlegen, wie ein genetischer Anfang zu konstruieren wäre, der jedenfalls Hyle und Gefühl schon brachte, aber auch Ichpol als Pol von Affektion und Aktion.” (Husserl 2006 (Hua Mat. 8), 352) Cf. Husserl 2006 (Hua Mat. 8), 3, 20, 49, 53, 183, 187, 189f., 352, 367, 368.

<sup>792</sup> Zahavi 1999, 135.



## 6.6 Conclusion: The feeling of being oneself

In this chapter, I addressed the question of how being features in the most basic form of self-experience, the minimal self. First, I discussed some recent criticisms of the minimal self-view that raised what I dubbed the too-thin-charge. On this view, the minimal self is so thin and formal that it cannot account for the experience of one's own individual existence and so fails to deliver a qualified description of selfhood. By contrast, I argued that Zahavi's minimal self is well-equipped to countenance the question of individuality. I then discussed another possible criticism that could be applied to Zahavi's reflexivist view about the minimal self-view: Lotze's view that the experience of our individual existence is rooted in a self-feeling, the feeling of one's own immediate and absolute value, or in short, self-concern. Following this view, I proposed to integrate feeling, as the way in which one's own existence is experienced, into the minimal self account. To this aim, I discussed Husserl's notion of the primordial way of lived presence and his attempt to describe feeling as the way in which the 'I' is partaking already on the most basic process of experiencing. However, as I showed, Husserl did not find a way to further describe the feeling participation of the 'I' in the primordial form of lived presence. I then suggested, applying Ratcliffe's notion of existential feeling, that feeling is best described as experience of possibility and that feelings are involved in experience in terms of protentional consciousness. Doing so, I offered a possible way for feeling to be integrated into the minimal self account. I further tried to illustrate this by a discussion of Husserl's notion of interest, showing that he takes protentional consciousness in the primordial lived presence to be characterized by an all-founding interest that is directed both towards the constitution of an object and, by the same token, the continuation of the experiential process as a whole. I then considered a possible objection against the view that self-concern and feeling forms part of experience as such: 'neutral' experiences. Referring to Heidegger and Lotze, I argued that those experiences we often might deem 'neutral' are neutral at best when compared with more pronounced feelings. On a closer examination, I emphasized that even the 'neutral' experiences that are typical for our human life are not fully neutral and include some feeling determined by the experienced possibility of death and by the experienced possibility of pleasure and pain. Only by virtue of severe abstraction can we conceive of possible—though artificial—experiences that are truly neutral insofar as they involve no experienced possibility of pleasure or pain and imply a total indifference regarding the experiential process. Imagining such an abstract experience, I argued, we must abstract from all sense of being. Hence, when we imagine such a neutral experience, what we conceive of is only phenomenal consciousness, though involving self-consciousness, but lacking any sense of being. By contrast, I finally proposed, if we want to conceive of the experiencer as being real, we need to consider what I call the *feeling of being oneself*, which I take to be intrinsic to the experiential process as such. While I would acknowledge that

there is no absolute knock-out argument that could be brought into the field, proving once and for all, as it were, that all phenomenal consciousness and self-consciousness necessarily involves feeling, I have tried to make a corresponding claim plausible. In the following, last chapter of the dissertation, I want to further support that claim by specifying the concept of the feeling of being oneself, by contrasting it with recent accounts of self-feeling, and by applying it to the discussion of the psychopathological phenomenon of depersonalization. For some have construed depersonalization as a case that lacks subjective character or feeling and so would presumably take it to challenge my conception of the feeling of being oneself.

## Chapter 7

# Specifying the Feeling of Being Oneself and Depersonalization

Feelings are elements or determinations of myself. They are the I-contents or I-qualities. They constitute the I, namely the I that I—do not think or infer—but immediately experience, that accompanies myself in any moment of my life, the immediate I of consciousness or the immediately experienced ‘subject’. We have called it already, because it is given in the feelings, the feeling-I (*Gefühls-Ich*). It is the same if we denote it as I-feeling (*Ichgefühl*).<sup>793</sup>

I call it the individual feeling of life. The single feelings and affects that are evoked by the contact with the external world are just the waves above the ground of this general emotional holistic phenomenon . . . . . The feeling of life comes closest to the category of moods. . . . However, it is not simply a member in the row of what we otherwise call moods, but rather lies below them, as it were.<sup>794</sup>

The preceding chapter introduced the notion of the feeling of being oneself in order to describe the feeling dimension as the most basic form of self-experience: our existence is experienced, I argued, as a feeling of being oneself. With this emphasis on feeling, I aimed not to append an additional feature onto the most basic form of self-experience. Instead, I argued that the most basic form of self-experience involves a feeling dimension. Protentional consciousness as a consciousness of possibility is never fully neutral. Rather, it is always loaded with significance—with something that has meaning *for me* and matters *to me*, the one who is experiencing. It is with ‘being concerned’ and with ‘care for my own future experience’ that I feel my being or my experiential existence.

This chapter further determines the specifics of such an approach to self-feeling. The first section does so by describing various features of my own proposal in light of other contemporaneous accounts of self-feeling. The second section offers a further argument to corroborate the view that self-experience encloses a feeling dimension or a feeling of being

---

<sup>793</sup> Lipps 1901, 14, my translation: “Dagegen sind Gefühle Elemente oder Bestimmungen meiner selbst. Sie sind Ichinhalte oder Ichqualitäten. Sie constituieren das Ich, nämlich das Ich, das ich—nicht denke oder erschliesse, sondern unmittelbar erlebe, das mir in jedem Momente meines Lebens vorschwebt, das unmittelbare Bewusstseins-Ich oder das unmittelbar erlebte ‘Subject’. Wir nannten es schon, weil es in den Gefühlen gegeben ist, das Gefühls-Ich. Es ist Dasselbe, wenn wir es als Ichgefühl bezeichnen.”

<sup>794</sup> Oesterreich 1910, 317, my translation: “Ich nenne es das individuelle Lebensgefühl. Die einzelnen Gefühle und Affekte, wie sie die Berührung mit der Außenwelt in uns hervorruft, sind nur die Wellen über dem Untergrunde dieses allgemeinen emotionalen Gesamtphänomens . . . . . Das Lebensgefühl steht am nächsten offenbar der Kategorie der Stimmungen. . . . Es steht jedoch nicht einfach in einer Reihe mit dem, was wir sonst Stimmung nennen, sondern es liegt sozusagen noch unterhalb dieser.”

oneself—self-feeling, more briefly. The main rationale of this argument is that a notion of self-feeling (of the kind I propose) is helpful, if not urgently needed, in accounting for a particular psychopathological phenomenon: depersonalization. Depersonalization is a condition wherein afflicted individuals feel detached from their experiences. Often, detachment is so strong that it may even appear to them that they do not exist at all. As a phenomenon, depersonalization is highly relevant to philosophy because it appears to present a case in which the experiential life of an individual may lack any self-experience whatsoever. Depersonalization further serves as a decisive example of the high variability of the experience of being, underscoring the existence of something like an experience of being to begin with. I discuss the phenomenon in light of a recent, radical proposal according to which depersonalization is attributed to an absence of subjective character.<sup>795</sup>

Contrary to this proposal, I argue that the experience of depersonalization does not lack all subjective character even though it is severely modified. Instead of simply denying all subjective character, I suggest we need a more refined understanding of the meaning and structure of subjective character and precisely how it is modified by depersonalization. Drawing from an early and rather unknown account of depersonalization by Konstantin Oesterreich (1910), I propose that it is the feeling dimension of subjective character that we need to look at in order to understand what, exactly, is modified in depersonalization. To anticipate, by shedding light on the alteration of self-feeling in depersonalization, I ultimately make the following claims:

- that despite a modification of its subjective character, the experience of depersonalization *continues to involve self-experience*;
- that the *subjective character of experience is modified* because it lacks a ‘normal’ feeling of being oneself;
- that experience involves *disturbing feelings of non-existence*. These should not be equated with an absence of self-feeling, but rather with a modification of the latter;
- that *variation in the experience of depersonalization is owed to variation in self-feeling*, corroborating the view that feeling constitutes consciousness of existence.

Accordingly, the case of depersonalization helps further elucidate the relationship between pre-reflective self-awareness and the feeling of being oneself.

---

<sup>795</sup> Billon 2016.

## 7.1 Specifying the ‘feeling of being oneself’

In the history of philosophy, the view that self-consciousness involves or consists of feeling is far from predominant. Today, it continues to be located at the periphery of mainstream philosophy. Nonetheless, there is increasing acceptance of such a view in debates that have occurred over the last two decades (proliferated by, for instance, Antonio Damasio 1994, 1999). But even though a growing number of researchers and thinkers have converged around the view that feeling is central to self-consciousness, it is important not to conflate the different proposals. Thus, the first part of this chapter specifies my own notion of the ‘feeling of being oneself’, by describing key features of it in order to clarify how it relates to alternative accounts for self-feeling.

### 7.1.1 Feeling and consciousness

The preceding chapter introduced the notion of the ‘feeling of being oneself’ within a discussion of the most basic form of self-experience. Here, I aimed to demonstrate that even the most basic form of self-experience involves a feeling dimension. But whether feeling is something that is necessarily experiential is another question entirely. In fact, some thinkers who emphasize feeling as forming the ground for experience and self-experience seem to endorse a view according to which feeling or affectivity, in its rudimentary forms, is something non-experiential.

Consider, for instance, Colombetti’s notion of “primordial affectivity”<sup>796</sup>. The way she determines the basic dimension of affectivity resembles strongly what I have described in the last chapter:

The mind, as embodied, is intrinsically or constitutively affective; you cannot take affectivity away from it and still have a mind. Affectivity as discussed in this chapter [=primordial affectivity, P.S.] refers broadly to a *lack of indifference*, and rather a *sensibility* or *interest* for one’s existence.<sup>797</sup>

However, appearances deceive as Colombetti understands affectivity not necessarily in reference to *feelings*. For her, “primordial affectivity need not be conscious”<sup>798</sup> as “it is enacted by the whole organism, and indeed even by organisms that lack a nervous system”<sup>799</sup>. This obviously contrasts starkly with my proposal of the *feeling of being oneself*, which I take to be essentially experiential. Yet, one might ask, are both proposals

---

<sup>796</sup> Colombetti 2014, Ch. 1.

<sup>797</sup> Colombetti 2014, 1.

<sup>798</sup> Colombetti 2014, 21.

<sup>799</sup> Colombetti 2014, 21.

necessarily incompatible with one another? After all, Colombetti herself emphasizes the distinction between affectivity and feeling.

However, I find this separation implausible for various reasons that ultimately all lead to a single concern: the risk of a category mistake (*metabasis eis allo genos*) that consists of transferring what are intrinsically phenomenal notions to the non-conscious sphere. This is manifest in the following quotation:

[T]he simplest living systems already realize a relationship with themselves and the world in which they are situated that entails purposefulness and concern for their existence. But such purposefulness and concern need not be accompanied by consciousness; rather, they ought to be understood as properties of a specific organization that sets up an *asymmetry* between the living system and the rest of the world, which consists in a perspective or point of view from which the world acquires meaning.<sup>800</sup>

Colombetti not only ascribes *interest* and *affectivity* to non-conscious living systems, but also *concern for their existence* and a *perspective* or *point of view*. Later in the chapter, she eventually claims that non-conscious primordial affectivity is “the bringing forth of a *world of significance*”<sup>801</sup>. Yet none of these italicized terms can be transferred to the non-conscious sphere without losing what is essential about them. Certainly, even non-conscious living systems interact with what we (conscious living systems) experience as the surrounding world. Perhaps Colombetti and her enactivist approach is right in that experience only emerges out of such a fundamental interaction between organism and environment<sup>802</sup>. But such interaction hardly means there is a world *for* the organism, or that this world could be somehow given without being given in an experiential manner. There may be living systems that, upon inspection, seem to *require* us to describe their behavior using terminology such as ‘perspective’, ‘point of view’, ‘interest’, ‘significance’, ‘purposefulness’, and ‘affectivity’. But instead of speculating about a non-conscious variant for each of the terms, I suggest we are better off assuming that the living system has a form of phenomenal consciousness—however rudimentary it may be<sup>803</sup>. Here, I do

---

<sup>800</sup> Colombetti 2014, 2.

<sup>801</sup> Colombetti 2014, 21, my emphasis.

<sup>802</sup> A convincing case is made by Fuchs 2012. He, too, argues that experience and self-experience are based on the organization of the living organism and its interaction with the environment. And he, too, emphasizes continuity between life and lived experience. But in his rendition of this continuity, he does not claim that notions applicable to lived experience—such as *Befinden* and *conation*—can be applied to sheer, non-conscious life processes and living systems. Rather, he suggests they *are* the life processes of the organism themselves insofar as they become conscious (Fuchs 2012, 152). They are the manifestation or continuation of life processes on the level of lived experience. In contrast, Colombetti introduces a doubling of notions such as affectivity, interest or perspective—one corresponding to the non-conscious, the other to the conscious sphere.

<sup>803</sup> I take it then that when ascribing a perspective, world, self-concern, and affectivity to an organism, we are either dealing with a form of empathy or projection (of our experiential condition onto an organism that is fundamentally different). *Tertium non datur*.

not intend to reject the “thesis of the *deep continuity* of life and mind”<sup>804</sup>. But what strikes me as problematic is the belief that non-conscious organisms could have a world, a perspective, an interest, or a concern while conscious organisms might have the same. The only difference for the latter, on top of all that, is they “also...exhibit awareness”<sup>805</sup>. Whatever non-phenomenological sense of the world a living system might have from interacting with the environment and processing corresponding information, it does not amount to having a world.<sup>806</sup> Even though there may be continuity between non-conscious and conscious life to the extent that there is no hiatus between them, the emergence of a world, perspective, interest, concern, and affectivity presents a shift in quality. I believe they all need phenomenality of some form—even if the emergence of phenomenality is best understood as the outcome of interaction between an organism and its environment.

Obviously, this all touches upon decisive philosophical questions that are being controversially discussed in current debates in the philosophy of the mind and cognitive sciences. My aim here is not (and cannot be) to provide a conclusive verdict on these matters. All I want to do is emphasize that the notion of the feeling of being oneself that I propose in this work is restricted to the conscious sphere. That is, I take feeling to be essentially experiential and linked to pre-reflective self-awareness.

### 7.1.2 *Seinsgefühl*: Existence, life, and the body

Most, if not all, thinkers who propose a notion of self-feeling emphasize that feeling is the kind of consciousness wherein we become originally aware of our existence, i.e., *that we are*. Even Kant, who emphasized that we do not gain any knowledge about the true nature of ourselves by examining our self-experience, stressed that we do undergo an experience of our existence. Crucially, he believed this experience is a feeling: “Now, it (the I; M.F.) is nothing more than a feeling of existence, without any concept and only a presentification (*Vorstellung*) of that to which all thinking (*relatione accidentis*) is related.”<sup>807</sup> As Manfred Frank highlights, many eighteenth-century thinkers—such as Hölderlin, Jacobi, Herder, Schleiermacher, Maine de Biran, Malebranche, and Rousseau, to name just a few—

---

<sup>804</sup> Colombetti 2014, xvi. Cf. Thompson 2007, ix, Fuchs 2012.

<sup>805</sup> Colombetti 2014, 2.

<sup>806</sup> As an implication of the preceding chapters on self-experientialism, I take this to apply *a fortiori* to the notion of selfhood. When Damasio speaks of a “proto-self” as “a coherent collection of neural patterns which map, moment by moment, the state of the physical structure of the organism in its many dimensions” (1999, 154), then Fuchs is right to call it a “*precursor* of self-experience” (2012, 161, emphasis mine). But as such it is not yet a *sense* of self. In my view, a sense of *self* emerges only with the reflexivity of experience. Similarly, even if a living organism might distinguish between what is in fact itself and its environment prior to any experience, in making this distinction in its action it is not yet showing reflexivity. One could say, roughly and *cum grano salis*, that the relationship of the non-conscious living organism to itself remains *de re* and only turn truly *de se* when it is self-aware in experience.

<sup>807</sup> Kant’s *Prolegomena* from 1783, quoted in Frank 2002, 41, my translation.

formulated similar views that referred to a feeling of one's existence<sup>808</sup>. For these thinkers, the central idea was that self-feeling reveals our existence as conscious subjects, i.e., that it somehow feels like *being* the one to whom objects are given, or to exist inasmuch as one experiences. A similar understanding also manifests in Lotze's account presented in the last chapter. Another later, prominent proponent of the view that self-consciousness consists essentially in feeling is Theodor Lipps, who states "... always once that experience (*Erleben*) takes place, I have an immediate awareness of it. I find myself as the experiencer or bearer. *I feel myself experiencing*"<sup>809</sup>. The feeling of one's own existence thus means a feeling of experiencing: self-feeling is a feeling of myself inasmuch as I experience.

But there is also a second, somewhat different, line of argument in theories of self-feeling—especially those advocated in the last couple of years of the debate. Proponents of this second line have a different focus for their defense of the notion of self-feeling. They talk about a feeling of *life*, being *alive*, and *vitality* instead of *being* or *existence*.<sup>810</sup> Rather than stress the feeling of being the subject of an ongoing experience, they point to the self-feeling of the living organism, its vitality, and its physiological function.

Consider, for instance, the distinction of "four levels of existential feelings"<sup>811</sup> proposed by Jan Slaby and Achim Stephan to describe the different ways we are consciously aware of ourselves through feeling. The *first* level of existential feeling, they suggest, is "mainly related to...one's organic being in a world in general"; the *second* to "one's relations, projects, existential contexts in general"; the *third* to "more specific social relations, projects, contexts, capabilities, coping potentials"; and finally, the *fourth* to "specific situations currently unfolding"<sup>812</sup>. Here, my interest lies not with the question of whether and how one can distinguish these different levels as sharply as the distinction might suggest<sup>813</sup>. I find the descriptions informative, presenting good examples of how existential feelings are intrinsically self-feelings. My concern lies in how they depict the first level. They give the following examples to clarify "pure existential feeling"<sup>814</sup>:

- Feeling alive
- Feeling that there is a world and that one is part of it
- Feeling that one has or is a body

---

<sup>808</sup> See Frank 2002, particularly 77–92.

<sup>809</sup> Lipps 1902, 6, my translation and emphasis: "Und immer findet dieses Erleben nicht nur statt, sondern ich habe auch davon ein unmittelbares Bewusstsein. Ich finde mich als den Erlebenden oder Habenden. *Ich fühle mich erlebend.*"

<sup>810</sup> For instance, Fuchs 2007, 2012, 2013a, b, Engelen 2014, Koch 2019, Slaby 2008, 2012, Slaby & Stephan 2008.

<sup>811</sup> Slaby & Stephan 2008, 510.

<sup>812</sup> Slaby & Stephan 2008, 510.

<sup>813</sup> Cf. Ratcliffe 2012, 43f.

<sup>814</sup> Slaby & Stephan 2008, 510.



- Feeling basic vital functions

As these examples show, Slaby and Stephan (2008) take the primary self-feeling as consisting of feelings that disclose one's own organismic and bodily constitution. In vital self-feelings, what we feel is not simply that we are bodily located *qua* spatiotemporal objects in the world. We feel that we are alive, i.e., that we partake in the world as living bodies.

Another relevant example is Fuchs (2012, 2013a), who claims that our most basic form of self-experience consists of self-feeling that is best described as a feeling of being alive. It is comprised of two components:

1. *vitality*: "an overall feeling for one's present life situation"<sup>815</sup>; and
2. *conation*: "a basic striving that manifests itself in drive, instinct, need and affect"<sup>816</sup>.

In this view, the feeling of being alive is primarily concerned with one's situation (or situatedness) in the world as a living organism. It is the seamless continuation of life processes that, if the organism has the capacity for consciousness, translate into lived experience. The feeling of being alive pervades all consciousness, becoming the root cause and phenomenological basis for any experience.

Both lines of understanding thus focus on different aspects of selfhood. The first emphasizes the feeling of being the one for whom something is experientially given. The second underscores awareness of one's life in terms of the body and vitality. The question therefore arises: how does my proposal of the feeling of being oneself relate to these two ways of understanding self-feeling? While both notions refer to different aspects of our experiential life, I see no reason why they shouldn't be compatible. However, that doesn't mean one can be reduced to the other. Likewise, neither can play the role of the other. The feeling of being the one to whom something is given is not congruent with having a sense of one's own vitality, physiological function, and biological processes. Nor can we understand the feeling of being the one who is undergoing an experience if we direct our attention to the feeling of being a living organism.

I propose that the notion of the feeling of being oneself is comprised of both aspects. Feeling oneself existing is a feeling of being both in terms of experiential selfhood (i.e., that one is undergoing experiences) *and* in terms of a living, bodily organism. To some, this may sound trivial. However, reducing our primary self-feeling to bodily and vital self-feelings creates at least one significant risk.

The risk emerges from conceiving of primary bodily self-feelings as feelings exclusively concerned with the physiological and vital status of an organism. This renders self-

---

<sup>815</sup> Fuchs 2012, 152.

<sup>816</sup> Fuchs 2012, 152.

feeling as a representation of an otherwise non-conscious biological life process. Let me further illustrate this with the concept of core affect described by James A. Russell (2009):

Core affect is a pre-conceptual primitive process, a neurophysiological state, accessible to consciousness as a simple non-reflective feeling: feeling good or bad, feeling lethargic or energized.<sup>817</sup>

Core affect comprises feelings of one's own bodily situation, or self-feelings on the current vital condition of the organism. It refers to the physiological states of an organism that exist regardless of whether the organism feels them or not. An organism may or may not feel hunger, for instance. But if the organism has a nutritional deficit, it will enter a certain state associated with specific biochemical and metabolic processes. These may or may not trigger a certain core affect; regardless, the bodily state exists. My hunger could go completely unnoticed, given the right drug, while my organism still undergoes the biological process of starving. If it emerges, core affect is (as Colombetti nicely describes it) "a perception of how the organism is faring"<sup>818</sup>.

But understood like this, self-feeling becomes a mere representation of the life processes of the body. In contrast, and in line with Fuchs' notion of lived experiences as being continuous with the life processes of the organism, I suggest we need to understand that even vital self-feelings are not exclusively about the physiological state of an organism (as if experience was a kind of detection or monitoring device). Vital self-feelings are, or inextricably interwoven with, a feeling of being the one for whom experiences unfold. They are thus also about the experiential process, or the experiential selfhood of the living organism. When feeling itself, the living body is always also a lived body in that it is an experiencing body. Vital self-feeling is not only about something that would take place if there were no awareness of it<sup>819</sup>. It is not simply an awareness of life (*Leben*), but always also self-awareness of lived experiences (*Erleben*), i.e., the *Seinsgefühl* of the experiencing self. If I feel cold, for example, I do not simply want to raise the temperature of my body; I want to *feel* warmer. Whenever an experiencing organism feels his basic vital functioning, it is always *itself*—the experiencing being—that is at stake.

Even in cases where a feeling *does* represent or phenomenally manifests a physiological state of the organism, it is *qua* experience always also about our experiential being. Again, take the example of hunger: when I experience hunger, this experience consists of an uncomfortable bodily feeling that seeks relief (which is also an experience). Or, when hunger is combined with appetite, it is the feeling of striving towards pleasure (which is also an experience). Hunger certainly signifies and refers to a bodily need. But

---

<sup>817</sup> Russell 2009, 1264.

<sup>818</sup> Colombetti 2014, 21.

<sup>819</sup> Cf. Ratcliffe 2005, 39.

when it is experienced, it is also—if not phenomenologically—about a certain experience. As such, it is also about me, the experiencing self<sup>820</sup>.

### 7.1.3 The value of existence

Referring to Lotze in the last chapter, I suggested that our experiential life is never bereft of the dimension of pleasure and pain. This is true even though there could be singular tokens of experience that, taken strictly in theoretical abstraction, may seem to have no clear and pronounced hedonic component. However, I argued that even these singular acts are embedded in an experiential field wherein pleasure and pain always contribute to what constitutes the horizon of experience. That is, the experiential process is such that pleasure and pain are always involved as possibilities. Once more, my view converges to a high degree with Locke's:

*Self* is that conscious thinking thing, (whatever Substance, made up of whether Spiritual, or Material, Simple, or Compounded, it matters not) which is sensible, or conscious of Pleasure and Pain, capable of Happiness or Misery, and so is concern'd for it *self*, as far as that consciousness extends.<sup>821</sup>

As Locke emphasizes, the experience of the possibility of pain and pleasure elicits (or amounts to) a self-concern. It is because hedonic experiences are always experienced as possible that the experiential life itself is soaked with the dimension of value. With pleasure on the experiential horizon, the experiential horizon itself becomes valuable. The experience of the possibility of pleasure signifies a *self-interest* that is prior to any *egoism*. Whereas the latter means the tendency to act in a way that benefits one's own interest, regardless of what that might mean for others (and is thus a much more complex phenomenon that involves higher levels or structures of the experience of a shared world), the former can be described with some further words from Locke: self-interest is "concern for Happiness the unavoidable concomitant of consciousness, that which is conscious of Pleasure and Pain, desiring, that that *self*, that is conscious, should be happy"<sup>822</sup>.

Crucially, the experience of the possibility of pleasure translates into self-interest inasmuch as the experience of the possibility of pleasure intrinsically strives towards fulfillment. That is, pleasure and its possibility do not require an additional motivational component. As a phenomenon, they are motivational, attractive, and (notably) valuable in themselves. Interest in pleasure is *self-interest* at the same time, because striving towards

---

<sup>820</sup> Note that with emphasizing that self-feeling is about the experiencing self I do not refer to an allegedly ontologically distinct entity "self" that could be detached from the living organism. Rather, with the experiencing self I refer to the experiential life of the organism in contrast to its mere biological life.

<sup>821</sup> Locke 2008/1689, 214.

<sup>822</sup> Locke 2008/1689, 218.

fulfillment of the possibility of pleasure means self-fulfillment. The attraction of the possibility of pleasure concerns me, the actual ‘experiencing me’ who is undergoing the current experience. This is what Locke means with the term “that that *self*, that is conscious, should be happy”. The experience of the possibility of pleasure is the experience of the possibility of myself experiencing pleasure; feeling the value of pleasure and its possibility is, by the same token, feeling one’s self to be valuable. Being concerned about pleasure and pain is being concerned about oneself, for striving towards pleasure is striving towards oneself feeling pleasure.

Accordingly, the value of existence that is experienced in the experience of the possibility of pleasure is intrinsically a self-value. The value of experience is not something that I could possess like an object or undergo as a singular token of experience. I cannot lose my experience and the value it has without losing myself. It is in this sense that one can say the value of existence is originally and most fundamentally experienced as self-value. This self-value is not dependent on a certain feature of oneself or on the concrete content of my experience. Instead, it refers to the process of experience as a whole. In this particular sense, it is absolute. To repeat Lotze’s dictum: it is a felt self-concern, one’s immediate self-interest and self-value that constitutes the experience of one’s own existence.

The notion of self-value that I propose with Lotze should be distinguished from alternative notions of self-worth offered by others who propose to account for self-feeling. Consider, for instance, the proposal by Theodor Lipps. Influenced by Lotze, Lipps accepts that self-consciousness involves self-feeling. But for him, the aspect of self-value and self-interest is not prominent<sup>823</sup>. He thematizes self-worth as a specific sub-category of “feelings of worth (*Wertgefühle*)”<sup>824</sup>. For Lipps, feelings of self-worth are a specific kind of feeling in that they have a concrete ‘I’ as an object. They presuppose reflection and memory, and they are based on a certain degree of self-objectification. By reflecting on our own experiences and activities, we can develop a feeling of self-worth—namely, when we see ourselves in light of *x*. Feelings of self-worth are thus evaluative, involving consideration of oneself vis-à-vis certain standards, experiential contents, or individual features<sup>825</sup>.

In contrast, the notion of felt self-value that I suggest with Locke and Lotze refers to the value of the experiential life and selfhood as such. It neither hinges on nor concerns individual features. That said, I believe it is difficult to fully phenomenologically isolate and separate (a) the felt self-interest or self-value that I take as intrinsic to the experiential process from (b) the more evaluative feelings of self-worth concerning oneself in regard to

---

<sup>823</sup> Lipps 1901, 1902.

<sup>824</sup> Lipps 1902, 175.

<sup>825</sup> Lipps 1902, 167, 176ff. Sometimes, it sounds like Lipps also takes feelings of self-worth to accompany all egological activities. However, he ultimately fails to clarify whether he believes there are also pre-reflective feelings of self-worth. Cf. Lipps 1902, 178.

certain individual features and experiences. Precisely how they hinge together is an open question that requires further investigation<sup>826</sup>. But what can be said, I suggest, is that selfhood is experienced as valuable in an immediate sense—even prior to any reflective evaluation of oneself in light of characteristic features or certain standards of experiential life<sup>827</sup>.

#### 7.1.4 The entwinement of feeling, temporality, and self

Two further aspects of my proposal about the feeling of being oneself (worked out in the preceding chapter) that I want to highlight and distinguish from alternative accounts of self-feeling are: (1) the intrinsic relationship between the temporality of consciousness and feeling; and (2) the feeling nature of selfhood. To illustrate these points, the following is a sketch of another recent proposal for self-feeling.

In his book *Self-feeling*, Gerhard Kreuch asks whether self-consciousness can be understood as a feeling<sup>828</sup>. The answer, he believes, is positive. He notes that self-feeling is “in the background of all our experiences”<sup>829</sup> and “affectively discloses our individual existence in this world”<sup>830</sup>. Along with the critics of Zahavi’s minimal self-view discussed in the last chapter, he shares the idea that “‘for-me-ness’ remains an empty and formal notion”; it “describes a formal feature of our experience but no more”<sup>831</sup>. Despite his criticism of Zahavi, he suggests that the twin accounts for self-feeling and the minimal self can be “both true and complement each other”<sup>832</sup>.

All this seems to fit well with my account for the feeling of being oneself. At the same time, there are some decisive differences. First, I do not subscribe to the view that the minimal self (as described by Zahavi) is too formal, empty, or in need of “material content”<sup>833</sup> in order to count as self to begin with. Rather, I claim that the minimal self as it has been described by Zahavi is not *neutral* itself; it involves a felt self-value or self-

---

<sup>826</sup> In this dissertation, I cannot pursue such a task. A relatively unknown philosopher, who has dealt with this issue, is Else Voigtländer, a student of Theodor Lipps who endorsed a notion of pre-reflective feelings of self-worth and tried to conceptualize their interplay with more complex and socially structured forms of feeling of self-worth. Cf. Vendrell-Ferran 2015, 2020a, b, 2021, forthcoming, Voigtländer 1910. A related view can be found in Cooley 2009/1922/1902. Another resource can be seen in Max Scheler who offered a theory of layers of feeling to which different kind of values correspond. Cf. Scheler 1921, Geniusas 2015.

<sup>827</sup> Presumably, rather than bringing an additional feeling of self-worth about that would stand, as it were, separately next to the basic felt self-value, reflection and socially-embedded phenomena modify the dimension of felt self-value as a whole.

<sup>828</sup> Kreuch 2019.

<sup>829</sup> Kreuch 2019, 158.

<sup>830</sup> Kreuch 2019, 158.

<sup>831</sup> Kreuch 2019, 157.

<sup>832</sup> Kreuch 2019, 158.

<sup>833</sup> Kreuch 2019, 37, 38.

interest that discloses one's own existence. It is this particular aspect of the feeling of being oneself that has been neglected by Zahavi. Both Kreuch and Zahavi admit (and I would agree) that any phase of experience always involves some experiential content. But I disagree that any phase of self-experience needs to involve some *self*-content in order to present a case of self-experience at all<sup>834</sup>. In my view, once again stressing the Lockean and Lotzean nature of my proposal, it is the experience of the possibility of pleasure and pain intrinsic to the experiential process that is felt as a self-concern. This is what discloses my individual existence: the experience of being oneself as the one whose life is at stake *here and now*. And this feeling of being oneself, I claim in line with Zahavi, does not hinge on a specific content and so arises whenever there is some experiential content without which there would not be any experience to begin with<sup>835</sup>.

Kreuch believes it is self-feeling that provides the content for an otherwise empty and formal structure of selfhood: "self-feeling fills the formal structure with material content and makes your way of being distinct from anyone else's"<sup>836</sup>. Yet I argue that it is *not* the fact that self-feeling, as specific content, is qualitatively distinct from other aspects that constitute the individuality of my existence. Rather, it is the mere fact that I am affected by experiential content. This engenders a feeling of being the one to whom something is given, i.e., an experience of my individual existence.

Thus, I suggest the temporal self-manifestation of the experiential flow, the basic 'for-me-ness' of experience (in Zahavi's terms), the experience of some content, and a feeling of being oneself are all intrinsically linked with one another. Whenever something is experientially given to me, it means some content is given *for me*. This consists of the auto-appearance of the flow, which elicits a feeling of being oneself. The following quote further illustrates the point:

I feel myself standing in this current now. Thereby, I find myself in continuous transition, not only with regard to the past, but also with regard to what is coming. In transition, I don't find myself sharply delimited, in no way tearing off, as if there were breaks and I would be restarting. I am a being that flows, which is also a flowing towards the coming.<sup>837</sup>

---

<sup>834</sup> Cf. Kreuch 2019, 37f.

<sup>835</sup> Note that, here, I am only saying that to have a feeling of being oneself at all doesn't hinge on a specific content. That doesn't mean experiential content wouldn't have an impact on the feeling of being oneself and its concrete phenomenal how. It surely has, as I will further elaborate in Section 7.1.6. But here the point is only that the feeling of being oneself emerges not due to a specific content but owing to the mere fact that *some* content is given to *me* at all, i.e., that there is experience at all.

<sup>836</sup> Kreuch 2019, 166.

<sup>837</sup> Volkelt 1925, 21f., my translation: "Indem ich mich als in diesem Jetzt stehend fühle, finde ich mich nicht nur nach der Vergangenheit hin in stetigem Übergehen, sondern auch nach vorwärts hin finde ich mich als keineswegs scharf begrenzt, als keineswegs abreißend, aufhörend und nach einer Pause neu anfangend. Auch nach vorwärts hin bin ich ein Weiterfließendes."

It is the consciousness of time and the temporality of consciousness that amounts to both the selfhood or ‘for-me-ness’ of experience as well as the feeling of being oneself. In contrast, Kreuch declines to elucidate the inextricability of time, self, and feeling. For him, the ‘for-me-ness’ of experience and self-feeling do not seem to hinge together. Kreuch sees Zahavi’s ‘for-me-ness’ as only a formal feature of experience that, he admits, forms part of any experience: “[n]o matter what experience you have, it is immediately clear that it is you having this experience” as “every experience is had ‘minely’”<sup>838</sup>. This includes self-feeling, which for Kreuch is one type of experience among others: “thus, the formal feature of ‘mineness’ applies to many kinds of experiences, including self-feeling”<sup>839</sup>.

But if Kreuch puts it like that, it seems he sees no direct connection between either (1) selfhood and self-feeling or (2) temporality and self-feeling. In fact, it does not really come as a surprise when he concedes on the last page of his book that he did not sufficiently address the aspect of temporality and selfhood. He notes that “he did not say all too much about how self-feeling relates to the temporality of our existence or our consciousness of time” and that “the relation between self-feeling and ‘self’ ... was only superficially discussed”<sup>840</sup>. To capture the most basic form of self-feeling, I think we need to acknowledge an intrinsic relationship between the temporal self-manifestation of the experiential flow and the selfhood of experience that comes with a feeling of being oneself<sup>841</sup>. In my view, it is the self-awareness of the experiential process that constitutes the being of selfhood (self-experientialism) and to which an awareness of one’s existence belongs—a feeling of being oneself. This feeling is *part* of the temporal self-manifestation inasmuch as one’s future, originally experienced in terms of protentional consciousness, is not experienced in a neutral way. There is always self-concern about what is to come regardless of what specific content is given to me and what I take to be my possible future. It is not only that feeling *qua* experience is temporally structured, but that the temporality of consciousness—my existence as experiencer—involves intrinsically feeling.

### 7.1.5 Openness and involvement: Egological participation through feeling

With this emphasis on the intrinsic-feeling character of the experiential process and its temporal self-manifestation, my proposal on the feeling of being oneself could seem to

---

<sup>838</sup> Kreuch 2019, 157.

<sup>839</sup> Kreuch 2019, 158.

<sup>840</sup> Kreuch 2019, 252.

<sup>841</sup> Kreuch’s rich book includes many interesting insights. However, when it comes to the most basic form of self-experience and self-feeling, I think his approach lacks a clearer explanation of why self-feeling is a *self*-feeling. To account for this, I think we need to do more than acknowledge that “the two accounts [self-feeling and Zahavi’s minimal self, P.S.] are both true and complement each other” (Kreuch 2019, 158). We need to integrate the notion of self-feeling into the minimal self account, as I have tried in the last chapter.

approach a rather classical phenomenological proposal for self-feeling: Michel Henry's notions of "auto-affectation"<sup>842</sup> and "radical immanence"<sup>843</sup>. Henry endorses these notions on the grounds that feeling—and more specifically, *self*-feeling—underpins all experience and self-experience. He further defends a position similar to the view of self-experientialism that I developed in the second part of the dissertation<sup>844</sup>. It is therefore not surprising that whenever Zahavi formulated a more metaphysical claim about self-experience constituting the being of selfhood, he refers to Henry<sup>845</sup>. Like Lotze and his successors, Henry distinguishes between sensation (*Empfindung*) and feeling (*Gefühl*); like Husserl, he also differentiates affection from feeling. Using different words, Henry introduces the distinction under the labels of "sensibility" and "affectivity"<sup>846</sup>. For the former, he refers to the phenomenon of "feeling something...through the intermediary of a sense and ultimately of the internal sense"; for the latter, he invokes the phenomenon of "auto-affectation", i.e., "the form of the essence wherein the essence is affected not by something else but by itself"—thus becoming a "feeling of self"<sup>847</sup>.

Yet Henry's proposal is not simply different in terminology only; it also involves substantive differences. Unlike the view I develop in line with Lotze and Husserl, Henry does not see the distinction as one of different (though intertwined) moments within a unitary structure for experience. Rather, he claims that both phenomena involve distinct structures: "[a]ffectivity has nothing to do with sensibility, with which it has constantly been confused, but it is rather structurally heterogeneous to sensibility"<sup>848</sup>. I agree with Henry that self-feeling is phenomenologically distinct from any sensory phenomenon (e.g., a sensation such as 'red' or 'hard', etc.). But distinguishing between sensation and feeling does not necessarily require prioritizing one over the other, as if one were independent of the other but not vice versa. Yet this seems to be Henry's position.

For Henry, the structure of affectivity or auto-affectation occurs prior to the structure of sensibility. It exists before the existence of something is received through the senses or internally:

---

<sup>842</sup> Henry 1973, 467.

<sup>843</sup> Henry 1973, 468.

<sup>844</sup> See, for instance, Henry 1973, 465: "That which senses itself, in such a way that it is not something which feels but the very fact of thus sensing itself, in such a way that its 'something' is constituted as follows: the self-feeling of self, the experiencing of self, the being-affected by self, this is the Being and the possibility of self." Moreover, he emphasizes that the "feeling of self ... constitutes the Being of the Self" (Henry 1973, 465).

<sup>845</sup> See, for instance, Zahavi 2006, 71, 106, 116, 125, Zahavi 2014, 18.

<sup>846</sup> Henry 1973, 463.

<sup>847</sup> Henry 1973, 463.

<sup>848</sup> Henry 1973, 463.



[T]he auto-affection in the immanence of pure affectivity, this is the Being-self of the subject as affective and concrete Self, the original Self of affection which as such makes all affection possible, even sensible affection.<sup>849</sup>

He believes the ipseity of auto-affection, the fundamental affectivity wherein ‘that which affects’ and ‘that which is affected’ coincide perfectly, is the only reason something *else* can appear to begin with:

[I]t is this Self..., which forms the opposition, in such a way that it is from it, from a Self, that the opposition opposes what it opposes; it is to it that the opposed-Being presents itself, in such a way that it is the Self which receives what can be received only by a Self, thus rendering possible every opposition and every reception in general and their identity as well.<sup>850</sup>

In other words, ‘I can only be affected by something else because I already am affected by myself’.

Although the Husserlian position by Zahavi (the view I have defended throughout the thesis) sounds similar, it is actually quite different from Henry’s take. Zahavi’s view consists of the idea that primary self-affection occurs *in the context* of affection through something other<sup>851</sup>. This is the idea that presumably gave rise to the title of his work *Self-Awareness and Alterity* (1999). My proposal modifies Zahavi’s account but remains loyal to it. As I have repeatedly stressed, self-feeling is part of the temporal self-manifestation of experience but does not underlie it as a precondition. Such a precondition would be external to the temporal structure of consciousness, as Henry seems to suggest: “Far from being able to give basis to the essence of ipseity, the internal sense presupposes it (the Self of auto-affection, P. S.) as that which makes its very structure possible”<sup>852</sup>. For him, inner time-consciousness is not the structure of primary self-awareness. Instead, self-awareness in terms of auto-affection makes possible inner time-consciousness through which all possible content is originally given.<sup>853</sup>

Auto-affection is both fully heterogeneous to the internal sense *and* a presupposition for it. Henry construes self-feeling as something that must be present before the whole temporal process of experiential manifestation can even be established. He thus insinuates that self-feeling is independent of the experiential process and its phenomenological structure. But issues similar to those faced by Kant soon emerge. Kant considered the possibility of an ‘intellectual intuition’ of the ‘I’, i.e., the possibility of self-awareness prior to any empirical consciousness or a self-awareness prior to and independent of all content.

---

<sup>849</sup> Henry 1973, 467.

<sup>850</sup> Henry 1973, 467f.

<sup>851</sup> Zahavi 1999, Ch. 7.

<sup>852</sup> Henry 1973, 468.

<sup>853</sup> Cf. Staudigl 2012, 345.

Before I discuss the issues emerging from Henry's proposal, the implicit Kantian setup is as follows. Henry asks:

Is it not here sufficient, when the problematic finally gets down to the absolutely fundamental ontological dimension to which it appeals, to read in it, in auto-affection, that which it is, i.e. the essence which constitutes auto-affection and makes it possible? The simplest experience, namely, that which it inaugurates before the ecstasy and in it, the immediate experience of self, the original feeling which the essence has of itself, can this not be recognized and grasped?<sup>854</sup>

The Kantian question that is hidden in this quotation is: Can we recognize and grasp the structure of self-awareness? After all, it ought to be essentially prior to the ecstasis of time, i.e., independent of the internal sense. So, is this possible? As discussed in Chapter 3, Kant's answer is negative. He believes the true nature of pure self-awareness independent of all empirical content necessarily remains elusive to us. But Henry agrees with Fichte, who believes we *can* recognize and determine the kind of consciousness that corresponds to an immediate and pure self-awareness. Where Fichte famously refers to an intellectual *activity* (a form of absolute self-conscious spontaneity) in and through which the 'I' posits itself, Henry detects a fundamental *passivity* of self as the ground for all consciousness:

The internal structure of immanence was ultimately understood and described as the passivity of Being with regard to itself, as original ontological passivity. Herein, in its structure, is defined in a strict and determined way what ultimately is Being itself, how, bound to itself, deprived of the power of breaking this bond, deprived of every power concerning its own Being (for example, that of positing it, of willing it or understanding it, of surpassing it in any way whatever) hence not free and as such essentially passive, passive with regard to itself, it coheres with self in the absolute unity of the adequate experience which constitutes it and makes it what it is.<sup>855</sup>

For Henry, pure self-awareness (i.e. self-awareness in radical immanence) is one of perfect coincidence that is passive throughout. Crucially, this passivity motivates his view that the original self-awareness is the affective—or feeling, in my terminology—kind. Feeling or affectivity is the most basic ground for consciousness because consciousness from which any possible appearance emerges is inherently passive. This self-passivity, which corresponds to the perfect self-identity, exists prior to any temporal self-differentiation.

While my account is similar to Henry's in many ways, they thus amount to very different proposals. I believe Henry's approach confronts a number of issues that, from my view, render it unconvincing. Let me discuss the two most important ones. First of all, I

---

<sup>854</sup> Henry 1973, 462.

<sup>855</sup> Henry 1973, 468.

submit that he runs into what I take to be an unresolvable dilemma. He must take auto-affection as existing prior to any temporalization and as something non-phenomenal or unconscious. Accordingly, it can be understood as something indeed distinct from what Henry (and Kant) call the internal sense. But in this case, the kind of self-givenness or self-familiarity auto-affection ought to present remains elusive. Even more problematically, something that itself is non-phenomenal could hardly be the ground for phenomenal consciousness. Henry cannot and does not maintain that auto-affection is something underlying the ground for consciousness.<sup>856</sup> Alternatively, Henry can try to defend the view that auto-affection *is* the ground of all givenness. It would so form a *phenomenal part* of the process of givenness. This indeed seems to be his position, as the following quotation in his later *Phénoménologie matérielle* (1990) reveals:

This self-givenness...is structurally different from 'relating-to'. It is not itself a 'relating-to' but insurmountably excludes it from itself. It is not outside of itself but in itself, not transcendence but radical immanence. And it is only on the basis of this radical immanence that something like transcendence is possible. Seeing is actualized only as a nonseeing... . This nonseeing, this unseen, this invisible, is not the unconscious. It is not the negation of phenomenality but its first phenomenalization. It is not a presupposition but rather our life in its non-ek-static but yet undeniable pathos.

Yet if Henry does indeed believe that auto-affection phenomenologically forms part of the experiential process and so enables it to begin with (*without* being a pre-supposition), why and how could he then maintain that auto-affection is throughout external and independent of temporally structured experience? Regardless of whether this is the case, the picture of a possible soul-like self-familiarity untouched by any 'otherness' emerges. Henry seems to suggest that auto-affection is a possible consciousness that can be separated—not only abstractedly, but also ontologically—from the experiential process unfolding temporally. But just as I emphasized in the discussion of Kant, the sheer 'for-me-ness' without anything else that is *for me* makes no sense either semantically or phenomenologically. That would certainly be an empty and formal abstraction.

Secondly, Henry describes it as auto-affection to overcome the empty and formal character of a purely immanent, perfectly coinciding self-awareness—and to offer a positive determination thereof. Owing to its passive structure, auto-affection ought to consist of feeling. But is this connection between passivity and feeling plausible in the context of an allegedly non-ek-static pure immanence? I do not believe it is. The connection between passivity and feeling is one we know from the kind of experience that *is* temporally

---

<sup>856</sup> He would run into the same problematic as I argued in Chapter 3 Kant does. But while Kant had tried to conceive of pure self-awareness as a *spontaneous activity of thinking* that is throughout non-empirical, Henry tries to conceptualize pure self-awareness in terms of a *passive affectivity* independent of any empirical givenness.

structured. If there is passivity that elicits feeling, it is the kind of passivity that we undergo in the temporally structured experiential process. Regarding this second point, there are several reasons why.

For one, we undergo the experience of being presented with something whether we like or not: this is the passivity of the receptivity of any experiential manifestation of  $x$ . And this includes the self-passivity of self-affection that Henry seems to be targeting, or the passivity of not being able to separate oneself from one's own process of experiencing. In other words, 'I have no say in whether or not  $x$  is given to me'. This is a passivity, of course, in regards to  $x$ . But by the same token, there is also a passivity in regards to myself. My being essentially consists of experiencing some  $x$  (if not  $x$ , then  $y$ ). I have absolutely no say in the fact that my being is the experiencing kind. All of my activities extend from my experiences, but the fact *that* my existence is an experience does not extend from any of my activities.

For another, it is precisely our own temporal being that we cannot escape or transcend in any way. Self-passivity *par excellence* is temporal in that I cannot avoid my anticipated future. Whatever is to be given will be given to me. It is the temporal horizon that reveals my inability to evade the experiential process, which (according to self-experientialism, which Henry shares) constitutes my very being in turn. To say that I am passive in regards to myself means to say that I am passive in regards to the temporal process of experience. At the very least, that could be a way to make sense of Henry's notion of fundamental passivity in regards to oneself. But contrary to his intention, this self-passivity is ecstatic. A non-ek-static (and thus pre-temporal and independently conscious) self-passivity is, I suggest, a construction. Moreover, *why* the passivity of sheer, non-temporal self-coincidence should give rise to feeling to begin with is unclear in this construction. Every reason to connect passivity with feeling is owed to a passivity that is temporally structured and thus stripped from the notion of a non-temporal self-passivity. In contrast, it can be more easily understood why the temporal process of experience and the passivity I experience within it elicits feeling: 'I am passive throughout the temporal unfolding of *this* experiential process, which is essentially my own being and my concrete and individual existence. It is *this* experiential process in which pleasure and pain are possible experiences and, intrinsically, possibilities for me—possibilities that in themselves are never experienced as neutral'.

What broader conclusions can be drawn from such a rejection of Henry's proposal of a pure, felt self-coincidence in radical immanence? The antithesis of Henry's view is a notion that self-feeling doesn't just occur on its own, in an enclosed sphere of selfhood as it were. Instead, it emerges from a fundamental openness: openness towards the givenness of  $x$ , the fact that experience means the capability of being affected by something else; and openness towards one's own future (which is precisely the experience of being capable of

affection). I suggest it is in protentional consciousness that we experience both the possibility of being affected by something else *and* our own future as experiencing beings. And, it is this fundamental openness that is originally experienced in feeling. Following this line of reasoning, feeling is what involves us in the world. It is the way through which we participate in the process of manifestation—how we have our existence in experience.

### 7.1.6 Variability of the feeling of being oneself: Scope and graduality

In the preceding sections, I offered the following specifications for the feeling of being oneself. It is:

- how we experience our existence
- a reference to our existence as experiencing beings
- how we experience the value of our existence as experiencing beings
- originally the experience of the possibility of pleasure and pain
- not the presentation of a certain kind of experiential content, but connected to protentional consciousness
- not independent from all experiential content, but forms part of the temporal horizon of experience from which experiential content is given
- a presentation of how we are open to and involved in the world
- shaping how we participate in the process of manifestation.

But there is one aspect of the feeling of being oneself that I have not yet addressed in detail: its concrete, phenomenal *how*. It is clearly an important aspect, which gives rise to a few related questions. Is the feeling of being oneself universal and present in any moment of consciousness? Can it vary depending on experiential context? What is the scope of the feeling of being oneself? Does it only concern myself insofar as I am the experiencer of my experiential life? Or does it extend to myself insofar as I am a member of a historically and culturally shaped world shared with others? This dissertation cannot provide a full-fledged account for the feeling of being oneself in all of its aspects, but my proposal so far has already implicitly touched upon these questions. So, some tentative answers can be made explicit.

I stressed that self-feeling neither corresponds to nor originates in an allegedly perfect self-coincidence that ought to lie before any affection by something else. I argued that self-feeling only arises in the context of the givenness of experiential content *x*. Self-feeling thus cannot be isolated and considered separately from what is experientially given or how it is given. I further argued that the feeling of being oneself forms part of the temporal horizon of experience in terms of protentional consciousness, or the consciousness of one's possible and anticipated future. Given this, it should be clear that the feeling of being oneself cannot be a specific feeling *x* that is universal and present in every moment in our experiential life in the same manner. It is true that I am anticipating

myself at every moment in the sense of the continuation of my existence in experience when protending what will be given to me in the moment to come. But experiential content varies continuously, and so does protentional consciousness. What could possibly be my own future changes the whole time. Although my anticipated future is always one of experiential kind, exactly what I anticipate coming is never the same. Therefore, I suggest, self-feeling, the feeling of being myself, varies continuously depending on what I take to be possible for me.<sup>857</sup>

Crucially, I do not think it is possible to subtract and isolate a *general* feeling of being oneself (being an experiencer) from a more *specific* feeling of being myself (being this particular individual experiencer). We do not have two self-feelings, one general and one specific. When I introduced the feeling of being oneself, I did not mean a specific token of feeling that is omnipresent in experiential life. Rather, I was referring to a type of feeling that is omnipresent in experiential life. I aim to establish the view that the feeling of being oneself is an indispensable dimension. But the concrete phenomenal ‘how’ of self-feeling may vary depending on current life context and who one experiences oneself to be.

Let me further illustrate the variability in and relevance of the feeling of being oneself by referring to an account for self-experience that does not explicitly propose self-feeling: Galen Strawson’s phenomenological distinction between what he calls Episodics and Diachronics. In order to make its case, his proposal needs to refer to self-feeling. In Strawson’s view, “[w]e lie on a long psychological spectrum which stretches from strongly *Diachronic* people or *Endurers*, at one end, to strongly *Episodic* people or *Transients*, at the other”<sup>858</sup>. As he further emphasizes, people are usually not fixed on any position of this continuum<sup>859</sup>. Sometimes, we may consider ourselves as one and continuous with our past selves. In other moments of our life, our sense of identity may be more punctual and be structured in the way of Episodics:

They experience themselves\* [= the self that is present in the moment at stake, P.S.] as something that was not there in the further past and will not be there in the further future (‘further’ is intentionality indeterminate).<sup>860</sup>

I agree with Strawson that experiences of this kind are quite common to most people. There can be and often are moments in life where someone can be puzzled about why she has ended up in some city or another, or why she started some career rather than another. In such moments, she may well remember how it had happened that these life decisions were

---

<sup>857</sup> This, I believe, is the fundamental truth that lies in Ratcliffe’s Heideggerian approach of existential feelings. How I find myself in the world is depending on what I experience to be possible for me, and it is *this* what determines what or better who exactly I find myself to be.

<sup>858</sup> Strawson 2017, 96.

<sup>859</sup> Strawson 2017, 96.

<sup>860</sup> Strawson 2017, 96.

made. She may even find herself capable of “empathic access”<sup>861</sup> to her past selves, thoughts, and feelings.

And yet, she might also have a sense that she *is* not one of those past selves. Nor, for that matter, is she any one of those selves that might inhabit her future. Such a punctual existence can occur during normal daily life, such as when taking the tram and trying to relax or noticing how tired she is. It can happen during a party at which someone literally just lives in and for the moment while talking about being around others. Most people have experienced leaving a party alone for a while to go to the pool, for instance, or the beach, the garden, or the streets. In these moments, some might ponder their lives, reflect on the party, or try to uncover their next steps in life. But some might simply wonder how strange and absurd it is that they even exist, cherishing a little moment of solitude in which they experience their life as consisting of just that very moment.

What is it, exactly, that renders such an experience either diachronic or episodic? Even though Strawson neither explicitly answers nor poses this question, his deliberations reveal it is feeling that determines the experience one undergoes. In terms of episodic or diachronic life, one might think the difference lies in a different temporal horizon whereby episodic life is simply narrower than diachronic life. For the latter, events and happenings in the past may feature in the background of experience. However, the former is determined by aspects of the present situation. Whereas Diachronics recall much of their past, Episodics are less aware of the happenings that preceded the current situation. While Strawson considers himself an Episodic who indeed has a tendency to recall only a little of his past, he emphasizes that memory is *not* a decisive factor:

Neither inconsistency nor poor memory is necessary for the episodic experience of life. John Updike writes ‘I have the persistent sensation, in my life and art, that I am just beginning’. These are the words of a man who has an extremely powerful personal memory and a highly consistent character. I have the same persistent sensation, and learn from Updike that it is nothing essentially to do with my extremely poor personal memory.<sup>862</sup>

This is not to say he has no memories at all. The point is that the way he remembers his past is quite different from the way Diachronics do. The difference lies not in the content of memory, or whether memories themselves are even recalled. Rather, it lies in the way he *feels* about them:

In fact, I am strongly inclined to say that the events in question didn't happen to me—to me\*, i.e. to that which *I feel myself to be*, in having self-experience—at all. These memories are of course distinctive in their ‘from-the-inside’ character, and they certainly happened to the human being that I (also) am; but

---

<sup>861</sup> Schechtman 2003.

<sup>862</sup> Strawson 2017, 31, my emphasis.

it simply does not follow that they present, or are experienced, as things that happened to me\* just characterized.<sup>863</sup>

What Strawson expresses here is that the person I experience myself as being depends on how I feel *about* what I experience. In this case, my feelings involve the memories that I recollect. If Strawson's descriptions of the episodic and diachronic self-experience are correct and we indeed undergo these variations in how we experience ourselves, then I suggest we should also acknowledge that these variations exist in the feeling of being oneself. Accordingly, the distinction between Episodics and Diachronics illustrates the variability of the feeling of being oneself. I argue that those who find Strawson's distinction persuasive accept by the same token (even if only implicitly) differences in the feeling of being oneself. Therefore, they accept something like the feeling of being oneself to begin with.

The following section further corroborates my proposal by defending the view that we should accept that there are feelings of being oneself that can vary. Additionally, the feeling of being oneself is never absent—even in extreme psychological conditions such as the psychopathological phenomenon of depersonalization, and even if maximally altered. In fact, depersonalization is interpreted as an experiential condition characterized by a complete lack of either self-experience or self-feeling. But I will argue that depersonalization is misconstrued if understood in terms of a lack of self-experience or self-feeling, lending support to the proposal offered here.

## 7.2 Minimal self, the feeling of being oneself, and depersonalization

Although the phenomenon of depersonalization was described in the second half of the nineteenth century, the best way to describe and conceptualize the pathological self-experience involved is still subject to debate. In this section, I discuss two lines of conception that can both be said to challenge my proposal of self-experience and the feeling of being oneself.

### *1) Depersonalization as a lack of subjective character*

The first line of conception is a recent offering. It suggests that the phenomenology of depersonalization lacks a subjective character. Different experiential symptoms associated with the condition, such as feelings of not being there, can be traced back to the loss of a sense of self. Such a conception might challenge both Zahavi's view of the minimal self

---

<sup>863</sup> Strawson 2017, 51.



and my notion of a feeling of being oneself; a central claim holds that the experience of depersonalization unfolds in an anonymous manner, giving rise to feelings of non-existence.

## 2) *Depersonalization as a lack of self-feeling*

More classical accounts, by contrast, have argued that depersonalization is primarily an emotional disorder. A subjective character is not completely absent in depersonalization, but the feeling of life or self-feeling goes missing in depersonalization. Such a conception is compatible with Zahavi's minimal self-view but still calls into question the idea that experience always involves the dimension of self-feeling. If depersonalization is characterized by a loss of self-feeling, this lends support to the notion that self-feeling is *not* an essential aspect of experience.

I believe both conceptions have their merits, but I ultimately argue that both fail to properly account for the phenomenology of depersonalization. I see the first conceptualization of depersonalization as too crude in that it denies all subjective character within the experience of depersonalization. The second, more classical approach to self-experience in depersonalization is more nuanced and conservative than the first. But in my view, it still doesn't get the picture right. Instead of describing depersonalization as a loss of self-feeling, I see 'feelings of non-existence' (often reported in depersonalization) as still part of the feelings of being oneself. What we see in depersonalization is an *alteration* of self-feeling rather than a complete absence of self-feeling. Rejecting both lines, I develop my own approach to the phenomenology of depersonalization that I then compare with Zahavi's interpretation of the disorder. As it turns out, our approaches relate to each other like two sides of the same coin. But first, let me introduce the phenomenon in question: What is depersonalization?

### 7.2.1 Depersonalization as a loss of self: An anti-thesis to the minimal self-view?

Depersonalization consists mostly of a pathological modification of self-experience. Here, "the individual feels that his or her own feelings and/or experiences are detached, distant, not his or her own, lost, etc."<sup>864</sup>. This typically involves a severe alienation from one's own body, memories, and environment—all associated with a form of emotional numbing.<sup>865</sup>

---

<sup>864</sup> World Health Organization 1993, ICD-10, F48.1.

<sup>865</sup> As Sierra & David (2011) and Sierra & Berrios (2001) point out, these four symptoms are stable and omnipresent in first-person reports to be found between the first descriptions of depersonalization in the late 19<sup>th</sup> century until today.

I see but I don't feel. I taste but it means nothing to me. ...I can't admire a painting—it affords me no pleasure. I seem to be walking about in a world I recognize but don't feel. I saw Big Ben alight last night, normally a moving sight to me, but it might have been an alarm clock for all I felt. It all makes me feel so lonely and cut off. It's the terrible isolation from the rest of the world that frightens me. ...My husband and I have always been happy together, but now he sits here and might be a complete stranger. I know he is my husband only by his appearance—he might be anybody for all I feel towards him. That's what worries me most of all.<sup>866</sup>

Emotional numbness does not necessarily mean the loss of all capacity for feeling in general.<sup>867</sup> But it often consists of a disturbing condition in which one experiences sensations as not really being part of oneself, or rather oneself not as partaking in the sensations one undergoes. For that matter, one does not partake in any kind of experience. As several subjects describe,

It was as if it was not me walking, it was not me talking, as if it was not me living ...I can look at me, I am somehow bothered by my body, as if it wasn't me, as if I lived on the side of my body, on the side of myself if you like. I don't know how to explain.<sup>868</sup>

I was completely unable to tell whether I was still present or whether I was the part that was gone.<sup>869</sup>

Feeling estranged from one's own experiences affects how one's own body is felt. This can translate into detachment from others and the world in general<sup>870</sup>. Such self-alienation and detachment may also engender a more global subversion of a sense of reality culminating in 'derealization'<sup>871</sup> wherein "I'm not part of anything and so nothing seems real"<sup>872</sup>. As this quote illustrates, a kind of self-loss may form the core of depersonalization. This then spreads out and extends to a loss of connection with the environment and reality. The question that interests me is how best to characterize the self-loss that is seemingly at the core of depersonalization. More specifically, does depersonalization present a case in

---

<sup>866</sup> Bockner 1949, 968.

<sup>867</sup> Sometimes there can be complaints of patients "that they are capable of experiencing neither pain nor pleasure, love and hate have perished with them" (Schilder 1928, quoted in Ackner 1954, 843). However, sometimes it is also emphasized that "[p]atients still feel physical pain, but may report being indifferent to it, oddly cut off from it" (Colombetti & Ratcliffe 2012, 145). Moreover, fear and anxiety seem to be prevalent in persons with depersonalization (Colombetti & Ratcliffe 2012, 148).

<sup>868</sup> Janet & Raymond 1898, 70.

<sup>869</sup> Roberts 1960, 481.

<sup>870</sup> Cf. Colombetti & Ratcliffe 2012.

<sup>871</sup> World Health Organization 1993, ICD-10, F48.1

<sup>872</sup> Anonymous patient in Ackner (1954, 866).

which the minimal self is absent? And if so, wouldn't that significantly undermine Zahavi's case for the latter?

### 7.2.2 Billon and Guillot: What's left of the sense of self in depersonalization?

To assess the extent to which depersonalization may present a challenge for the minimal self-view, I discuss two recent approaches that consider the minimal self absent from the psychopathological condition. These approaches correspond with the first line of conception, which sees depersonalization as a lack of subjective character.

In a recent article, Alexandre Billon asks whether there is “a common phenomenological core, an ‘experiential nucleus’, that would explain why they [depersonalized patients, P.S.] feel the way they do”<sup>873</sup>. He suggests that there is indeed a feature common to the experience of depersonalization: lack of the *subjective character* of experience. In his view, this “lack of subjectivity”<sup>874</sup> can explain the core symptoms found in depersonalization: disembodiment, emotional numbness, and “death and nonexistence experiences”<sup>875</sup> (and detachment from the world, in derealization). This absence is further detailed as follows:

- *Disembodiment*: “if my bodily sensations didn't clearly feel mine, it could seem to me as if they were not my bodily sensations and as if the bodily part in which I feel them were not my bodily part”<sup>876</sup>;
- *Emotional numbing*: “If the subjectivity recedes from their emotions, they might feel as if *they* had no emotions at all”<sup>877</sup>;
- *Experience of nonexistence*: “[I]f someone underwent a sweeping withdrawal of the subjective character of his experiences, he would cease to properly *feel himself* and that his awareness of himself would be severely impaired. ... [I]t is precisely this kind of impairment of self-awareness, caused by a deficit in subjectivity, that the feeling of nonexistence reflects.”<sup>878</sup>

According to Billon, then, alteration of the general experiential character underlies the alleged loss of self and all that is associated with it.

Bearing in mind the various notions that are usually associated with the term ‘subjective character’, the question remains open as to what exactly is meant with a ‘lack of subjective character’ and a ‘lack of subjectivity’? Guillot suggests we consider her

---

<sup>873</sup> Billon 2016, 373.

<sup>874</sup> Billon 2016, 375.

<sup>875</sup> Billon 2016, 375.

<sup>876</sup> Billon 2016, 375.

<sup>877</sup> Billon 2016, 375.

<sup>878</sup> Billon 2016, 377.

tripartite differentiation of ‘subjective character’, according to which we must distinguish between ‘for-me-ness’, ‘me-ness’, and ‘mineness’:

While I find the general direction of Billon’s argument very compelling, I think that the three-way distinction I propose in lieu of the misleadingly monolithic notion of subjective character might allow for a more conservative, and perhaps more intuitive, description of the pathology.<sup>879</sup>

Against this background, she describes depersonalization as a psychopathological “condition in which experience lacks both *me-ness* and *mineness*, but retains *for-me-ness*”<sup>880</sup>. That is, “*something* of the subjective character of the abnormal experience”<sup>881</sup> is not lost and in light of this, Billon’s view seems overly strong. Even so, both Billon and Guillot stress that what is missing in depersonalized experience is a sense of self. Their disagreement could thus be construed as mostly terminological. Both would accept that patients with depersonalization surely keep on having experiences of which they are aware, even though they may feel cut off from them. But I suggest that with Guillot’s reformulation, Billon’s claim becomes quite strong. In other words, Guillot’s precision reveals how strong Billon’s claim really is.

To see this strength, we must set aside Zahavi’s understanding of ‘for-me-ness’ for a while. For Zahavi, ‘for-me-ness’ already involves a minimal sense of self. But for Guillot, ‘for-me-ness’ signifies an experiential state in which there is no sense of self at all. It is a completely anonymous type of experience and therefore non-egological throughout (to use Zahavi’s terms). So, is that really a more intuitive conceptualization of the phenomenon of depersonalization? Are we dealing with a complete ‘lack of subjectivity’ here? Despite some merits, this interpretation of depersonalization ultimately does not sit well with its phenomenology. To illustrate, it is helpful to compare the kind of experience we might find in non-human animals with the experience of depersonalization. If a guinea pig or a dog has phenomenal consciousness then, according to Guillot’s tripartite framework, such an experience lacks ‘me-ness’ and ‘mine-ness’. In lacking reflective capacities (to say nothing of conceptual mastery), a guinea pig or a dog cannot make itself the object of awareness or of its experience as its own. But since for such an animal, as experiencing, it will somehow be like to undergo its experiences, in Guillot’s view, the animal’s experiences would involve ‘for-me-ness’. Accordingly, the profile for subjective character within the experience of depersonalization (in terms of Guillot’s reformulation of Billon’s approach) would be the same as with non-human animals. Both lack any sense of self because the

---

<sup>879</sup> Guillot 2017, 42.

<sup>880</sup> Guillot 2017, 42.

<sup>881</sup> Guillot 2017, 42.

subjective character, in these cases, consists of nothing more than the awareness of experience or ‘for-me-ness’. Now, there seem to be two ways to respond to that.

- 1) One could emphasize that in the experience of conscious non-human animals really there is also more than merely for-me-ness, i.e., some kind of me-ness, however construed. And, that saying that, in the experience of conscious non-human animals, there is only for-me-ness, is a wrong depiction of their experiential life.
- 2) One could emphasize that for a depersonalized human person it means something different to undergo experiences that only bear for-me-ness compared to the experience of a non-human animal that bears only for-me-ness.

Clearly, the first response does not help Guillot make her case at all. It seems highly implausible to ascribe a more subjective character to a guinea pig and a dog than plain ‘for-me-ness’ while maintaining that all depersonalization retains is plain ‘for-me-ness’. The second possible response likewise fails to buttress Guillot’s position. If she is forced to concede a difference between human and non-human experience only in terms of ‘for-me-ness’, a question arises as to what this phenomenological difference consists of. Surely the phenomenological difference would concern subjective character. It thus seems she would inevitably be forced to admit that the subjective character of depersonalized human experience contains *something more* than plain ‘for-me-ness’ construed as an anonymous self-awareness of experience. Guillot’s position, therefore, is too strong; it involves a description of depersonalized experience that still cuts out too much of its subjective character.

In fact, numerous first-person accounts from individuals with depersonalization are found in the literature. These suggest that despite admittedly severe changes and loss (of partial aspects) of the subjective character of experience, a good deal of it remains—even though it might have a significantly altered phenomenological profile. Before I elaborate on this more, let me briefly sum up and provide a short preview on my alternative perspective. Billon, I suggest, is right about change in the subjective character of experience. This may explain many of the other core symptoms of depersonalization. But Guillot is right to point out that Billon’s depiction is too radical as some aspects of subjective character are retained even in depersonalization. Yet Guillot’s reformulation is likewise too strong and excessive. Again, I agree that something of the subjective character of experience in depersonalization is lost (or rather modified). However, I do not believe that the depersonalized experience is completely bereft of any sense of self. Hence the question of what exactly is modified and how, along with what is retained, prevails.

My alternative answer to the question is as follows. The experience of depersonalization still incorporates a sense of self, albeit one that is altered. The central

aspect of the alteration of the sense of self consists in a modification of the feeling of being oneself. What is lost—if one wants to speak of depersonalization as a condition in which something is lacking—is the *normal* feeling of being oneself. A lack of the normal feeling of being oneself is not tantamount to the absence of all feeling, but is rather itself felt. The lack or weakening of the normal feeling of being oneself can be decoded phenomenologically as an experience in which pain and pleasure are not simply absent or dampened down. It also involves the experience of a fundamental inability to experience pleasure or pain, and thus to partake in one’s own experiences. Crucially, persons with depersonalization are still aware of the *general* possibility of experiencing pain and pleasure; they only feel it is not possible *for them* as individuals. It is this, I argue, that gives rise to feelings of non-existence—of partaking in the process of experiential manifestation without participating in it. Far from lacking subjective character and/or self-feeling, the depersonalization experience presents us with an extreme kind of self-feeling that involves self-experience (including its essential dimension of feeling).

But before I reach this conclusion, let me first proceed in my discussion and rejection of the two concepts of depersonalization. Consider the first-person reports brought into the field by Billon to demonstrate the lack of subjective character in depersonalization<sup>882</sup>. He gives the following examples of features absent from the language of the patients:

- the “feeling of ‘I’, the experience of ‘me’”<sup>883</sup>;
- a “feeling of myself”<sup>884</sup>; and
- an “awareness of myself”<sup>885</sup>.

All of these features clearly indicate that the subjective character of experience has undergone a significant change. However, it is not clear whether they indeed imply the absence of a *sense of self*. Those who refer to a modification in self-feeling might still be said to have self-awareness. They still ‘know’ they are someone or that they were someone, and that they no longer seem to feel this way. They still might have an awareness of how it was or feels to be someone, yet they do not currently feel this. The pain of their loss of self is a vague awareness that their selfhood does not feel real. One cannot deny the existence of *x* without having a sense of what *x* is and what it would mean for *x* to be existent. This holds true for believing *p* as well as feeling *p*. “Nonexistence experiences”<sup>886</sup>, as Billon calls them, thus presuppose at least a vague sense of self, namely, oneself as non-existent.

---

<sup>882</sup> Billon 2016, 373f.

<sup>883</sup> Simeon & Abugel 2006, 143.

<sup>884</sup> Janet & Raymond 1898, 73.

<sup>885</sup> Krishaber 1873, 171.

<sup>886</sup> Billon 2016, 371.

Accordingly, I suggest reports that refer to a feeling of the loss of self do not support the view of no sense of self at all. Rather, they support the disturbing and extreme feeling of a lack of *life, being, or existence*.

But how about the other expressions? What is a lack of the ‘experience of me’ or ‘awareness of myself’? Aren’t these indicative of lacking a sense of self? A closer examination of the contexts from which these are taken is informative, and does not necessarily support Billon’s reading. The lack of the ‘experience of me’, for example, is from the case of Suzanne Segal. Suzanne is a woman who suffers from what seems to be out-of-body experiences. These consisted of a feeling of location outside the body in a “cloud of awareness”, which she refers to as “the witness”<sup>887</sup>. For some time, she experienced finding herself somehow located at the witness’s position. But then the witness dissolved, which significantly affected her self-experience:

The disappearance of the witness meant the disappearance of the last vestiges of the experience of personal identity. The witness had at least held a location for a ‘me’, albeit a distant one. In the dissolution of the witness, there was literally *no more experience of a ‘me’* at all. The experience of personal identity switched off and was never to appear again.<sup>888</sup>

As this longer quotation shows, the patient is trying to express the loss of a *sense of identity*. Although a sense of identity is a specific sense of self, it is not equivalent to all senses of self. Hence, losing a sense of identity need not amount to a loss of all sense of oneself<sup>889</sup>.

How about the other quotation that supposedly refers to a loss of ‘awareness of myself’? Here, it is important to note that Billon offers a translation of what is originally a French text by Maurice Krishaber. The original version states, “*J’ai perdu la conscience de mon être*”<sup>890</sup>. The longer and more literal translation is ‘I have lost the consciousness of my being’. Krishaber emphasizes on the same page that the same feeling is also sometimes expressed by patients as “*je ne suis pas moi-même*” (literally, ‘I am not myself’). Hence, it is not at all clear that ‘awareness of oneself’ is lost in terms of the lack of a sense of self. Rather, how one is aware of oneself has such an alien character that it almost feels as if one were not there at all.

The following first-person report tries to express precisely this: “It is as though I had transcended personality”, the patient said, “As if ‘myself’ had receded to an image which I regarded objectively, and which is not identified with the whole of me”<sup>891</sup>. There

---

<sup>887</sup> Simeon & Abugle 2006, 143.

<sup>888</sup> Segal 1998, 54, my emphasis.

<sup>889</sup> Recall, for instance, Galen Strawson’s notion of an episodic life compared to a diachronic life (as mentioned earlier).

<sup>890</sup> Krishaber 1873, 171.

<sup>891</sup> Shorvon 1946, 784.

is thus no absence of any subjective character or sense of self. Instead, there is clearly an alteration of the sense of self. This modification concerns one's own existence, or the way one is. The first-person reports referred to by Billon and Guillot do not offer strong support for their view that the subjective character of experience is lacking. They are more supportive of the claim that depersonalization severely undermines the feeling of being oneself, but a sense of self remains. Billon himself is aware of the distinction between subjective character (sense of self) on the one hand, and the feeling of being (the feeling of non-existence in depersonalization) on the other hand. In fact, he argues that the lack of subjective character or sense of self explains the feeling of non-existence:

Now it is arguable that if someone underwent a sweeping withdrawal of the subjective character of his experiences, he would cease to properly *feel himself* and that this awareness of himself would be severely impaired. According to the present interpretation, it is precisely this kind of impairment of self-awareness, caused by a deficit in subjectivity, that the feeling of nonexistence reflects.<sup>892</sup>

I believe there is something fundamentally true about Billon's argument, namely, that the feeling of being oneself ('feel himself') hinges upon pre-reflective self-awareness ('the subjective character of his experiences'). This is precisely what I proposed in the preceding chapter; it is pre-reflective self-awareness that constitutes selfhood. The feeling of being oneself is consciousness of one's own existence.

However, the rationale for his argument seems somewhat flawed or at least insufficient. This can be illustrated by considering Billon's suggestion as a syllogism:

$\neg p \rightarrow \neg q$	If there is 'withdrawal of the subjective character', then a person will 'cease to properly <i>feel himself</i> '.
$\neg q$	Patients with depersonalization do not feel themselves.
$\neg p$	The experience of patients with depersonalization lacks subjective character.

Note that I am not suggesting that Billon is proposing any kind of deduction or inference in a strict sense. But the wrong conclusion demonstrates how pointing to the lack of feeling oneself alone does not suffice to launch the far less obvious and stronger claim that experience lacks all subjective character (and, bearing in mind Guillot's precision, any sense of self). Yet this seems to be Billon's strategy. First, he declares "death and nonexistence experiences"<sup>893</sup> *qua* a symptom and pathological experience of depersonalization. He takes this as an explanandum. Next, he describes the experience of

---

<sup>892</sup> Billon 2016, 377.

<sup>893</sup> Billon 2016, 375.



self in persons suffering from the condition as the following: “It seems to them that they do not themselves have any mental or phenomenal state”<sup>894</sup> but this “will not explain nonexistence experiences”<sup>895</sup>. Unfortunately, Billon fails to specify why he believes so. Instead, he attempts to launch his own explanation for ‘nonexistence experiences’ by claiming they are the outcome of the lack of a subjective character of experience. He then refers to first-person reports that express feelings of non-existence and self-alienation, rather than a lack of the subjective character of experience (or sense of self). In other words, the first-person reports he alludes to do not help elucidate what an experience that lacks any subjective character or sense of self might look like.

Another example is the following first-person report that, according to Billon, shows the lack of subjective character. The patient recounts “seeing life as if it were played like a film in a movie. But in that case where am I? Who is watching the film?”<sup>896</sup> This seems to be an experience in which the sense of self is severely weakened. But still, there remains a dim awareness of oneself as being absent. The awareness becomes even more palpable if one reads the text surrounding the short quotation:

I look at *my mind* from within and feel both trapped and puzzled about the strangeness of *my existence*. *My thoughts* swirl round and round constantly probing the strangeness of *selfhood*—why do *I* exist? Why am *I me* and not someone else? ... *I hear myself* having conversations and wonder where the voice is coming from. *I imagine myself* seeing life as if it were played like a film in a cinema. But in that case, where am I? Who is watching the film?<sup>897</sup>

Considering the context out of which Billon took the report to show the lack of selfhood within the experience of the afflicted person, it becomes abundantly clear that the person is far from lacking all sense of self or undergoing an experience bereft of subjective character. Instead, the person is deeply puzzled by her selfhood and estranged from herself. That surely cannot be equated with a complete absence of the sense of self. Any attempt to construe such an experience as one that is throughout anonymous and selfless is bound to fail. Billon’s and Guillot’s position is thus untenable insofar as they overlook the extent to which a sense of self is still involved in depersonalization. Their strategy of explaining deviant self-feelings involved in depersonalization by emphasizing an alleged lack of subjective character, i.e., the first conception of depersonalization, is not convincing.

---

<sup>894</sup> Billon 2016, 375.

<sup>895</sup> Billon 2016, 375.

<sup>896</sup> Billon 2016, 376.

<sup>897</sup> Simeon & Abugel 2006, 15, my emphasis.

### 7.2.3 Oesterreich's alternative approach: Inhibition of the feeling of life

Let me turn to the second concept of depersonalization. One proxy of this line of reasoning is an alternative account of depersonalization that conceives of the alteration of self-experience in a different way than Billon and Guillot. Based on an elaborate theory of self-experience stemming from a highly-detailed discussion of the different feeling-aspects of self-experience, this alternative account does more justice to the phenomenology of depersonalized experience: Traugott Konstantin Oesterreich's notion of a "feeling of life (*Lebensgefühl*)"<sup>898</sup> or "ubiquitous psychic feeling (*psychisches Gemeingefühl*)"<sup>899</sup> in his *Phänomenologie des Ich in ihren Grundproblemen* (1910). Although the approach dates back to the early twentieth century and has received little attention since, I propose that it is highly informative for the current debate. For one, it provides more plausible descriptions of the experiential condition of depersonalization. For another, it offers distinctions that are helpful for discriminating between and understanding different (albeit interrelated) aspects of self-experience. Despite these merits, I detail below how Oesterreich's proposal must nonetheless be adapted for the sake of relevance.

It is important to note that Oesterreich's terminology differs slightly from contemporary conventions and the way I have talked about 'self', 'I', and 'feeling'. Using his own nomenclature, I present key aspects of Oesterreich's account and then discuss how it can inform my own. Oesterreich's proposal is highly influenced by Lotze and Lipps. It can thus be considered an extension of the self-feeling approach presented in the preceding chapter. Among other things, it offers an application of self-feeling to the case of depersonalization. Like Lotze and Lipps, Oesterreich supports the view that we must distinguish between sensations and feelings. For him, feeling is the experiential dimension through which the existence of 'I' becomes conscious<sup>900</sup>. A central argument revolves around the issue of depersonalization wherein the latter provides both the explanandum as well as the phenomenon with the power to test theories of self-consciousness. The upside of this argument is that if a theory does not fit well with the phenomenology of depersonalization, it can (and should) be dropped.

Applying depersonalization as a kind of litmus test, Oesterreich rules out a set of different approaches to self-consciousness along with the role of feeling therein. He argues that any appeal to sensation or complexions of sensations fails to ground self-feeling, for instance, given that sensations *qua* sensation may persist in depersonalization<sup>901</sup>. The same

---

<sup>898</sup> Oesterreich speaks of the "individual (*individuelles*)" (1910, 317), "central (*zentrales*)" (1910, 322) feeling of life or the "general personality-feeling (*allgemeines Persönlichkeitsgefühl*)" (1910, 317).

<sup>899</sup> Oesterreich 1910, 317.

<sup>900</sup> Crucially, similar to self-experientialism, and contrary to Lotze and Lipps, Oesterreich takes the 'I' to be throughout experiential. Cf. Oesterreich 1910, 232.

<sup>901</sup> Oesterreich 1910, 49, 59.

observation applies to bodily sensations or cenesthesia. Although patients with depersonalization may feel utterly estranged from their bodies, they typically retain bodily sensations or awareness of organic and vital processes.<sup>902</sup> Another example is the idea that a person's experience of the will provides the foundation for consciousness of one's own existence. As Oesterreich points out, a complete dampening down of the will—abulia—need not amount to an experience of depersonalization.<sup>903</sup> One can go through states of severe avolition or weakness of will without ceasing to feel alive, even though dampened will often correlates with lessened feeling.

In contrast, Oesterreich emphasizes, the same cannot be said for feeling:

[T]he core of this condition [depersonalization, P.S.] consists in a general inhibition of the whole *life of feeling*. ... I emphasize that also the central *feeling of life* undergoes such an inhibition.<sup>904</sup>

In this view, feeling is significantly altered to the effect that it almost dissolves. Feelings of death and non-existence then arise. Put differently, it is not the subjective character of experience as such that is absent. Rather, it is only the feeling participation of 'I' (as I called it with Husserl in the preceding chapter). A person may still have normal sensations; what is missing are the feelings that normally accompany these sensations. Deprived of feelings of pleasure and pain, sensations seem awkward and alien. They seem as though not belonging to oneself, because oneself is not partaking in the experience through feeling as one normally does. Oesterreich thus explains the experience of depersonalization as a general obstruction of the feeling character of experience. Obstruction also affects a person's feeling of life, thereby giving rise to feelings of non-existence.

Like Billon, Oesterreich detects a change in the subjective character of experience: The fact that *x* is given to me (whether through sensory perception, imagination, memory, or thinking) does not elicit any feeling in me. The same can be said for any other potential objects and experiences; no 'feeling of myself' being there emerges. On the contrary, it is the lack of feeling in any experience that amounts to the feeling of not being there at all. This is because I am not partaking in any of my experiences. At the same time, Oesterreich avoids Billon's denial of all subjective characteristics of experience in depersonalization:

Even when depersonalized persons often say: they wouldn't possess any 'I' any longer, then this is a deception. Even in complete athymia and abulia there

---

<sup>902</sup> This, among others, is also the reason why I insisted in the last section that vital self-feelings are not all what counts as basic self-feelings. For, in depersonalization, a person may feel 'dead' even though she is clearly feeling her estranged body as vitally intact.

<sup>903</sup> Oesterreich 1910, 222.

<sup>904</sup> Oesterreich 1910, 322, my translation: "[D]er Kern dieses Zustandes [besteht] in einer allgemeinen Hemmung des gesamten Gefühlslebens. ... [I]ch [betone] ausdrücklich, daß auch das zentrale Lebensgefühl eine solche Hemmung erfährt." Cf. also Oesterreich 1908.

always remain numerous acts of perception and thinking unharmed, which are similarly of subjective, egological nature. Depersonalization is not an absolutely egoless state.<sup>905</sup>

Oesterreich stresses that it is important to distinguish between ‘I’ and ‘personhood’, which are often identified problematically.<sup>906</sup> In depersonalization, according to him, the latter is affected and the former prevails.<sup>907</sup> As he puts it,

What is dissolved in depersonalization is the self-phenomenon alone, i.e. the normal feeling of life, but not the ‘I’, which merely finds itself in a very abnormal state.<sup>908</sup>

Oesterreich further points out that there is, of course, something paradoxical about it and we can find many contradictions in the first-person reports of depersonalized patients:

The loss of the normal personality-consciousness therefore ensues on the one hand the complaint: I don't feel myself any longer, I don't exist anymore. But the other side of the situation is: that the ‘I’ nevertheless still does exist, but it finds itself in a peculiar, very abnormal state, which is expressed in the first-person reports: I am someone else.<sup>909</sup>

This double aspect of selfhood explains the contradictions found in the experiences of individuals with depersonalization and their expressions. On the one hand, all feeling seems to be gone. On the other hand, the absence of all feeling is itself a feeling. This seeming paradox is sometimes expressed as the “feeling of non-feeling” (*Gefühl der Gefühllosigkeit*)<sup>910</sup>. The feeling of non-existence is still a feeling, and thus a mode of individual being. Ratcliffe observes the fact that a numb hand lacking feeling is not simply devoid of all feeling. Rather, physical sensation simply *feels* absent.<sup>911</sup> The feeling of non-

---

<sup>905</sup> Oesterreich 1910, 322f., my translation: “Wenn die Depersonalisierten aber oft auch sagen: sie besäßen kein Ich mehr, so ist das eine Täuschung. Selbst bei voller Athymie und Abulie bleiben immer noch zahlreiche Wahrnehmungs- und Denkakte erhalten, die ja ebenfalls subjektiver, ich-hafter Natur sind. Es handelt sich auch in der Depersonalisation nicht um einen absolut ichlosen Zustand.”

<sup>906</sup> Oesterreich 1910, 321.

<sup>907</sup> In fact, for Oesterreich (1910, 323), all experience is egological: “Egolessness only emerges when there is no psychic life left.” My translation: “Ichlosigkeit tritt erst ein, wenn überhaupt kein psychisches Leben mehr vorhanden ist.”

<sup>908</sup> Oesterreich 1910, 323, my translation: “Was aber in der Depersonalisation aufgehoben ist, ist allein das Selbstphänomen, das normale Lebensgefühl, nicht jedoch das Ich, das sich lediglich nunmehr in einem sehr abnormen Zustand befindet.”

<sup>909</sup> Oesterreich 1910, 324, my translation: “Der Verlust des normalen Persönlichkeitsbewußtseins hat sonach einerseits die Klage zur Folge: ich fühle mich nicht mehr, ich existiere nicht mehr. Die andere Seite der Sachlage: daß das Ich gleichwohl noch existiert, sich aber in einem besonderen, ganz abnormen Zustand befindet, ergibt in den Krankheitsberichten den Ausdruck: ich bin ein anderer.”

<sup>910</sup> Schneider 1920, 285.

<sup>911</sup> Ratcliffe 2008, 56.

existence cannot simply be equated with a lack of feeling and a non-existent 'I'. Instead, the feeling of non-existence is a state of 'I'.

Oesterreich's position and the claims involved can be summed up as follows:

- (1) There is a feeling dimension distinct from sensations;
- (2) One's own existence and personality is experienced in feeling, notably in form of a general feeling of life;
- (3) A weak or absent feeling translates into a weak or absent feeling of life, which manifests as feelings of non-existence;
- (4) Feelings of non-existence do not indicate the absence of 'I', but rather determine the state of 'I';
- (5) We should distinguish between 'I' and 'self/personality' as only the latter is absent in depersonalization.

In conclusion, Oesterreich sees depersonalization as a case wherein the subjective character of experience has changed significantly. Changes include the modification of a central feature: the feeling of life that usually accompanies all experiences. But such modification does not amount to a complete loss of subjective character.

#### 7.2.4 Self-experience and the feeling of being oneself in depersonalization

Oesterreich's view on depersonalization can be further assessed in terms of how it might inform my own proposal of self-feeling. First, I wish to emphasize that I concur with much of Oesterreich's deliberations. Still, there are some aspects that I would describe differently. For example, Oesterreich makes a Lippsian-inspired distinction between 'I' and 'personality/self'. I agree with his distinction or, to put it better, I agree with what the distinction ought to express. But I find the labelling somewhat confusing. Oesterreich wants to say that even in depersonalization, experiences such as perceptions and thoughts have subjective character. There are objects that are given *to me*—that is, they have 'for-me-ness' in Zahavi's sense. When undergoing an experience in a state of depersonalization, I am still aware that I am undergoing these experiences. Put differently, self-awareness of experience remains unharmed in depersonalization and this is what makes Oesterreich's 'I' to begin with.

Since the reflexivity of experience constitutes selfhood, which Oesterreich believes prevails in depersonalization, I think it would be clearer to admit that selfhood is not

dissolved in depersonalization. When Oesterreich claims the “self-phenomenon”<sup>912</sup> is absent in depersonalization, he invites confusion as reflexivity is sufficiently intact even in depersonalization. It is compatible with expressions such as ‘my experiences do not really seem to be mine’, ‘my experiences feel alien’, and ‘I feel absent in my experiences’. Instead of saying the ‘I’ prevails in depersonalization, we might better say (egological) pre-reflective self-awareness is involved *even* when depersonalization occurs. Even these experiences include a sense of self at least in terms of the minimal self, if not more.<sup>913</sup> When speaking about someone undergoing depersonalization, we are still dealing with a clear case of selfhood. Therefore, I think it is misleading to declare dissolution of the ‘self-phenomenon’ in depersonalization.

The second disagreement concerns the feeling aspect that is altered in depersonalization. Oesterreich identifies the ‘self-phenomenon’ with ‘personality-consciousness’ and ‘feeling of life’. For him, all notions refer to the same phenomenon that he takes to be absent in depersonalization. I have already pointed out how the ‘self-phenomenon’ cannot be omitted from the depersonalization experience, even though it might have undergone a significant change. But what about the other two? I agree that both can be seen as severely weakened or even absent in depersonalization. One may experience the world *as if* one weren’t a person, or as if one had no palpable personal identity stretching over time. And one may experience even the present moment as though one is not really there, *as if* one were not alive at all. Yet I am less inclined to accept the two (personality-consciousness and feeling of life) as identical. I can see how a lack of a normal feeling of life impinges on one’s self-experience as a person; if I feel dead and non-existent, surely, I will not feel a sense of myself as someone with certain characteristics who has a certain history and who will persist in the future. But I do not think the reverse is true in a similar way.

Does lack of a sense of personal identity necessarily amount to lack of the feeling of life? I can see how someone who has lost her sense of identity might have a feeling of perishing existence, a loss of identity that gives rise to feelings of death. Compared to a full-fledged personal identity, the mere momentous selfhood that only comprises the current situation to some may seem to be a poor, rare case of selfhood resembling a feeling of being nothing. But here, the necessary link is absent. Consider again the episodic life described by Galen Strawson:

---

<sup>912</sup> Oesterreich 1910, 323.

<sup>913</sup> I would even go farther and emphasize that individuals undergoing depersonalisation have a much more complex sense of self than the mere self-givenness of experience. They for sure know what it is like to be a person and crave for the feeling to be such. They are very much aware of their constitution as selves, as well as what they are lacking.

Episodics, looking out from the present, have very little sense that the I that is a mental presence now was there in the past and will be there in the further future.<sup>914</sup>

Diachronics may feel there is something chilling and empty in the Episodic life, but the principal thing about it is simply that it is more directed on the present. The past is not alive in memory, as Diachronics may find, but it is alive—Episodics might say more truly alive—in the form of the present: insofar as it has shaped the way one is in the present.<sup>915</sup>

Episodics do not feel a strong sense of identity. Instead, they take their being and existence to consist of the present, i.e., the present consciousness. Yet this does not necessarily mean they feel dead or not alive. At one point, Oesterreich himself highlights the presence of a feeling of life at times when we are not concerned about anything in particular, such as when we mindlessly gaze at nothing.<sup>916</sup> Simply equating consciousness of oneself as a person with the feeling of life seems a strong claim. We should also bear in mind that Lotze, who clearly figures in the background of Oesterreich's account, referred to self-feeling in order to account for rudimentary forms of self-experience that precede explicit or more cognitively-shaped forms of consciousness of self (e.g., as in non-human animals). It is not even clear whether persons suffering from depersonalization indeed lose all awareness of their personality and personhood. They may very well remember things of the past and recognize that they somehow have a history in life, even though they feel entirely detached from this history (considered as the experiencer they are at the time of the act of recollection). All this suggests that the identification of 'personality-consciousness' with 'feeling of life' is too simplistic. The idea that both are simply absent in depersonalization, as if this were a dichotomous variable that only allows for a 'yes' or 'no', is likewise reductive—and, as Billon's and Guillot's proposals, too crude.

In my own account for self-feeling, the notion of the *feeling of being oneself*, I make a slightly different proposal: Neither a sense of self nor the feeling of being oneself are completely absent in depersonalization. Afflicted individuals are still aware of their experiences. They are concerned about themselves and troubled by the fact that they feel differently, cut off from the world and from themselves. Accordingly, they retain a feeling of their being. In my view, then, first-person reports of 'feelings of non-existence' are expressions of a certain feeling of being oneself.<sup>917</sup> The *normal* feeling of being oneself is modified in depersonalization. It can be seen as absent, but the same cannot be said of all feeling of being oneself. What, phenomenologically speaking, makes the normal feeling of

---

<sup>914</sup> Strawson 2017, 51.

<sup>915</sup> Strawson 2017, 52.

<sup>916</sup> Cf. Oesterreich 1910, 322.

<sup>917</sup> Again, and to stress, it would be wrong to simply classify and identify verbal expressions of *feelings of non-existence* with an alleged *non-existence feelings*.

being oneself absent? How can we determine what features of the feeling of being oneself (as I have conceptualized them above) are modified in depersonalization?

In Section 7.1, I argued that the feeling of being oneself refers to one's own experiential existence—not merely to the vitality of one's body and its organic functional status. In the self-feeling of depersonalization, one's own experiential existence is one of alienation. One feels cut-off from what should be one's own experiential process. It is not that one experiences a lack of an experiential process. But while undergoing and being in the process, one does not really participate. It is an extreme kind of impotence that condemns oneself to being a mere observer of one's passing experiences without a satisfying feeling of involvement. This lack of egological participation and involvement, I suggest, is due to an alteration of feeling owed to a modification of the horizon of possibility (the possibility of experiences of pleasure and pain). The following first-person expressions are reported by Bockner, who describes depersonalization as the “feeling that one is no longer oneself”<sup>918</sup>. He examines the case of a woman with depersonalization who emphasized:

I really love music, but now I can't bear it because it doesn't stir me. I hear it with my ears in a rather distant way, but there's no pleasure attached.

I should feel deep pleasure from these roses my husband has sent me, but they mean nothing to me. I was an expert at arranging flowers, but not now.

I'll eat anything put before me. We had walnut cake which I normally adore—but it might have been a piece of dry bread. I eat, not for pleasure, but only to live.

I was always soft and feminine, but now I feel hard, as though nothing could shock me any longer. I'm beyond feeling. Meeting a friend I say the conventional things, but feel nothing and don't mean what I say.

I can't admire a painting—it affords me no pleasure.<sup>919</sup>

As these descriptions underscore, depersonalization can be described as “a chronic state of the lack of conscious pleasure”<sup>920</sup>. Note that this condition is more than the mere absence of pleasure, which is why the aspect of chronicity is important to bear in mind. Depersonalized patients, I suggest, are acutely aware of chronicity in that they experience the impossibility of feeling pleasure. Their condition is whatever experiences seem to be possible for them, and pleasure is not among them. According to my account of self-feeling,

---

<sup>918</sup> Bockner 1949, 969.

<sup>919</sup> Bockner 1949, 969.

<sup>920</sup> Ackner 1954, 858.



that is where a normal feeling of being oneself stands and falls. I earlier argued that with Locke and Lotze, it is the capability for pleasure and pain—the non-neutral experience of the possibility of pleasure and pain—that constitutes and so shapes the feeling of being oneself. Once again, we must take care to describe depersonalization accurately. I am suggesting that they experience the impossibility for their own selves to experience pleasure and pain. However, that does not mean they are blind to the possibility of an experience of pleasure and pain. They are well aware that they have experienced pleasure before, or that pleasure is an experience possible for others. Hence, they do experience a general possibility of pleasure and pain, but they do not feel that it is possible for them.

This particular phenomenological profile of the possibility of pleasure and pain gives rise to the experience of depersonalization and the seemingly paradoxical structure of the ‘feeling of non-feeling’ and ‘being as if non-existent’ involved.<sup>921</sup> On the one hand, the depersonalized do not experience pleasure or pain as possible for them. This makes them lack the normal feeling of being oneself in that they are not involved in their own experiences. Deprived of pleasure and pain in the normal sense, patients no longer feel that their experiences concern them. If selfhood consists of partaking in experience, then corresponding self-feeling is dampened down. Yet on the other hand, they remain aware that pleasure (and pain) are generally possible experiences. Such an awareness is not neutral, and essentially given in feeling. Persons so-afflicted still value the experiential existence of selfhood that, in depersonalization, is experientially manifest in a felt self-concern for one’s own possibility of pleasure and pain.

Patients with depersonalization are worried and concerned not about specific, concrete possibilities of pleasure and fulfilment, but about the very core of their egological being: participating in the process of experiential givenness. Although that means they might lack the experience of the individual possibility of pleasure, their mode of being is still one that essentially involves self-concern. In fact, the paradox intrinsic to the condition is often fueled by the absence of the possibility of pleasure and of felt participation in life. This is precisely that what elicits psychological pain of the most intensive kind. The experience of the absence of an individual possibility for pleasure and pain is not something we experience in a neutral way, even though persons suffering from depersonalization may

---

<sup>921</sup> In their study, based on empirical findings, Mula et al. (2010) have argued that we need to distinguish between affective depersonalization and anhedonia (for instance, in major depression). In line with this, I would like to suggest that we might be able to differentiate phenomenologically between different forms of the loss of the capability for pleasure. Differences might be quantitative, for instance, between a dampening down of pleasure and an almost total absence of pleasure. However, I would also propose that what distinguishes these forms is the way how the *possibility* of pleasure is experienced. It makes a difference whether I experience pleasure as impossible ‘because of *x*’ or ‘as long as *x* prevails’ vs. ‘independently of any possible circumstance’. Here, these more subtle but potentially impactful differences are not relevant. What matters is that depersonalization presents an extreme case in that the possibility of pleasure is only one of general or theoretical kind but isn’t experienced as one that would apply to the individual case of oneself.

complain about the fact that everything seems neutral to them. Experiencing everything as neutral is *not* a neutral experience. It is fundamentally an experience of loss—the loss of a meaningful life in which one normally would participate in terms of feeling, which in turn may engender the feeling that one does not exist at all. In short, I think it is a mistake to construe depersonalization in terms of a lack of either the subjective character of experience or self-feeling. Rather, depersonalization involves a significant alteration of the feeling of being oneself. Alteration consists of a change that lies beyond the normal variability of self-feeling. But even if this change presents as an extreme case, it is still one that is located *within* the dimension of self-feeling.

### 7.2.5 Zahavi's account of depersonalization and the feeling of being oneself

The preceding sections defended the view that depersonalization does not undermine the minimal self-account. Similarly, it should not be a surprise that the proposal I have developed here is also compatible with Zahavi's account of depersonalization. Still, comparison between his account and my understanding of depersonalization is informative. In order to describe the phenomenology of depersonalization in a way that sits well with the claim of a minimal self-view, Zahavi must refer to *some* alteration in the subjective character of experience. He must also do so without maintaining that experience in depersonalization lacks all subjective character, which would be contrary to his minimal self-view. I suggest it is the feeling of being oneself that is altered in depersonalization. Even though Zahavi does not refer to a feeling of being oneself or self-feeling *expressis verbis*, the way he construes his account for depersonalization reveals implicit acknowledgement of the feeling dimension and its relevance to self-experience in some way. At the very least, the feeling dimension is relevant to an experience of life or being (which I would argue is best accounted for in terms of feeling). This acknowledgement further corroborates my proposal.

The first thing to notice is that Zahavi's strategy for explaining experiential change in depersonalization resembles Lotze's argument in favor of feeling as an inextricable aspect of self-experience. Zahavi addresses depersonalization as a possible challenge to the minimal self-view, emphasizing that "the feeling of depersonalization might ultimately be due to an exacerbation of self-awareness, a kind of ultra-reflection, rather than due to a lack or loss of self-awareness"<sup>922</sup>. The idea is that a heightened self-objectification through excessive reflection may give rise to an increasing self-alienation and even complete the detachment of oneself from one's own experiences. Such a condition, Textor emphasizes,

---

<sup>922</sup> Zahavi 1999, 155. Cf. Sass 1994, Schilder 1951.

is (taken *cum grano salis*) comparable to Lotze's angel who has perfect reflective knowledge of himself but lacks any self-feeling:

[H]e has no affect at all. He is still able to self-ascribe experiences, yet he does not *own* the experiences he self-ascribes; they do not seem to him to be his. They are for him objects like any other object of which he has knowledge. While Michael [how Textor calls Lotze's perfect angel, P.S.] has perfect knowledge of himself, he and what belongs to him have no special and incomparable value for him. He does not care about the mental properties of which he has non-observational knowledge any more than about any other properties he knows about. His knowledge of himself is, therefore, only an *indifferent* representation, and he lacks self-consciousness.<sup>923</sup>

Lotze's point is if the self-consciousness of a conscious being only exists in the reflexivity of self-awareness without any self-feeling, it will not experience its own being and existence. It will lack *Fürsichsein*. To such a conscious being, "the inwardness and innerness would remain unknown and incomprehensible"<sup>924</sup>. Lotze is claiming that without self-feeling, there is a lack of *real* self-experience. Oneself would appear to oneself in an indifferent (and therefore object-like) manner, a mode in which one *is* not for oneself. This is an experience wherein one's own existence does not even come to mind ('for-me-ness' without *Fürsichsein*). Roughly, then, the equation Lotze proposes is the following:

Lack of self-feeling = reflective self-objectification = lack of *Fürsichsein*

Zahavi's argument concerning depersonalization refers to a very similar explanatory mechanism:

Ultra-reflection = reflective self-objectification = feeling of depersonalization

In both cases, the view is that even an extensive or exclusively reflective self-awareness cannot bring about a sense of one's life and being. If our self-experience approximates a state of reflecting on one's own experiences, they themselves lose the character that *lived* experiences usually have.

Whereas Lotze's point is that reflective self-awareness without feeling does not give rise to consciousness of one's own existence, Zahavi highlights that focusing on a reflective attitude towards one's own experiences and entering a state of ultra-reflection may actually undermine a person's participation in her own experiential process: "The perceptions, feelings, and thoughts of its social self are compulsively monitored with such a critical

---

<sup>923</sup> Textor 2018, 165.

<sup>924</sup> Lotze 1884/1856, 280, my translation: "unbekannt und unverständlich würde ihr die Innigkeit bleiben".

detachment that they lose spontaneity and become lifeless and unreal...<sup>925</sup>. It should be noted that Zahavi is not saying reflective self-awareness lacks all feeling. Instead, he points out that “[e]ven normal reflection might be accompanied by a feeling of detachment and dissociation”<sup>926</sup>. For him “depersonalization can be viewed as an extreme form of self-objectification that occurs when self-observation turns into self-reification”<sup>927</sup>. This may lead us to feel “that there is something wrong with ourselves, and finally feel that we are no longer ourselves”<sup>928, 929</sup>.

Why is all this relevant? The understanding of depersonalization offered by Zahavi implies recognition of the feeling dimension of self-experience. He notes that through excessive self-monitoring, one’s lived experience becomes ‘lifeless’ and ‘unreal’. He thus reveals that he takes experience as normally unfolding in a way that makes it appear to be full of life. Under non-pathological circumstances, experience is usually lived through in a way that makes it feel real. But with excessive reflection, our own experiences can develop an alien character in that we feel oddly detached from them.

How can Zahavi describe the transition from a lived experience unfolding in a normal, non-pathological way to the pathological self-experience of depersonalization? He does so by implicitly referring to a change in feeling. Feeling makes experience seem full of life and real. Feeling can also make it seem detached and unreal. What matters here is not the fact that the change in feeling might be (partially) owed to the transition from a pre-reflective or reflective towards a deviant, ultra-reflective self-experience. The point is that this transition covaries with a change in feeling. Patients with depersonalization do not complain about the fact that they have a hyper-reflective attitude. They complain about the self-alienating feeling associated with such an attitude.<sup>930</sup> Zahavi cannot help but recognize this particular change—and I suggest that he indeed does so. But if so, it would only be

---

<sup>925</sup> Zahavi 1999, 156.

<sup>926</sup> Zahavi 1999, 156.

<sup>927</sup> Zahavi 1999, 156.

<sup>928</sup> Zahavi 1999, 155.

<sup>929</sup> See also Cianuica, Charlton and Farmer (2020), who have further developed the idea of self-objectification in depersonalization.

<sup>930</sup> It might also be argued that ultra-reflection in depersonalization is rather the attempt to compensate for and thus a by-product of an altered self-feeling. In this vein, one could interpret the altered self-feeling to be primary, motivating a secondary, heightened reflectivity that ought to repair the loss of self-familiarity. Failing in developing a clearer self-feeling, a process of continuously increasing reflectivity may ensue and engender the kind of self-estrangement typical for depersonalization. In fact, quite different connections between ultra-reflection and an altered self-feeling are intelligible. For a perspective on the different ways experiences can hinge together, see Sass 2014 or Schmidt 2018. Here, however, all what matters is what the phenomenon in question, depersonalization, phenomenologically consists of. Crucially, whatever is primary in depersonalization from an aetiological perspective, the suffering of persons so-afflicted typically concerns the shift in self-feeling. *That* is the phenomenological datum which must be acknowledged regardless of what specific explanation for the phenomenon one favors.

consistent on his end to acknowledge that self-experience involves feeling even in the normal, non-pathological case.

### 7.3 Conclusion: Depersonalization as a modification of self-feeling

This chapter had two aims. In the first part, I presented my account for the feeling of being oneself in light of alternative notions of self-feeling. In doing so, I noted that self-feeling refers to the feeling dimension intrinsic to self-experience. I clarified that I do not consider self-feeling to be a certain content of consciousness. Instead, it is an aspect of how the experience of any content unfolds. That feeling thus reveals our own individual existence as experiencing beings. Accordingly, I distinguished my notion from other accounts that see self-feeling as primarily and originally awareness of one's own bodily and organismic functions. Moreover, I argued that feeling varies depending on how we find ourselves in the world. Following Ratcliffe's notion of existential feelings, I posited that the sphere of experienced possibilities essentially shapes our existence—and therefore how we feel ourselves to be. Our existence in experience, i.e. our existence as an experiencer, is intrinsically experienced as valuable. This is something that always manifests in a felt self-concern. Experience of the possibility of pleasure and pain originally constitutes the horizon of felt self-concern.

In the second part, I aimed to shed light on the psychopathological phenomenon of depersonalization. In the past, depersonalization has often served to criticize or motivate conceptions of selfhood and self-experience. More specifically, I discussed two lines of reasoning brought forward to understand the condition of depersonalization. If their rationale were convincing, both lines would significantly undermine my proposal for self-feeling. In the first view, experience in depersonalization lacks subjective character. So, it is bereft of self-experience. In the second, less radical view, not all subjective character is missing in depersonalization. Only self-feeling, the individual feeling of life characteristic of normal self-experience, is absent.

On the contrary, I argued that conceptualizing depersonalization as a form of anonymous or selfless experience (or as a form of feeling-free experience) is unconvincing. I defend the view that those afflicted by depersonalization show great self-concern about how they live through their experiences. In doing so, they demonstrate the retention of a(n altered) sense of self. They also report changes in self-perception that are best described as a modification rather than a simple loss of self-feeling. Even though they experience pleasure and pain as impossible for themselves either as individuals or as singular experiences, this general impossibility provides the experiential ground for the emotional pain that broadly characterizes their condition. Far from being a state of indifference,

depersonalization comes with strong feelings about the indifference that afflicted persons feel in terms of their singular experiences. Put differently, the depersonalized are not indifferent about their felt indifference. They care about their experiential being. They are concerned for themselves. Depersonalization is actually an extreme case of experiencing—one that is located within the dimension of feeling, given that it consists of a certain feeling of being oneself.

Having rejected the two lines of reasoning, I then offered a reformulation of depersonalization that involves a phenomenological description by applying my own proposal of self-feeling to the case. Doing so provides neither a final nor conclusive argument in favor of my understanding of self-feeling. I propose that my discussion of depersonalization simply corroborates my case for the feeling dimension of self-experience. Far from posing a serious challenge to the view that feeling is an intrinsic aspect of self-experience, depersonalization turns out to be a phenomenon that requires some reference to self-feeling and the significant changes it may incur. Only then can it be understood properly.

## Conclusion

### Self-Realization and How I Feel about Myself

What is self-experience and how does it relate to selfhood? Addressing this question, which I suggested should be at the center of the debate about the self, I discussed several possible responses and defended self-experientialism. Self-experientialism, in the way I described it by building on the phenomenological tradition and its current proponents, presents the view that selfhood consists in self-experience. On this view, it is by experiencing ourselves that we are selves. The essential feature of the being of selfhood is pre-reflective self-awareness that is—according to the minimal self-view—intrinsic to experience. I am self-aware insofar as something is *given* to me and I am myself insofar as something is given *to me*. It is this form of pre-reflective reflexivity in which selfhood realizes itself in the most original way. In this particular sense, self-experience is the *modus essendi* and *modus cognoscendi* of selfhood. For, on the view of self-experientialism, selfhood is real by realizing itself in self-experience, i.e., by becoming aware of itself as the one undergoing a given experience. Moreover, self-experience being the *modus cognoscendi* also means that by reflecting on self-experience, we can determine what selfhood is. After introducing the idea of self-experientialism, I addressed the question whether such a formal concept of selfhood might not be too thin. In this context, I defended the minimal self as a formal notion but argued that its phenomenology also involves a sense of one's own existence, a sense of being real. This kind of consciousness, I suggested, is part of the minimal self in terms of a feeling of being oneself. More specifically, I offered the view that feeling is integral to the temporal self-manifestation of the stream of consciousness, which is understood as the phenomenological structure of the minimal self.

How did I yield this position? What is the rationale of the overall argument? The general line of reasoning was to first motivate an experiential approach to selfhood. I did this by providing an argumentation *ex negativo*. To this aim, I discussed what I take to be the three main forms of skepticism about self-experience that all, each in a different way, undermine an experiential approach to selfhood. I found that none of the skeptical views, which stand in contrast to self-experientialism, offers a convincing account of selfhood. If my rejection of these skeptical views is right, then I suggest we are better off pursuing an experiential approach to selfhood, one that defines selfhood in terms of experience. I identified Zahavi's minimal self as such an account but emphasized that the ontological implications and commitments of his notion of self were still not fully clear. I tried to carve them out by specifying that Zahavi's view of the experiential self consists in self-experientialism. To launch the minimal self-view in the way Zahavi does, I argued, one

must acknowledge self-experientialism. For, only if the self-reflexivity of pre-reflective self-awareness is taken as the most basic form of self-experience Zahavi can call it a minimal self. Or, to put it differently, if Zahavi is saying that pre-reflective self-awareness amounts to a minimal self then this means that for selfhood to be prevalent in the first place, as a minimum criterium, at least the condition of pre-reflective self-awareness must be met. This also implies that selfhood is essentially experiential, given that pre-reflective self-awareness only occurs in the context of something being experientially given.

Although the claim that self-experientialism is implied in the minimal self-view is sparsely noticed, let alone discussed, in the broad and highly evolved debate about the minimal self, I take it to be rather uncontroversial. Some may see in this the reason why precisely there is no need in discussing it. That, however, I suggest would be a fundamental mistake. Guillot, I argued, is right in that this is not a trivial thesis and that it needs further argumentation.

What are the arguments that I offered in order to buttress self-experientialism? Crucially, I wish to emphasize that I don't believe there is one master or knock-out argument in favor of any of the presented views of the self. After all, the issue of selfhood belongs to the thorniest subjects in philosophy, creating an eternal debate about how we should understand ourselves and what we are. Even so, the different discussions and arguments I presented, I suggest, in conjunction with each other, speak in favor of self-experientialism.

*1) Ex negativo: Skeptical accounts about self-experience are unattractive.*

First, I would emphasize that self-experientialism is partly motivated by the fact that competing theories tend to commit—explicitly or implicitly—to one of the three forms of skepticism about self-experience, all of which I denoted as unattractive options for a theory of selfhood. Rejecting the skeptical views that were meant or are implicitly adverted to Cartesianism, thus, pushes us into the direction of phenomenological self-experientialism. The minimal self and the implied phenomenological self-experientialism provides an alternative to both Cartesianism and the skeptical reactions to it. That was the overall and general strategy to argue in favor of self-experientialism. To show that self-experientialism is the best alternative to Cartesianism and skepticism, I discussed other recent experiential approaches to selfhood and found that they are either not convincing or failing because they amount implicitly to the third form of skepticism about self-experience.

*2) Self is prior to personal identity: A more robust Lockean account.*

In this context, I also addressed a typical line of argument against experiential approaches to selfhood, namely the need to account for personal identity and the claim that a Lockean—or experiential—approach to personal identity is ill-suited to solve the issue of



personal identity. By providing a modest interpretation of Locke with regard to personal identity, I argued that this issue is secondary to the question of selfhood and that the latter constrains the former—rather than the other way around. First, we need to define what selfhood is, only then can we address the question under which circumstances we might speak of personal identity. Although this is not necessarily and unambiguously an argument *in favor* of the kind of self-experientialism I suggested, it is very much compatible with self-experientialism and so helps taking out one thorn out of the flesh of self-experientialism. Importantly, it helps undermining a central motivation for accounts that make an attempt to define selfhood in non-experiential terms.

3) *Genealogy I: Semantics of self as the bearer is derivative of experiential selfhood.*

A more positive and direct argument given in favor of self-experientialism was of a genealogical kind. This argument starts from the assumption that we—in everyday life—take selfhood to consist in—some way to be specified—being the bearer of experiences, which McClelland considers a platitude about selfhood. To be a self means to undergo experiences, experiences belong to the self. In this context, I distinguished between a substantial and an experiential understanding of being the bearer of experience. I argued that any talk about being the bearer is motivated by and traces back to the experiential understanding of the bearer: the experience of the fact that different experiences are given *to me* as I find myself in experience. This experiential understanding is primary to the substantial understanding of the self as an alleged extra-experiential bearer of experience. But not only is the substantial understanding of self derivated of the experiential understanding, a substantial self cannot possibly play the role of the self as which we experience us *in* experience—and will never meet the criteria of its job description. An extra-experiential or non-experiential self understood as that to which something is experientially given makes no sense. The ‘to me’ of givenness refers to myself as the experiencer as I find myself *in* experience.

4) *Self-experience doesn't purport to refer to something extra-experiential.*

The last point can also be put in a slightly different way. It has sometimes been argued—as was the case with Metzinger—that self-experience and phenomenal selfhood as a phenomenon purports to refer to what is classically called the soul, a persisting, incorruptible substance that transcends experience while really self-experience so-understood doesn't refer to anything at all. I argued that the idea that self-experience must have an extra-experiential referent is inappropriate. But not only that. I also emphasized that, by drawing from a discussion of McGilvary's critique of William James as well as Zahavi's Husserlian notion of the temporal self-givenness of the stream of consciousness, the most basic phenomenon of self-experience doesn't refer to anything else than itself, i.e.,

the stream of consciousness. Self-awareness of the stream of consciousness, as a phenomenon, doesn't purport to refer to something external to experience. That view was further corroborated by an analysis of the Kantian notion of the thing in itself. While the naïve realism of our everyday experience might, to some degree, motivate such a concept of the perceptual object being there independently of whether we experience it or not, this doesn't apply to self-experience as a phenomenon at all. We don't experience ourselves in such a way that suggests that we, as we find ourselves *in* experience, exist whether we experience ourselves or not. Although we might feel that we can 'survive' periods of unconsciousness, in no way do we experience ourselves in a way that suggests we would proceed to be there if all experience was irrevocably taken from us. The phenomenon of self-experience, I suggested, consists in experiencing oneself as being essentially *in* experience.

*5) Argument from the phenomenon of self-care and the feeling of being oneself.*

Another line of argument pointing in the same direction referred to self-care and what I introduced in the later parts of the dissertation as the feeling of being oneself. The upshot of this argument was this: We never primarily care about anything non-experiential that might be involved as an immaterial (whatever that might be) or neurophysiological underpinning of our experience and self-experience when we care about *ourselves*. The primary concern for ourselves is directed at our experiential being, our being in experience as experiencing. It is our experiencing that originally matters to us in self-concern.

*6) Genealogy II: Self-experientialism legitimizes ex post everyday intuitions about selfhood.*

Furthermore, I offered an argument from a pragmatic and phenomenologically adapted reflection of a Humean kind. Hume's strategy in reflecting on his experience was to show that the soul, as classically conceived, does not and could not possibly show up in experience. Sometimes this has been taken to be a proof of the fact that experience is bereft of self-experience. However, I not only argued that the No-Self-Picture is phenomenologically inaccurate. I also suggested that one could adapt Hume's inquiry into the phenomenology of experience. Instead of presupposing a certain notion of selfhood and test whether it is warranted by the phenomenology of experience, one could start with a phenomenological reflection on experience and ask whether there is anything in experience that could motivate and explain why we talk about self-experience and selfhood in the first place. So, the idea would be to examine experience in search for the best candidate of self-experience. If pressed for an answer, I submit, most would agree that the experience of reflection on one's own experience would count as a very good, if not the best, candidate. In reflecting on my experience *p*, I do have a sense of myself as having undergone *p*. But,

as a next step, the question would be whether reflective self-experience is the most basic or original form of self-experience. And this is where a further phenomenological analysis of reflective consciousness, as Zahavi has provided it, would come into play. As Zahavi, convincingly in my view, has shown, reflective consciousness is only possible because experience involves pre-reflective self-awareness. If that is correct, the argument goes, we can determine pre-reflective self-awareness as the most basic form of self-experience. The next step then is to ask: What does this kind of self-experience tell us about selfhood? Since it is described as the temporal self-givenness of the stream of consciousness, it simply refers to the experiential process itself and not to anything extra-experiential. Put differently, what is most likely the best candidate for self-experience refers us to the experiential process, the argument goes. Accordingly, we have good reason to conceptualize selfhood in experiential terms. In sum, the general line of this argument then is this: We begin by pragmatically acknowledging our existence and taking talk about selfhood seriously. We therefore ask where this talk and intuition stems from and find the root of it in reflective experience and the reflexivity involved. We then further inquire phenomenologically this form of reflexivity and then come to yield an insight into a more original form of pre-reflective reflexivity, which we identify as the most basic experience of reflexivity. Next, we interpret this kind of experience of reflexivity as the most original self-experience and root of any talk about selfhood. We thus investigate the phenomenon of this pre-reflective reflexivity and come to see that it refers to the experiential process and not to anything extra-experiential.

In the third and last part of the dissertation, I addressed the line of criticism against the minimal self-view according to which the notion of the minimal self is too thin and doesn't provide a satisfactory account of selfhood. Since I made the case for self-experientialism by building on Zahavi's minimal self, evidently, such a line of criticism concerns self-experientialism alike. Indeed, if the mere self-awareness of experience is taken to be too formal and thin in order to present the essential feature of selfhood, then, equally, the main idea of self-experientialism is put into question. One shared general strategy of critics of the thinness of the minimal self-concept is to point out that, because the minimal self is structurally the same in every experiencer, the notion is too weak in that it does not suffice to individuate the different experiencers. That, however, is a requirement for any notion of self. I argued that Zahavi's account is well-prepared to respond to this line of critique. However, I also offered a further interpretation of the minimal self, one that integrates feeling into the approach. That, I suggested, might not only have the advantage to further accommodate the critics' worries that the minimal self is too weak a notion. I also emphasized that for the minimal self to present the most basic form of being of selfhood, an experience of being, a consciousness of one's own existence should be integral to the

minimal self in some way. If self-awareness is what makes *me* most originally *someone*, an *experiencing being* in the first place, then somehow, even on the most basic level of selfhood, the existence of oneself must feature in the experience, i.e., as an aspect of experience. As Zahavi's position stands, it is not fully clear how a consciousness of existence is involved in the minimal self. In reference to a wide-spread tradition of associating consciousness of being with feeling, I explored and suggested that feeling can be integrated into the minimal self-approach. The view I proposed then was that it is through feeling, as an aspect of the minimal self, that even the most basic form of selfhood involves an experience of one's own existence. To do this, I offered an interpretation of Lotze's *Fürsichsein* and Ratcliffe's notion of *existential feeling* in the context of Husserl's account of *inner time-consciousness*. While appeal to feeling might not be necessary in order to describe the minimal self and its structure as the most basic sense of self or self-consciousness, I suggested, feeling—in terms of a feeling of being oneself—can help to make plausible the claim that the minimal self presents the most basic form of selfhood and its being. Finally, I discussed whether the psychopathological phenomenon of depersonalization presents a challenge for my interpretation of the minimal self. As a response to such an anticipated objection, I highlighted that, if anything, the phenomenon of depersonalization corroborates my understanding of the minimal self and the feeling of being oneself.

Overall, I defended three major genuine claims that each correspond—more or less—to one of the different parts of the dissertation:

- (1) the claim that self-experience is informative with regard to the existence and being of selfhood (*epistemological claim*);
- (2) the claim that self-experience presents the existence and being of selfhood (*metaphysical claim*);
- (3) the claim that self-experience incorporates a consciousness of one's own existence in terms of a feeling of being oneself (*phenomenological claim*).

I shall note that I am well aware that all three claims are highly controversial and that it is likely that many will not subscribe to all three of them. However, I wish to note that these views, although interrelated and mutually supporting, can be—in some cases at least—held independently from each other. For instance, one could acknowledge the epistemological claim but remain cautious with regard to the metaphysical claim. One might not be persuaded by the argumentation for self-experientialism and consider a different alternative to the skeptical views about self-experience. It is also possible to follow the idea of self-experientialism but reject the phenomenological claim that self-experience involves a

feeling of being oneself. Moreover, one could also subscribe to the idea that self-experience incorporates a feeling of being oneself, but construe of the latter in a different way so as to make it compatible with a skeptical view about self-experience. That is, one could subscribe to the phenomenological claim but remain skeptical about self-experience and self-experientialism. However, it doesn't strike me as possible to accept the metaphysical claim, i.e., self-experientialism, *and* deny the epistemological claim. If one defends that self-experience presents the being of selfhood, then one must also defend that self-experience is informative with regard to the being of selfhood.

What follows from this relative independency? Most importantly, it means that taking issue with one part of the argumentation or claim doesn't yet suffice to make other arguments and claims that the dissertation thesis offers fall. That said, perhaps someone might find in fact all three of the claims at fault. Yet, I want to suggest, even in that case, there is a number of insights that this work offers that might be useful for further debate of different *topoi* relevant to the philosophy of self. Let me provide the major and most important examples. This will also allow me to shed light on how the investigation of this dissertation thesis could be further extended in the future in order to respond to remaining open questions or to provide new perspectives.

### *1) Conceptual framework and synopsis of possible positions on self-experience*

As I have emphasized, there are three main skeptical views about self-experience that can be and have been formulated as attacks on one or more than one of the Cartesian triad of views:

- (1) that we have self-experience
- (2) that we can infer from self-experience that we exist
- (3) that we can infer from self-experience what our metaphysical nature is: thinking being.

That is, the three skeptical views stand in a systematic relationship to Cartesianism about self-experience and self. As far as I know, the spectrum of these positions and their overall and general relationship have not been carved out so far in a systematic manner. Rather, in the history of philosophy, typically one of the ways has dominated a certain phase or approach. Yet, independently of the history of philosophy and how it has unfolded, the three skeptical views and self-experientialism broadly construed<sup>931</sup> present the main if not the whole spectrum of possible positions on selfhood. That claim might surprise some. For, some accounts of selfhood might not make any explicit appeal to self-experience at all. But, in response, I would contend that any view on the self also must somehow map onto

---

<sup>931</sup> With the term 'self-experientialism broadly construed', I include Cartesianism and Lockeanism, and other conceptions that are similar but not identical with the kind of phenomenological self-experientialism that I defended on the basis of Zahavi's concept of the minimal self.

the debate about self-experience, i.e., will have an *implication* for the role attributed to self-experience. Any account of selfhood will imply either the view that self-experience is informative with regard to selfhood—or not. It is not possible to offer a certain conception of selfhood and then remain neutral on the question whether self-experience reveals the nature of selfhood. Neither is it possible to remain neutral about *both* self-experience and the notion of selfhood—and to call this a conception of selfhood. As a conception of selfhood, it is possible though to claim that we have self-experience while remaining neutral about the nature of selfhood. That, however, corresponds to the third skeptical view. Accordingly, whatever position one might endorse with regard to the self, it will either amount to self-experientialism of some form or to a skeptical view about self-experience. Depending on to which class of account a proposal belongs, to be persuasive, it must find a response to the corresponding challenges, as I have carved them out for each line of thinking. Hence, even in case of a rejection of the three claims I defend, my analysis might prove helpful for the assessment of any future conception of selfhood as well as of the infinite number of existing conceptions I could not address in the dissertation. It is in this sense that I present a novel conceptual framework that might help to navigate the wild and cloudy sea of the philosophy of self. Given that the debate on the self will unlikely ever come to an end, it is important to know its points of gravitation. My proposal is to consider self-experientialism broadly construed and the three skeptical views as such poles.

<b>1<sup>st</sup> skepticism</b>	<b>2<sup>nd</sup> skepticism</b>	<b>3<sup>rd</sup> skepticism</b>	<b>self-experientialism</b>
<ul style="list-style-type: none"> <li>- accepts self-experience</li> <li>- denies reality of selfhood (existence)</li> <li>- denies that self-experience reveals being of selfhood</li> </ul>	<ul style="list-style-type: none"> <li>- denies self-experience</li> </ul>	<ul style="list-style-type: none"> <li>- accepts self-experience</li> <li>- accepts reality of selfhood (existence)</li> <li>- denies that self-experience reveals being of selfhood</li> </ul>	<ul style="list-style-type: none"> <li>- accepts self-experience</li> <li>- accepts reality of selfhood (existence)</li> <li>- accepts that self-experience reveals being of selfhood</li> </ul>
<p>Concludes selfhood is a hallucination.</p>	<p>Concludes selfhood is simply unreal or we have no reason from our experience to speak about selfhood.</p>	<p>Concludes that our experience of ourselves as existing is correct but holds we cannot infer from that what selfhood really is.</p>	<p>Concludes that our self-experience reveals existence and being of selfhood. Phenomenological self-experientialism interprets self-experience as the being of selfhood.</p>

## 2) *New perspective for the research in the history of philosophy*

The conceptual framework distinguishing between forms of self-experientialism and forms of skepticism about self-experience might also allow for a new perspective on the history of philosophy, especially since Descartes. Identifying the different proposals to be found in the history of philosophy with one of the lines described by the framework may help to generate and discuss hypotheses about the systematic relationship with other approaches. Moreover, the conceptual distinctions of the framework can also help to compare different interpretations of a thinker's proposal.

### 3) *Self-experience constrains the question of being*

A major part of the dissertation has consisted in working out where the three skeptical views about self-experience go wrong. As it turned out, even though the views are fairly different, they share—at least the proxies I discussed for each—what I argued is a misplaced assumption, building on the phenomenological tradition: the assumption that the self ought to be taken as a special kind of object. In the case of the first skeptical view, the Hallucination model, the idea was that if selfhood ought to be called real, there would have to be a certain physically measurable thing to which self-experience refers. There is no such object, *therefore* selfhood ought to be hallucination. The second skeptical view, in terms of the No-Self-Picture based on a certain interpretation of Hume, argues that if we really had self-experience, then this would mean that the self would show up just like perceptual objects do, namely in form of sensory impressions. But there are no sensory impressions that correspond to an alleged self-object and *therefore* we cannot even be said to have self-experience. The third skeptical, notably in terms of the Kantian elusiveness view, treats the self like an object. For Kant, the self-experience that he is willing to grant underlies the same conditions as object-experience. Just like in the case of perceptual objects of the physical world, we can only know of and about ourselves insofar as we appear to ourselves. In both cases, the thing in itself eludes our experiential perspective and *therefore* we cannot know the nature of selfhood. Although all these views can be considered as responses to the Cartesian understanding of the self as a *res cogitans*, a thinking *thing* and *substance*, it is not at all clear why for selfhood to be real it must amount to an object of a certain kind. In fact, emphasizing and defending main tenets of the phenomenological tradition, I have suggested that self-experience, in its most basic form, doesn't consist in a form of objective self-experience. Selfhood is real in self-experience, even though not necessarily as an object of experience. One important repercussion of this is the notion that 'being real' and 'being an object' do not coincide. However, the question arises as to what exactly do we mean by 'being'? That is an important question for which further research and analysis would be required. Post-Husserlian phenomenologists such as Martin Heidegger, Jean-Paul Sartre, Maurice Merleau-Ponty, or Michel Henry have addressed related questions while typically trying to avoid Cartesian reifications of selfhood. To deal with the notion of being, however, is an issue that presents a project of its own, and which, accordingly, couldn't be accomplished here. Yet, I wish to emphasize that, while this may read as a limitation of the positions that I defend here, I rather suggest that it is a strength not to presuppose a certain understanding of being, as the three skeptical views, as well as Descartes, have done. All these proposals, in some way, include the view that for something to be real it must be an object. By contrast, that assumption is undermined by phenomenological self-experientialism. In future research, it could be further investigated how phenomenological self-experientialism may inform more general ontological debates and how it affects our

understanding of being more generally. My suggestion is that the notion of selfhood should constrain the notion of being, rather than the other way around.

#### 4) *The question of priority: Self vs. personal identity*

Defending self-experientialism, I argued that the question of what selfhood is should be settled before we start asking questions about personal identity, which, of course, some might criticize. But I think that the very question of which of the issues—self vs. personal identity—has priority over the other has not yet sufficiently been addressed in the general debate. This is an aspect that could be extended in future research. In this dissertation, I have only tried to draw attention to the fact that there is often a silent assumption, at least in mainstream analytic philosophy of mind, that the issue of personal identity is primary, which is far from being a trivial assumption. The question of how best to describe personal identity under the conditions of phenomenological self-experientialism still requires a more extended discussion, but, for being fruitful, it would presuppose that there is already an agreement over how the issues of self and personal identity relate to each other.

#### 5) *Minimal self and self-feeling*

Just as it is the case with the philosophy of self in general, the debate about the minimal self is endless, or, to put it with words of more dignity, eternal. Not surprisingly, therefore, light is shed on ever new aspects and problems, with the aim to either undermine or corroborate the thesis that all experience involves a basic sense of self. I have addressed recent criticisms that have attacked the minimal self-view for not having sufficiently shown how pre-reflective self-awareness and the being of selfhood hinge together. Defending self-experientialism, my aim was to provide support for the minimal self-view. Doing so, my focus was on very basic forms of self-experience. In future research, one could further extend the discussion on how my notion of the feeling of being oneself, which I argued is integral even to the minimal self, can help to elucidate also other—notably, higher—forms of self-experience, especially also in social and societal contexts. Doing so, further comparison with and differentiation between competing as well as mutually complementing notions of self-feeling could be informed by both the idea of self-experientialism and the notion of the feeling of being oneself.<sup>932</sup>

---

<sup>932</sup> As I have emphasized in Chapter 6 and 7, there are quite heterogeneous approaches to self-feeling that target different aspects of self-experience and also stem from different philosophical traditions. It could be worth trying to integrate the different notions of self-feeling in a comprehensive theory of self-feeling. Hopefully, my concept of the feeling of being oneself could be helpful, given that I have combined different theorems and lines of thinking in order to develop it. Notably, I have built on Lotze's notion of *Fürsichsein* and tried to integrate ideas of it with Husserl-inspired as well as Heidegger-inspired concepts such as the minimal self and existential feelings. Perhaps this attempt for an integrative view may also help to foster more comprehensive systematic and historical debates that are still pending. For instance, Lotze's philosophy (and emphasis on feeling) highly influenced (a) William James (Kraushaar 1936, 1938, 1939,



## 6) *Feeling und sensation: Self-feeling and the cognitive phenomenology debate*

One important aspect in the development of the notion of the feeling of being oneself was the distinction between feeling (*Gefühl*) and sensation (*Empfindung*) that is stressed by proponents of self-feeling such as Lotze or Lipps, among others. Building on this, I argued that all experience involves a feeling of being oneself, which, given that it is to be distinguished from sensation, cannot be described simply by referring to sensory phenomenology. Therefore, what I want to suggest is that the notion of the feeling of being oneself might have implications for the cognitive phenomenology debate, corroborating the idea that emotional phenomenology (broadly construed) involves a sui generis kind of phenomenology that cannot be reduced to sensory phenomenology (or cognition—as the other extreme would have it).<sup>933</sup> Future research could apply arguments in favor of self-feeling to the cognitive phenomenology debate and contrast it with those in favor of positions that maintain that all phenomenology is sensory phenomenology and so present a reductive view of experience.<sup>934</sup>

## 7) *The feeling of being oneself and psychopathology*

After having introduced the notion of feeling of being oneself in the context of my discussions of the minimal self-view, I went on to challenge the notion in the context of a psychopathological phenomenon: depersonalization. I argued that not only is a feeling of being oneself involved in the condition, supporting the view of the entwinement of minimal self and feeling of being oneself, it is also a specific feeling of being oneself that constitutes and thus, at least partially, explains depersonalization. I concluded that my notion of the feeling of being oneself holds up to the challenge of depersonalization. Future research could apply these insights and ways of analysis to other psychopathological phenomena, investigating how the feeling of being oneself varies across different and heterogeneous conditions. One important research topic could consist in carving out how shifts in other

---

1940) and those inspired by him (e.g., Cooley 2009/1922/1902), (b) Theodor Lipps, Max Scheler, and the early phenomenologists or so-called “Munich phenomenology” (Salice 2020/2015, Smid 1982, Vendrell-Ferran 2008, 2015), but also (c) the phenomenologists Husserl and Heidegger (Sullivan 2018/2005). Moreover, Lotze’s theory of self-feeling, given his focus on very basic feelings of pleasure and pain, is also compatible with many of the approaches to self-feeling to be found in the recent scientific literature that focuses on self-feeling in the context of embodied selfhood. Although there are significant differences between these approaches, which could be worked out by shedding light on how each of them involve a theory of self-feeling, the integrative notion of the feeling of being oneself could help to lay bare what they actually share. Similarly, it could be worth investigating whether the notion of the feeling of being oneself (as described here) could build a bridge between recent understandings of self-feeling and those formulated in the context of German Romanticism (Frank 2002).

<sup>933</sup> See, for instance, Montague 2009, Dewalque 2017, Slaby 2017, Vendrell-Ferran 2008, Ch. 4 and 6.

<sup>934</sup> See, for instance, Carruthers & Veillet 2011, Prinz 2011.

aspects of experience—such as intersubjectivity, temporality, embodiment, narrativity, normativity, and others—affect the self-feeling of a person.<sup>935</sup>

---

<sup>935</sup> Cf., for instance, Fuchs 2013b, Fuchs & Pallagrosi 2018, Gallagher 2013, Gallagher & Daly 2018.

## References

- Ackner, B. (1954). Depersonalization. II. Clinical syndromes. *The British Journal of Psychiatry*, 100, 854–872.
- Albahari, M. (2011). Nirvana and ownerless consciousness. In: M. Siderits, E. Thompson & D. Zahavi (eds.): *Self, No Self? Perspectives from Analytical, Phenomenological, and Indian Traditions*. Oxford: Oxford University Press, 79–115.
- Allison, H. E. (1990). *Kant's theory of freedom*. Cambridge: Cambridge University Press.
- Alweiss, L. (2010). Against Cartesian mistrust: Cavell, Husserl and the other mind sceptic. *Ratio*, 23 (3), 241–259.
- Apps, M. A. J. & Tsakiris, M. (2014). The free-energy self. A predictive coding account of self-recognition. *Neuroscience & Biobehavioral Reviews*, 41, 85–97.
- Armstrong, D. M. (1993). *A Materialist Theory of the Mind*. London: Routledge.
- Bayne, T. (2008). The phenomenology of agency. *Philosophy Compass*, 3 (1), 182–202.
- Bayne, T. (2010). *The Unity of Consciousness*. Oxford: Oxford University Press.
- Bayne, T. & Montague, M. (eds.) (2011). *Cognitive Phenomenology*. Oxford: Oxford University Press.
- Beaudry, J.-S. (2019). Are animals persons? Why ask? *Journal of Animal Ethics*, 9 (1), 6–26.
- Becker, O. (2001). Husserl and Descartes. *The New Yearbook for Phenomenology and Phenomenological Philosophy*, I, 351–356.
- Beni, M. D. (2019). *Structuring the Self*. London: Palgrave Macmillan.
- Benovsky, J. (2014). I am a lot of things: A pluralistic account of the self. *Metaphysica*, 15 (1), 113–127.
- Benovsky, J. (2017). Buddhist philosophy and the no-self view. *Philosophy East and West*, 67 (2), 545–553.
- Benovsky, J. (2018). *Mind and Matter. Panpsychism, Dual-Aspect Monism, and the Combination Problem*. Cham: Springer.
- Billon, A. (2014). Why are we certain that we exist? *Philosophy and Phenomenological Research*, 91 (3), 723–759.
- Billon, A. (2016). Making sense of the Cotard syndrome: Insights from the study of depersonalization. *Mind & Language*, 31 (3), 356–391.
- Blackmore, S. (1999). *Blackmore The Meme Machine*. Oxford: Oxford University Press.
- Blackmore, S. (2016). Delusions of consciousness. *Journal of Consciousness Studies*, 23 (11-12), 52–64.
- Blanke, O. & Metzinger, T. (2009). Full-body illusions and minimal phenomenal selfhood. *Trends in Cognitive Science*, 13 (1), 7–13.
- Bockner, S. (1949). The depersonalization syndrome: Report of a case. *The British Journal of Psychiatry*, 95, 968–971.

- Bortolan, A. (2020). Affectivity and the distinction between minimal and narrative self. *Continental Philosophy Review*, 53, 67–84.
- Bower, M. (2014). Husserl's theory of instincts as a theory of affection. *The Journal of the British Society for Phenomenology*, 45 (2), 133–147.
- Breazale, D. (2013). *Thinking Through the Wissenschaftslehre. Themes from Fichte's Early Philosophy*. Oxford: Oxford University Press.
- Cassam, Q. (1997). *Self and World*. Oxford: Oxford University Press.
- Carruthers, P. (2000). *Phenomenal Consciousness: A Naturalistic Theory*. Cambridge: Cambridge University Press.
- Carruthers, P. (2005). *Consciousness: Essays from a Higher-Order Perspective*. Oxford: Oxford University Press.
- Carruthers, P. (2016). Higher-order theories of consciousness. In: E. N. Zalta (ed.): *The Stanford Encyclopedia of Philosophy* (Fall 2016 Edition), URL: <https://plato.stanford.edu/archives/fall2016/entries/consciousness-higher>.
- Carruthers, P. & Veillet, B. (2011). The case against cognitive phenomenology. In: Bayne, T. & Montague, M. (eds.): *Cognitive Phenomenology*. Oxford: Oxford University Press, 35–56.
- Caston, V. (2002). Aristotle on consciousness. *Mind*, 111, 751–815.
- Chalmers, D. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, 2 (3), 200–219.
- Christias, D. (2019). Two modes of givenness of pre-reflective self-consciousness. *Philosophical Explorations*, doi: 10.1080/13869795.2019.1663248.
- Chudnoff, E. (2015). *Cognitive Phenomenology*. Oxon/New York: Routledge.
- Ciaunica, A., Charlton, J. & Farmer, H. (2020). When the window cracks: Transparency and the fractured self in Depersonalisation. *Phenomenology and the Cognitive Sciences*, doi: 10.1007/s11097-020-09677-z.
- Clark, A. (2002). Is seeing all it seems? Action, reason and the grand illusion. *Journal of Consciousness Studies*, 9 (5-6), 181–202.
- Clark, A. (2015). *Surfing Uncertainty. Prediction, Action, and the Embodied Mind*. Oxford: Oxford University Press.
- Clowes, R. W. & Gärtner, K. (2020). The pre-reflective situational self. *Topoi*, 39, 623–637.
- Claesges, U. (1964). *Edmund Husserls Theorie der Raumkonstitution*. Den Haag: Martinus Nijhoff.
- Colombetti, G. (2011). Varieties of pre-reflective self-awareness. Foreground and background bodily feelings in emotion experience. *Inquiry: An Interdisciplinary Journal of Philosophy*, 54 (3), 293–313.
- Colombetti, G. (2014). *The Feeling Body. Affective Science Meets the Enactive Mind*. Cambridge, MA: MIT Press.

- Colombetti, G. & Ratcliffe, M. (2012). Bodily feeling in depersonalization: a phenomenological account. *Emotion Review*, 4, 145.
- Cooley, C. H. (2009/1922/1902). *Human Nature and the Social Order*. 7<sup>th</sup> ed. New Brunswick, NJ: Transaction Publishers.
- Cramer, K. (1974). "Erlebnis". Thesen zu Hegels Theorie des Selbstbewußtseins mit Rücksicht auf die Aporien eines Grundbegriffs nachhegelscher Philosophie. In: H.-G. Gadamer (ed.): *Stuttgarter Hegel-Tage 1970*. Bonn: Hegel-Studien (Beiheft 11), 537–603.
- Crowell, S. (2002). The cartesianism of phenomenology. *Continental Philosophy Review*, 35, 433–454.
- Dainton, B. (2000). *Stream of Consciousness. Unity and Continuity in Conscious Experience*. London & New York: Routledge.
- Dainton, B. (2008). *The Phenomenal Self*. Oxford: Oxford University Press.
- Damasio, A. (1994). *Descartes' Error. Emotion, Reason, and the Human Brain*. London: Picador.
- Damasio, A. (1999). *The Feeling of What Happens. Body, Emotion and the Making of Consciousness*. New York: Harcourt Brace.
- Dennett, D. (1992). The self as a center of narrative gravity. In: F. Kessel, P. Cole & D. Johnson (eds.): *Self and Consciousness: Multiple Perspectives*. Hillsdale, NJ: Erlbaum, 103–118.
- Descartes, R. (1982). *Gespräch mit Burman*. Lat./Germ. Transl. by H. W. Arndt. Hamburg: Meiner.
- Descartes, R. (1997/1641). *Meditations on First Philosophy*. Transl. by L. J. Lafleur. New Jersey: Prentice Hall.
- Dewalque, A. (2017). Emotional Phenomenology: Toward a Nonreductive Analysis. *Midwest Studies in Philosophy*, 41 (1), 27–40.
- Dilthey, W. (1894). *Ideen über eine beschreibende und zergliedernde Psychologie*. Sitzungsberichte der Königlich-Preussischen Akademie der Wissenschaften zu Berlin. Vol. 2. Berlin: Verlag der Königlichen Akademie der Wissenschaften.
- Dreyfus, G. (2011). Self and subjectivity: A Middle Way Approach. In: M. Siderits, E. Thompson & D. Zahavi (eds.): *Self, No Self? Perspectives from Analytical, Phenomenological, and Indian Traditions*. Oxford: Oxford University Press, 114–156.
- Drummond, J. (1984). On seeing a material thing in space: The role of kinaesthesia in visual perception. *Philosophy and Phenomenological Research*, 40, 19–32.
- Duncan, M. (2019). The self shows up in experience. *Review of Philosophy and Psychology*, 10 (2), 299–318.

- Duus, R. E. (2017). Personhood and first-personal experience. *Journal of Theoretical and Philosophical Psychology*, 37 (2), 109–127.
- Engelen, E.-M. (2014). *Vom Leben zur Bedeutung. Philosophische Studien zum Verhältnis von Gefühl, Bewusstsein und Sprache*. Boston and Berlin: De Gruyter.
- Epicurus (1993). *The Essential Epicurus. Letters, Principal Doctrines, Vatican Sayings, and Fragments*. Amherst, NY: Prometheus Books.
- Evans, C. O. (1970). *The Subject of Consciousness*. London a. New York: George Allen a. Humanities Press.
- Farrell, J. & McClelland, T. (2017). Editorial: Consciousness and inner awareness. *Review of Philosophy and Psychology*, 8 (1), 1–22.
- Fasching, W. (2009). The mineness of experience. *Continental Philosophy Review*, 42, 131–148.
- Fasching, W. (2011). ‘I am oft he nature of seeing’: Phenomenological reflections on the Indian notion of witness-consciousness. In: In: M. Siderits, E. Thompson & D. Zahavi (eds.): *Self, No Self? Perspectives from Analytical, Phenomenological & Indian Traditions*. Oxford: Oxford University Press, 193–216.
- Fichte, J. G. (1992). *Foundations of Transcendental Philosophy. (Wissenschaftslehre) Nova Methodo (1796/1799)*. Ithaca a. London: Cornell University Press.
- Foster, J. (1991). *The Immaterial Self. A Defense of the Cartesian Dualist Conception of the Mind*. London a. New York: Routledge.
- Frank, M. (1991a). *Selbstbewußtsein und Selbsterkenntnis*. Stuttgart: Reclam.
- Frank, M. (ed.) (1991b). *Selbstbewußtseinstheorien von Fichte bis Sartre*. Frankfurt a. M.: Suhrkamp.
- Frank, M. (2002). *Selbstgefühl. Eine historisch-systematische Erkundung*. Frankfurt a. M.: Suhrkamp.
- Frank, M. (2015). *Präreflexives Selbstbewusstsein. Vier Vorlesungen*. Stuttgart: Reclam.
- Frankfurt, H. (1987). Identification and wholeheartedness. In: F. D. Schoeman (ed.): *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*. Cambridge: Cambridge University Press, 27–45.
- Frankish, K. (2016). Illusionism as a theory of consciousness. *Journal of Consciousness Studies*, 23 (11-12), 11–39.
- Frankish, K. (ed.) (2017). *Illusionism as a Theory of Consciousness*. Exeter: Imprint.
- Frechette, G. (2013). Searching for the self: Early phenomenological accounts of the self-consciousness from Lotze to Scheler. *International Journal of Philosophical Studies*, 21 (5), 654–679.
- Friston, K. (2010). The free–energy principle: A unified brain theory? *Nature Reviews: Neuroscience*, 11 (2), 127–38.

- Frith, C. (1992). *The Cognitive Neuropsychology of Schizophrenia*. Hove: Psychology Press.
- Frith, C. (2007). *Making up the Mind: How the Brain Creates Our Mental Worlds*. Oxford: Blackwell.
- Fuchs, T. (2006). Implicit and explicit temporality. *Philosophy, Psychiatry & Psychology*, 12 (3), 195–198.
- Fuchs, T. (2012). The feeling of being alive. Organic foundations of self-awareness. In: J. Fingerhut & S. Marienberg (eds.): *Feelings of Being Alive*. Berlin: De Gruyter, 149–165.
- Fuchs, T. (2013a). The phenomenology of affectivity. In: B. Fulford et al. (eds.): *Oxford Handbook of the Philosophy of Psychiatry*. Oxford: Oxford University Press, 612–631.
- Fuchs, T. (2013b). Temporality and psychopathology. *Phenomenology and the Cognitive Sciences*, 12, 75–104.
- Fuchs, T. (2016). Intercorporeality and interaffectivity. *Phenomenology and Mind*, 11, 194–209.
- Fuchs, T. (2017a). *Ecology of the Brain. The Phenomenology and Biology of the Embodied Mind*. Oxford: Oxford University Press.
- Fuchs, T. (2017b). Self across time: the diachronic unity of bodily existence. *Phenomenology and the Cognitive Sciences*, 16 (2), 291–315.
- Fuchs, T. (2020). Delusion, reality, and intersubjectivity: A phenomenological and enactive analysis. *Philosophy, Psychiatry & Psychology*, 27 (1), 61–79.
- Fuchs, T. & Pallagrosi, M. (2018). Phenomenology of temporality and dimensional psychopathology. In: M. Biondi, M. Pasquini & A. Picardi (eds.): *Dimensional Psychopathology*. Cham, Switzerland: Springer, 287–300.
- Gallagher, S. (1986a). Lived body and environment. *Research in Phenomenology*, 16, 139–170.
- Gallagher, S. (1986b). Body image and body schema. A conceptual clarification. *The Journal of Mind and Behavior*, 7 (4), 541–554.
- Gallagher, S. (2005a). *How the Body Shapes the Mind*. New York: Oxford University Press.
- Gallagher, S. (2005b). Metzinger's matrix: Living the virtual life with a real body. *Psyche*, 11 (5), 1–9.
- Gallagher, S. (2013). A pattern theory of self. *Frontiers in Human Neuroscience*, 7, 443.
- Gallagher, S. (2017). Self-defense: Deflecting deflationary and eliminativist critiques of the sense of ownership. *Frontiers in Psychology*, 8:1612, doi: 10.3389/fpsyg.2017.01612.
- Gallagher, S. & Daly, A. (2018). Dynamical relations in self-pattern. *Frontiers in Psychology*, 9: 664, doi: 10.3389/fpsyg.2018.00664.
- Gallagher, S. & Zahavi, D. (2008). *The Phenomenological Mind. An Introduction to Philosophy of Mind and Cognitive Science*. London a. New York: Routledge.

- Garfield, J. L. (2016). Illusionism and givenness. *Journal of Consciousness Studies*, 23 (11-12), 73–82.
- Ganeri, J. (2012). *The Self. Naturalism, Consciousness, and the First-Person Stance*. Oxford: Oxford University Press.
- Ganeri, J. (2017). The self restated. *Philosophical Studies*, 174, 1713–1719.
- Gassendi, P. (2005). Fifth set of objections. From P. Gassendi to that distinguished gentleman René Descartes. In: J. Cottingham, R. Stoothoff, D. Murdoch (eds.): *The Philosophical Writing of Descartes*. Vol. II. Cambridge: Cambridge University Press, 179–240.
- Geniusas, S. (2015). Max Scheler and the stratification of the emotional life. In: L. Hagedorn & J. Dodd (eds.): *The Philosophy of Jan Patočka*. The New Yearbook for Phenomenology and Phenomenological Philosophy. Vol. 14, 355–376.
- Gertler, B. (2011). *Self-knowledge*. New York: Routledge.
- Ghin, M. (2005). What a self could be. *Psyche*, 11 (5), 1–10.
- Giles, J. (1993). The no-self theory: Hume, Buddhism, and personal identity. *Philosophy East and West*, 43 (2), 175–200.
- Gram, M. S. (1981). Intellectual intuition: The continuity thesis. *Journal of the History of Ideas*, 42 (2), 287–304.
- Green, G. W. (2010). Self-consciousness and temporality: Fichte and Husserl. In: V. L. Waibel, D. Breazeale & T. Rockmore (eds.): *Fichte and the Phenomenological Tradition*. Berlin a. New York: De Gruyter, 167–190.
- Grünbaum, T. (2012). First-person and minimal self-consciousness. In: S. Miguens, & G. Preyer (eds.): *Consciousness and Subjectivity*. Heusenstamm: Ontos Verlag, 297–320.
- Guillot, M. (2013). The limits of selflessness: Semantic relativism and the epistemology of *de se* thoughts. *Synthese*, 190 (10), 1793–1816.
- Guillot, M. (2016). Thinking of oneself as the thinker: the concept of self and the phenomenology of intellection. *Philosophical Explorations. An International Journal for the Philosophy of Mind and Action*, 19 (2), 138–160.
- Guillot, M. (2017). I me mine: On a confusion concerning the subjective character of experience. *Review of Philosophy and Psychology*, 8, 23–53.
- Gurwitsch, A. (1974). On thematization. *Research in Phenomenology*, 4, 35–49.
- Halligan, P. W. & Marshall, J. C. (1996). The wise prophet makes sure of the event first: Hallucinations, amnesia and delusions. In: P. W. Halligan & J. C. Marshall (eds.): *Method in Madness: Case Studies in Cognitive Neuropsychiatry*. Hove: Psychology Press, 237–266.
- Hamilton, W. (1870). *Lectures on Metaphysics and Logic*. Vol. 2. Edinburgh a. London: William Blackwood and Sons.



- Harman, G. (2011). The problem with Metzinger. *Cosmos and History: The Journal of Natural and Social Philosophy*, 7 (1), 7–36.
- Hayward, M., Berry, K. & Asthon, A. (2011). Applying interpersonal theories to the understanding of and therapy for auditory hallucinations: A review of the literature and directions for further research. *Clinical Psychology Review*, 31, 1313–1323.
- Head, H. (1920). *Studies in Neurology*. Vol. 2. Oxford: Oxford University Press.
- Heidegger, M. (1996). *Being and Time*. Transl. by J. Stambaugh. New York: State University of New York Press.
- Heidegger, M. (2006/1927). *Sein und Zeit*. 19th ed. Tübingen: Max Niemeyer.
- Heidegger, M. (1979). *Prolegomena zur Geschichte des Zeitbegriffs*. GA 20. Frankfurt a. M.: Klostermann.
- Henrich, D. (1966). Fichtes ursprüngliche Einsicht. In: D. Henrich & H. Wagner (eds.): *Subjektivität und Metaphysik*. Frankfurt am Main: Klostermann, 188–232.
- Henrich, D. (1971). Self-consciousness: A critical introduction to a theory. *Man and World*, 4, 3–28.
- Henry, M. (1973). *The Essence of Manifestation*. Transl. by G. Etzkorn. The Hague: Martinus Nijhoff.
- Hicks, G. D., Laird, J. & Doward, A. (1928). Symposium: the nature of self and self-consciousness. In: *Proceedings of the Aristotelian Society. Supplementary Volumes, Vol. 8: Mind, Matter, and Purpose*, 189–221.
- Himma, E. H. (2005). The problem of explaining phenomenal selfhood: a comment on Thomas Metzinger's self-model theory of subjectivity. *Psyche* 11 (5), 1–19.
- Hobson, A. (2005). Finally some one: Reflection on Thomas Metzinger's "Being No One". *Psyche* 11 (5), 1–8.
- Hohwy, J. (2013). *The Predictive Mind*. Oxford: Oxford University Press.
- Hohwy, J. (2016). The self-evidencing brain. *Noûs*, 50 (2), 259–85.
- Hohwy, J. & Michael, J. (2017). Why should any body have a self? In: F. de Vignemont and A. J. T. Alsmith (eds.): *The Subject's Matter. Self-Consciousness and the Body*. Cambridge, Massachusetts: The MIT Press, 363–391.
- Hood, B. (2011). *The Self Illusion: Why There Is No 'You' Inside Your Head*. London: Constable & Robinson.
- Howell, R. J. (2010). Subjectivity and the elusiveness of the self. *Canadian Journal of Philosophy* 40 (3), 459–483.
- Hume, D. (2007/1739). *A Treatise of Human Nature*. Vol 1. Oxford: Clarendon Press.
- Husserl, E. (1952). *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie. 2. Buch Phänomenologische Untersuchungen zur Konstitution*. Husserliana 4 (Hua 4). Den Haag: Martinus Nijhoff.

- Husserl, E. (1966a). *Zur Phänomenologie des inneren Zeitbewusstseins* (1893-1917). Husserliana 10 (Hua 10). Den Haag: Martinus Nijhoff.
- Husserl, E. (1966b). *Analysen zur passiven Synthesis. Aus Vorlesungs- und Forschungsmanuskripten 1918-1926*. Husserliana 11 (Hua 11). Den Haag: Martinus Nijhoff.
- Husserl, E. (1973a). *Zur Phänomenologie der Intersubjektivität II*. Husserliana 14 (Hua 14). Den Haag: Martinus Nijhoff.
- Husserl, E. (1973b). *Ding und Raum. Vorlesungen 1907*. Husserliana 16 (Hua 16). Den Haag: Martinus Nijhoff.
- Husserl, E. (1976). *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie. I. Buch Allgemeine in die reine Phänomenologie*. Husserliana 3/1 (Hua 3/1). Den Haag: Martinus Nijhoff.
- Husserl, E. (1982). *Cartesian Meditations. An Introduction to Phenomenology*. Transl. by D. Cairns. The Hague: Martinus Nijhoff.
- Husserl, E. (1991). *On the Phenomenology of the Consciousness of Internal time (1893-1917)*. Dordrecht a. o.: Kluwer.
- Husserl, E. (2001). *Die Bernauer Manuskripte über das Zeitbewusstsein (1917/1918)*. Husserliana 33 (Hua 33). Dordrecht a. o.: Kluwer Academic.
- Husserl, E. (2006). *Späte Texte über Zeitkonstitution (1929-1934). Die C-Manuskripte*. Husserliana Materialien 8 (Hua Mat. 8). Dordrecht: Springer.
- Indregard, J. J. (2017). Self-affection and pure intuition in Kant. *Australasian Journal of Philosophy*, 95 (4), 627–643.
- James, W. (1950/1890). *The Principles of Psychology*. Vol. 1. New York: Dover.
- Janet, P. & Raymond, F. (1898). *Névroses et idées fixes*. Vol. 2. Paris: Félix Alcan.
- Kant, I (1996/1781). *Critique of Pure Reason*. Cambridge a. Indianapolis: Hackett.
- Kiverstein, J. (2020). Free energy and the self: An ecological-enactive interpretation. *Topoi*, 39, 559–574.
- Kjosavik, D., Beyer, C. & Fricke, C. (2020). *Husserl's Phenomenology of Intersubjectivity. Historical Interpretations and Contemporary Applications*. London: Routledge.
- Koch, C. (2019). *The Feeling of Life Itself. Why Consciousness is Widespread but Can't be Computed*. Cambridge, MA: MIT Press.
- Kortooms, T. (2002). *Phenomenology of Time. Edmund Husserl's Analysis of Time-Consciousness*. Dordrecht: Kluwer.
- Kraushaar, O. F. (1936). Lotze's influence on the psychology of William James. *Psychological Reviews*, 43, 235–257.
- Kraushaar, O. F. (1938). What James' philosophical orientation owed to Lotze. *Philosophical Review*, 47, 517–526.

- Kraushaar, O. F. (1939). Lotze as a factor in the development of James's Radical Empiricism and Pluralism. *Philosophical Review*, 48, 455–471.
- Kraushaar, O. F. (1940). Lotze's influence on the pragmatism and practical philosophy of William James. *Journal of the History of Ideas*, 1, 439–458.
- Kreuch, G. (2019). *Self-Feeling: Can Self-Consciousness be Understood as a Feeling?* Cham: Springer.
- Kriegel, U. (2004). Consciousness and self-consciousness. *The Monist*, 87 (2), 182–205.
- Kriegel, U. & Williford, K. (eds.) (2006). *Self-Representational Approaches to Consciousness*. Cambridge, MA a. London: The MIT Press.
- Kriegel, U. (2006). The same order monitoring theory of consciousness. In: U. Kriegel & K. Williford (eds.): *Self-Representational Approaches to Consciousness*. Cambridge, MA a. London: The MIT Press, 143–170.
- Kriegel, U. (2009). *Subjective Consciousness: A Self-Representational Theory*. Oxford: Oxford University Press.
- Kriegel, U. (2015). *The Variety of Consciousness*. Oxford: Oxford University Press.
- Krishaber, M. (1783). *De la névropathie cérébro-cardiaque*. Paris: Masson.
- Kristjánsson, K. (2010). *The Self and Its Emotions*. Cambridge: Cambridge University Press.
- Krueger, J. (2011). The who and the how of experience. In: M. Siderits, E. Thompson & D. Zahavi (eds.): *Self, No Self? Perspectives from Analytical, Phenomenological & Indian Traditions*. Oxford: Oxford University Press, 27–55.
- Lakoff, G & Johnson, M. (1980). *Metaphors We Live By*. Chicago: Chicago University Press.
- Lakoff, G. (1987). *Women, Fire, and Dangerous Things. What Categories Reveal About the Mind*. Chicago a. London: University Chicago Press.
- Landgrebe, L. (1961). Husserls Abschied vom Cartesianismus. *Philosophische Rundschau*, 9 (2/3), 133–177.
- Lang, S. (2020). *Phänomenales Bewusstsein und Selbstbewusstsein. Idealistische und selbstrepräsentationalistische Interpretationen*. Hamburg: Meiner.
- Lee, N.-I. (1993). *Edmund Husserls Phänomenologie der Instinkte*. Dordrecht: Springer.
- Legrand, D. (2007). Pre-reflective self-consciousness. On being bodily in the world. *Janus Head*, 9 (2), 493–519.
- Letheby, C. (2020). Being for no-one. Psychedelic experience and minimal subjectivity. *Philosophy and the Mind Sciences*, 1 (I), 5, doi: 10.33735/philisci.2020.I.47.
- Levine, J. (2001). *Purple Haze: The Puzzle of Consciousness*. Oxford: Oxford University Press.
- Lewis, H. D. (1982). *The Elusive Self*. London & Basingstoke: The Macmillan Press.
- Limanowski, J. (2014). What can body ownership illusions tell us about minimal phenomenal selfhood? *Frontiers in Human Neuroscience*, 8:946, doi: 10.3389/fnhum.2014.00946.

- Limanowski, J. & Blankenburg, F. (2013). Minimal self-models and the free energy principle. *Frontiers in Human Neuroscience*, 7: 547, doi: 10.3389/fnhum.2013.00547.
- Lipps, T. (1901). *Das Selbstbewusstsein. Empfindung und Gefühl*. Wiesbaden: Bergmann.
- Lipps, T. (1902). *Vom Fühlen, Wollen und Denken. Eine psychologische Skizze*. Leipzig: Johann Ambrosius Barth.
- Locke, J. (2008/1689). *An Essay concerning Human Understanding*. Oxford: Oxford University Press.
- Lohmar, D. (2002). What does protention 'protend'? *Philosophy Today*, 46 (5), 154–167.
- López-Silva, P. (2019). Me and I are not friends, just acquaintances: on thought insertion and self-awareness. *Review of Philosophy and Psychology*, 10, 319–335.
- Lotz, C. (2007). *From Affectivity to Subjectivity. Husserl's Phenomenology Revisited*. Basingstoke a. New York: Palgrave Macmillan.
- Lotze, R. H. (1852). *Medizinische Psychologie oder Physiologie der Seele*. Leipzig: Weidmann'sche Buchhandlung.
- Lotze, R. H. (1884/1856). *Mikrokosmos. Ideen zur Naturgeschichte und Geschichte der Menschheit. Versuch einer Anthropologie*. Vol. 1. 4th ed. Leipzig: S. Hirzel.
- Lotze, R. H. (1885). *Microcosmus. An Essay Concerning Man and His Relation to the World*. English Translation in two volumes by E. Hamilton & E. E. C. Jones. Edinburgh: T. and T. Clark.
- Lotze, R. H. (1888/1864<sup>1</sup>). *Mikrokosmos. Ideen zur Naturgeschichte und Geschichte der Menschheit. Versuch einer Anthropologie*. Vol. 3. 4th ed. Leipzig: S. Hirzel.
- Lotze, R. H. (1889/1881). *Grundzüge der Psychologie: Dictate aus den Vorlesungen*. 4th ed. Leipzig: S. Hirzel.
- Lotze, R. H. (1989/1846). Seele und Seelenleben. In: *Kleine Schriften zur Psychologie*. Berlin: VEB Deutscher Verlag der Wissenschaften, 112–234.
- Luft, S. (2004). Husserl's theory of the phenomenological reduction: between life-world and Cartesianism. *Research in Phenomenology*, 34, 198–234.
- Lutz, A. & Thompson, E. (2003). Neurophenomenology. Integrating subjective experience and brain dynamics in the neuroscience of consciousness. *Journal of Consciousness Studies* 10 (9-10), 31–52.
- MacDonald, P. S. (1999). *Descartes and Husserl: The Philosophical Project of Radical Beginnings*. Albany: State University of New York Press.
- MacIntyre, A. (2007/1981). *After Virtue. A Study in Moral Theory*. 3<sup>rd</sup> ed. Notre Dame, IN: University of Notre Dame Press.
- Madell, G. (2015). *The Essence of the Self. In Defense of the Simple View of Personal Identity*. New York a. Oxon: Routledge.

- Maier, H. (1908). *Psychologie des emotionalen Denkens*. Tübingen: J. C. B. Mohr (Paul Siebeck).
- Martin, W. M. (2008). Descartes and the phenomenological tradition. In: J. Broughton & J. Carriero (eds.): *A Companion to Descartes*. Oxford: Blackwell, 496–512.
- McClelland, T. (2019). Against Virtual Selves. *Erkenntnis*, 84 (1), 21–40.
- McClelland, T. (forthcoming a). Self-representational theories of consciousness. In: U. Kriegel (ed.): *The Oxford Handbook of the Philosophy of Consciousness*. Oxford: Oxford University Press.
- McClelland, T. (forthcoming b). Four impediments to the case for mineness. In: M. Guillot & Manuel García-Carpintero (eds.): *Consciousness and the Sense of Mineness*. Oxford: Oxford University Press.
- McGilvary, E. B. (1907). The stream of consciousness. *The Journal of Philosophy, Psychology and Scientific Methods*, 4 (9), 225–235.
- McLaughlin, B. (2009). Monothematic Delusions and Existential Feelings. In: T. Bayne & J. Fernández (eds.): *Delusion and Self-deception: Affective and Motivational Influences on Belief Formation*. New York: Psychology Press, 139–164.
- Mehl, E. (2021). Phenomenology and the Cartesian tradition. In: D. de Santis, B. C. Hopkins, C. Majolino (eds.): *The Routledge Handbook of Phenomenology and Phenomenological Philosophy*. Oxon a. New York: Routledge, Ch. 4.
- Melle, U. (1983). *Das Wahrnehmungsproblem und seine Verwandlung in phänomenologischer Einstellung. Untersuchungen zu den phänomenologischen Wahrnehmungstheorien von Husserl, Gurwitsch und Merleau-Ponty*. The Hague a. o.: Martinus Nijhoff.
- Metzinger, T. (2003). *Being No One. The Self-Model Theory of Subjectivity*. Cambridge, MA: The MIT Press.
- Metzinger, T. (2005). Précis: being no one. *Psyche* 11 (5), 1–35.
- Metzinger, T. (2009a). *The Ego Tunnel*. New York: Basic Books.
- Metzinger, T. (2009b). Enlightenment 2.0: Interview with Thomas Metzinger. *Collapse* V, 187–215.
- Metzinger, T. (2011). The no-self alternative. In: S. Gallagher (ed.): *The Oxford Handbook of Self*. Oxford: Oxford University Press, 280–295.
- Metzinger, T. (2013). Why are dreams interesting for philosophers? The example of minimal phenomenal selfhood, plus an agenda for future research. *Frontiers in Psychology*, 4: 746, doi: 10.3389/fpsyg.2013.00746/full.
- Millière, R. (2017). Looking for the self: phenomenology, neurophysiology and philosophical significance of drug-induced ego dissolution. *Frontiers in Human Neuroscience*, 11, 245, doi: 10.3389/fnhum.2017.00245.
- Moran, D. (2000). *Introduction to Phenomenology*. London a. New York: Routledge.

- Moran, D. (2017). Intercorporeality and intersubjectivity: A phenomenological exploration of embodiment. In: C. Durt, T. Fuchs & C. Tewes (eds): *Embodiment, Enaction, and Culture: Investigating the Constitution of the Shared World*. MIT Press, Cambridge: MIT Press, 25–40.
- Montague, M. (2009). The logic, intentionality, and phenomenology of emotion. *Philosophical Studies*, 145, 171–192.
- Montague, M. (2016). *The Given. Experience and its Content*. Oxford: Oxford University Press.
- Mula, M., Pini, S., Calugi, S., Preve, M., Masini, M. et al. (2020). Distinguishing affective depersonalization from anhedonia in major depression and bipolar disorder. *Comprehensive Psychiatry*, 51, 187–192.
- Musholt, K. (2013). Self-consciousness and nonconceptual content. *Philosophical Studies*, 163 (3), 649–672.
- Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, 83, 435–450.
- Nida-Rümelin, M. (2017). Self-awareness. *Review of Philosophy and Psychology*, 8, 55–82.
- Niikawa, T. (2021). Illusionism and definitions of phenomenal consciousness. *Philosophical Studies*, 178, 1–21.
- Noonan, H. (2003). *Personal Identity*. 2<sup>nd</sup> ed. London: Routledge & Kegan Paul.
- Northoff, G. (2004a). *Philosophy of the Brain: The Brain Problems*. Amsterdam: John Benjamins.
- Northoff, G. (2004b). Why do we need a philosophy of the brain? *Trends in Cognitive Science*, 8, 102–107.
- O’Conaill, D. (2019). Subjectivity and mineness. *Erkenntnis*, 84, 325–341.
- Oesterreich, K. (1910). *Die Phänomenologie des Ich in ihren Grundproblemen*. 1. Bd.: Das Ich und das Selbstbewußtsein. Die scheinbare Spaltung des Ich. Leipzig: Johann Ambrosius Barth.
- O’Farrell, F. (1979). Intuition in Kant’s theory of knowledge. Part 1: Intuition in general in Kant’s theory. *Gregorianum*, 60 (3), 481–511.
- Olson, E. T. (1998). There is no problem of the self. *Journal of Consciousness Studies*, 5 (5-6), 645–657.
- Olson, E. T. (2007). *What Are We? A Study in Personal Ontology*. Oxford: Oxford University Press.
- Olson, E. T. (2008). *The Human Animal. Personal Identity Without Psychology*. New York a. Oxford: Oxford University Press.
- Panksepp, J. (1998). The periconscious substrates of consciousness: Affective states and the evolutionary origins of the self. *Journal of Consciousness Studies* 5 (5-6), 566–582.
- Parfit, D. (1971). Personal identity. *Philosophical Review*, 80, 3–27.
- Parfit, D. (1987). *Reasons and Persons*. Oxford: Clarendon Press.

- Peacocke, C. (2014). *The Mirror of the World. Subjects, Consciousness and Self-Consciousness*. Oxford: Oxford University Press.
- Peacocke, C. (2019). *The Primacy of Metaphysics*. Oxford: Oxford University Press.
- Perry, J. (1976). The importance of being identical. In: A. Rorty (ed.): *The Identities of Persons*. Berkely, CA: University of California Press.
- Persson, I. (2004). Self-doubt: why we are not identical to things of any kind. *Ratio*, 17 (4), 390–408.
- Pitt, D. (2004). The phenomenology of cognition, or What is it like to think that *p*? *Philosophy and Phenomenological Research*, 69 (1), 1–36.
- Pitt, D. (2018). Mental representation. In: E. N. Zalta (ed.): *The Stanford Encyclopedia of Philosophy* (Winter 2018 Edition) [URL: <https://plato.stanford.edu/archives/win2018/entries/mental-representation>].
- Price, H. H. (1940). *Hume's Theory of the External World*. Oxford: Oxford University Press.
- Prinz, J. (2011). The sensory basis of cognitive phenomenology. In: T. Bayne & M. Montague (eds.): *Cognitive Phenomenology*. Oxford: Oxford University Press, 174–196.
- Ramm, B. (2017). Self-experience. *Journal of Consciousness Studies*, 24 (11-12), 142–166.
- Ratcliffe, M. (2005). The feeling of being. *Journal of Consciousness Studies*, 12 (8-10), 43–60.
- Ratcliffe, M. J. (2008). *Feelings of Being. Phenomenology, Psychiatry and the Sense of Reality*. Oxford: Oxford University Press.
- Ratcliffe, M. J. (2010). Depression, guilt and emotional depth. *Inquiry: An Interdisciplinary Journal of Philosophy*, 53 (6), 602–626.
- Ratcliffe, M. J. (2012). The phenomenology of existential feeling. In: J. Fingerhut & S. Marienberg (eds.): *Feelings of Being Alive*. Berlin: De Gruyter, 23–53.
- Ratcliffe, M. J. (2013a). What is it to lose hope? *Phenomenology and the Cognitive Sciences*, 12, 597–614.
- Ratcliffe, M. J. (2013b). Some Husserlian reflections on the contents of experience. In: M. C. Haug (ed.): *Philosophical Methodology: the Armchair or the Laboratory?* London: Routledge, 353–378.
- Ratcliffe, M. J. (2016). Existential feeling and narrative. In: T. Breyer & O. Müller (eds.): *Funktionen des Lebendigen*. Berlin: de Gruyter, 169–192.
- Ratcliffe, M. J. (2017). *Real Hallucinations: Psychiatric Illness, Intentionality, and the Interpersonal World*. Cambridge, MA: The MIT Press.
- Ratcliffe, M. J. (2020). Existential feelings. In: T. Szanto & H. Landweer (eds.): *The Routledge Handbook of Phenomenology of Emotion*, 250–261.

- Reis Piedade, J. I. (2002): *Der bewegte Leib. Kinästhesen bei Husserl im Spannungsfeld von Intention und Erfüllung*. Wien: Passagen.
- Ricoeur, P. (1988). *Time and Narrative*. Chicago, IL: University of Chicago Press.
- Ricoeur, P. (1991). Narrative identity. *Philosophy Today*, 35 (1), 73–81.
- Ricoeur, P. (1992). *Oneself as Another*. Chicago, IL: University of Chicago Press.
- Rinofner, S. (2005). The limits of representationalism. A phenomenological critique of Thomas Metzinger's self-model theory. *Synthesis Philosophica*, 40, 355–371.
- Roberts, W. W. (1960). Normal and abnormal depersonalization. *The British Journal of Psychiatry*, 106 (443), 478–493.
- Römer, I. (2010). *Das Zeitdenken bei Husserl, Heidegger und Ricoeur*. Dordrecht: Springer.
- Römpp, G. (1992). *Husserls Phänomenologie der Intersubjektivität. Und ihre Bedeutung für eine Theorie intersubjektiver Objektivität und die Konzeption einer phänomenologischen Philosophie*. Dordrecht: Kluwer Academic Publishers.
- Romano, C. (2012). Must phenomenology remain Cartesian? *Continental Philosophy Review*, 45, 425–445.
- Rosenthal, D. M. (1993). Higher-order thoughts and the appendage theory of consciousness. *Philosophical Psychology*, 6, 155–166.
- Rosenthal, D. M. (1997). A theory of consciousness. In: N. Block, O. Flanagan & G. Güzeldere (eds.): *The Nature of Consciousness*. Cambridge, MA: MIT Press.
- Rousse, B. S. (2019). Self-awareness and self-understanding. *European Journal of Philosophy*, 27 (1), 162–186.
- Rudd, A. (2009). In defence of narrative. *European Journal of Philosophy*, 17 (1), 60–75.
- Rudd, A. (2012). *Self, Value, and Narrative. Kierkegaardian Approach*. Oxford: Oxford University Press.
- Russell, J. A. (2009). Emotion, core affect, and psychological construction. *Cognition & Emotion*, 23 (7), 1259–1283.
- Salice, A. (2020/2015). The phenomenology of the Munich and Göttingen circles. *The Stanford Encyclopedia of Philosophy* (Winter 2020 Edition), E. N. Zalta (ed.), <https://plato.stanford.edu/archives/win2020/entries/phenomenology-mg>.
- Sartre, J.-P. (1991/1936). *The Transcendence of the Ego. An Existentialist Theory of Consciousness*. New York: Hill and Wang.
- Sartre, J.-P. (1956). *Being and Nothingness. An Essay on Phenomenological Ontology*. New York: Philosophical Library.
- Sass, L. A. (1994). *The Paradoxes of Delusion*. London: Cornell University Press.
- Sass, L. A. (2014). Explanation and description in phenomenological psychopathology. *Journal of Psychopathology*, 20, 366–376.



- Schear, J. K. (2009). Experience and self-consciousness. *Philosophical Studies*, 144, 95–105.
- Schechtman, M. (1996). *The Constitution of Selves*. Ithica, NY: Cornell University Press.
- Schechtman, M. (2001). Empathic access: the missing ingredient in personal identity. *Philosophical Explorations*, 4 (2), 95–111.
- Schechtman, M. (2007). Stories, lives, and basic survival: a refinement and defense of the narrative view. In: D. Hutto (ed.): *Narrative and Understanding Persons*. Cambridge: Cambridge University Press, 155–178.
- Scheler, M. (1921). *Der Formalismus in der Ethik und die materiale Wertethik. Neuer Versuch der Grundlegung eines ethischen Personalismus*. 2<sup>nd</sup> ed. Halle a. d. S.: Max Niemeyer.
- Scheler, M. (1979). Tod und Fortleben. In: *Die Zukunft des Kapitalismus und andere Aufsätze*. München: Francke, 7–71.
- Schellenberg, S. (2016). *De se* content and *de hinc* content. *Analysis*, 76, 334–345.
- Schilder, P. (1923). *Das Körperschema. Ein Beitrag zur Lehre vom Bewusstsein des eigenen Körpers*. Berlin: Springer.
- Schilder, P. (1951). *Psychoanalysis, Man, and Society*. New York: Norton.
- Schlicht, T. (2018). Experiencing organisms: form mineness to subject of experience. *Philosophical Studies*, 175, 2447–2474.
- Schmidt, P. (2018). The relevance of explanatory first-person approaches (EFPA) for understanding psychopathological phenomena: The role of phenomenology. *Frontiers in Psychology*, 9:694, doi: 10.3389/fpsyg.2018.00694.
- Schmidt, P. (2019). Zeit und Leib bei Kant. Ein Blick auf die ‘Widerlegung des Idealismus’ im Lichte von Husserls Phänomenologie der Zeit. In: V. Waibel & M. Ruffing (eds.): *Akten des 12. Internationalen Kant-Kongresses ‘Natur und Freiheit’ in Wien vom 21.-25. September 2015*. Berlin: De Gruyter, 3591–3599.
- Schneider, K. (1920). Die Schichtung des emotionalen Lebens und der Aufbau der Depressionszustände. *Zeitschrift für die gesamte Neurologie und Psychiatrie*, 59, 281–286.
- Schwerin, A. (2007). Hume and the self: A critical response. *Journal of Scottish Philosophy*, 5 (1), 15–30.
- Segal, S. (1998). *Collision with the Infinite. A Life Beyond the Personal Self*. San Diego: Blue Dove Press.
- Serrahima Balius, C. (2019). *My Body is the Subject’s Body. In Defence of Experientialism about the Sense of Bodily Ownership*. PhD Dissertation. Universitat de Barcelona, Facultat de Filosofia.
- Seth, A. & Tsakiris, M. (2018). Being a beast machine: The somatic basis of selfhood. *Trends in Cognitive Sciences*, 22 (11), 969–981.

- Shear, J. (ed.) (1999). *Explaining Consciousness: The Hard Problem*. Cambridge, MA a. London: The MIT Press.
- Shoemaker, S. (1968). Self-reference and self-awareness. *The Journal of Philosophy*, 65, 556–579.
- Shoemaker, S. (1984). Personal identity: A materialist's account. In: S. Shoemaker & R. Swinburne (eds.): *Personal Identity*. Oxford: Basil Blackwell, 67–138.
- Shoemaker, S. (1994). Introspection and the self. In: Q. Cassam (ed.): *Self-knowledge*. Oxford: Oxford University Press, 118–139.
- Shorvon, H. (1946). The depersonalization syndrome. *Proceedings of the Royal Society of Medicine*, 39 (12), 779–791.
- Siegel, S. (2010). *The Contents of Visual Experience*. Oxford: Oxford University Press.
- Sierra, M. & David, A. S. (2011). Depersonalization: A selective impairment of self-awareness. *Consciousness and Cognition*, 20 (1), 99–108.
- Sierra, M. & Berrios, D. E. (2001). The phenomenological stability of depersonalization: Comparing the old with the new. *Journal of Nervous and Mental Disease*, 189 (9), 1107–1113.
- Siewert, C. (1998). *The Significance of Consciousness*. Princeton, NJ: Princeton University Press.
- Siewert, C. (2011). Phenomenal thought. In: T. Bayne & M. Montague (eds.): *Cognitive Phenomenology*. Oxford: Oxford University Press, 236–267.
- Sigwart, C. (1881). *Kleine Schriften. Zweite Reihe. Zur Erkenntnislehre und Psychologie*. Freiburg: Mohr.
- Simeon, D. & Abugel, J. (2006). *Feeling Unreal: Depersonalization Disorder and the Loss of Self*. Oxford: Oxford University Press.
- Slaby, J. (2008). Affective intentionality and the feeling body. *Phenomenology and the Cognitive Sciences*, 7 (4), 429–444.
- Slaby, J. (2012). Affective self-construal and the sense of ability. *Emotion Review*, 4 (2), 151–156.
- Slaby, J. (2017). More than a feeling: Affect as radical situatedness. *Midwest Studies in Philosophy*, 41 (1), 7–26.
- Slaby, J. & Stephan, A. (2008). Affective intentionality and self-consciousness. *Consciousness and Cognition*, 17 (2), 506–513.
- Smid, R. N. (1982). “Münchener Phänomenologie” – Zur Frühgeschichte des Begriffs. In: H. Spiegelberg & E. Avé-Lallemant (eds.): *Pfänder-Studien*. *Phaenomenologica* 84. The Hague a. o.: Martinus Nijhoff, 109–154.
- Soffer, W. (1981). Husserl's neo-Cartesianism. *Research in Phenomenology*, 11, 141–158.
- Staiti, A. (2015). On Husserl's alleged Cartesianism and Conjunctivism: a critical reply to Claude Romano. *Husserl Studies*, 31, 123–141.

- Staudigl, M. (2012). From the ‘metaphysics of the individual’ to the critique of society: on the practical significance of Michel Henry’s phenomenology of life. *Continental Philosophy Review*, 45, 339–361.
- Strasser, S. (1956). *Das Gemüt. Grundgedanken zu einer phänomenologischen Philosophie und Theorie des menschlichen Gefühlslebens*. Utrecht: Het Spectrum, 109–160.
- Strasser, S. (1977). *Phenomenology of Feeling: An Essay on the Phenomena of the Heart*. Pittsburgh: Duquesne University Press.
- Strawson, G. (1997). The self. *Journal of Consciousness Studies* 4, 405–428.
- Strawson, G. (1999). The self and the SESMET. *Journal of Consciousness Studies*, 6 (4), 99–135.
- Strawson, G. (2009). *Selves: An Essay in Revisionary Metaphysics*. Oxford: Oxford University Press.
- Strawson, G. (2010). *Mental Reality*. 2<sup>nd</sup> edition. Cambridge, MA: The MIT Press.
- Strawson, G. (2011). Cognitive phenomenology: Real life. In: T. Bayne & M. Montague (eds.): *Cognitive Phenomenology*. Oxford: Oxford University Press, 285–325.
- Strawson, G. (2017). *The Subject of Experience*. Oxford: Oxford University Press.
- Sullivan, D. (2018/2005). Hermann Lotze. *The Stanford Encyclopedia of Philosophy* (Winter 2018 Edition), E. N. Zalta (ed.), <https://plato.stanford.edu/archives/win2018/entries/hermann-lotze>.
- Summa, M. (2014). *Spatio-Temporal Intertwining. Husserl’s Transcendental Aesthetics*. *Phaenomenologica* 213. Dordrecht: Springer.
- Swinburne, R. (2019). *Are We Bodies or Souls?* Oxford: Oxford University Press.
- Taipale, J. (2014). *Phenomenology and Embodiment. Husserl and the Constitution of Subjectivity*. Evanston, IL: Northwestern University Press.
- Taipale, J. (2015). Beyond cartesianism: Body-perception and the immediacy of empathy. *Continental Philosophy Review*, 48, 161–178.
- Textor, M. (2018). ‘Enjoy your self’: Lotze on self-concern and self-consciousness. *History of Philosophy Quarterly*, 35 (2), 157–177.
- Taylor, C. (1989). *Sources of the Self*. Cambridge: Cambridge University Press.
- Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of the Mind*. Cambridge, MA: Harvard University Press.
- Thompson, H. (2005). Review of “Being No-one: The Self-Model Theory of Subjectivity”, *Essays in Philosophy*, 6 (1), Article 32.
- Tooley, M. (2012). Are nonhuman animals persons? In: T. L. Beauchamp & R. G. Frey (eds.): *The Oxford Handbook of Animal Ethics*. Oxford Handbook Online, doi: 10.1093/oxfordhb/9780195371963.013.0013.

- Trehub, A. (2007). Space, self, and the theater of consciousness. *Consciousness and Cognition*, 16, 310–330.
- Tugendhat, E. (1979). *Selbstbewußtsein und Selbstbestimmung*. Frankfurt a.M.: Suhrkamp.
- Turausky, K. E. (2014). Wherever you go, there you are: On individuating subjective phenomenology. *Essays in Philosophy*, 15 (2), Art. 2, doi: 10.7710/1526-0569.1506.
- Tye, M. (1995). *Ten Problems of Consciousness*. Cambridge, MA: The MIT Press.
- Van Gulick, R. (2001). Inward and upward: reflection, introspection, and self-awareness. *Philosophical Topics*, 28, 275–305.
- Van Gulick, R. (2004). Higher-order global states (HOGS): An alternative higher-order model of consciousness. In: R. Gennaro (ed.): *Higher-Order Theories of Consciousness: An Anthology*. Amsterdam: John Benjamins, 67–92.
- Van Gulick, R. (2006). Mirror mirror—Is that all? In: U. Kriegel & K. Williford (ed.): *Self-Representational Approaches to Consciousness*. Cambridge, Mass.: MIT Press, 11–39.
- Vendrell-Ferran, Í (2008). *Die Emotionen. Gefühle in der realistischen Phänomenologie*. Berlin: Akademie-Verlag.
- Vendrell-Ferran, Í (2015). The emotions in early phenomenology. *Studia phaenomenologica*, 15, 349–374.
- Vendrell-Ferran, Í (2020a). Else Voigtländer. In: H. Landweer & T. Szanto (eds.): *The Routledge Handbook of Phenomenology of Emotion*. London: Routledge.
- Vendrell-Ferran, Í (2020b). Feeling of self-worth in Else Voigtländer. *Encyclopedia of Concise Concepts by Women Philosophers*.
- Vendrell-Ferran, Í (2021). Envy, powerlessness, and the feeling of self-worth. In: A. Bortolan & E. Magrí (eds.): *Empathy, Intersubjectivity, and the Social World. The Continued Relevance of Phenomenology. Essays in Honour of Dermot Moran*. Berlin: De Gruyter.
- Vendrell-Ferran, Í (forthcoming). How to understand feelings of vitality: An approach to their nature, varieties and functions. In: S. Ferrarello (ed.): *Phenomenology of Bioethics: Technoethics and Lived Experience*. Cham: Springer.
- Voigtländer, E. (1910). *Vom Selbstgefühl. Ein Beitrag zur Förderung psychologischen Denkens*. Leipzig: R. Voigtländer.
- Volkelt, J. (1925). *Phänomenologie und Metaphysik der Zeit*. München: C. H. Beck.
- Waibel, V. L., Breazeale, J. D. & Rockmore, T. (eds.) (2010). *Fichte and the Phenomenological Tradition*. Berlin a. New York: De Gruyter.
- Warren, N. de (2009). *Husserl and the Promise of Time. Subjectivity in Transcendental Phenomenology*. Cambridge: Cambridge University Press.

- Washburn, M. C. (1976). Did Kant have a theory of self-knowledge? *Archiv für Geschichte der Philosophie*, 58 (1), 50–56.
- Watzl, S. (2018). Consciousness and no self? *Ratio*, 31 (4), 363–375.
- Welton, D. (2002): Zur Räumlichkeit des Bewußtseins. In: H. Hüni u. P. Trawny (eds.): *Die erscheinende Welt*. Berlin: Duncker & Humblot, 623–649.
- World Health Organization (1993). *The ICD-10 classification of mental and behavioural disorders*. Genève: World Health Organization.
- Wiese, W. (2018). *Experienced Wholeness. Integrating Insights from Gestalt Theory, Cognitive Neuroscience, and Predictive Processing*. Cambridge, MA a. London: The MIT Press.
- Wiese, W. (2019). Explaining the enduring intuition of substantiality. The phenomenal self as an abstract ‘salience object’. *Journal of Consciousness Studies*, 26 (3-4), 64–87.
- Wundt, W. (1874). *Grundzüge der physiologischen Psychologie*. Leipzig: W. Engelmann.
- Zahavi, D. (1996). *Husserl und die transzendente Intersubjektivität. Eine Antwort auf die sprachpragmatische Kritik*. Phaenomenologica 135. Dordrecht a.o.: Kluwer.
- Zahavi, D. (1999). *Self-Awareness and Alterity. A Phenomenological Investigation*. Evanston, Illinois: Northwestern University Press.
- Zahavi, D (2003). Inner time-consciousness and pre-reflective self-awareness. In: D. Welton (ed.): *The New Husserl. A Critical Reader*. Indiana University Press, 157–180.
- Zahavi, D. (2004). Time and Consciousness in the Bernau Manuscripts. *Husserl Studies*, 20/2, 99–118.
- Zahavi, D. (2005). *Subjectivity and Selfhood. Investigating the First-Person Perspective*. Cambridge, MA a. London: The MIT Press.
- Zahavi, D. (2007a). The Heidelberg School and the limits of reflection. In: S. Heinämaa, V. Lähteenmaki & P. Remes (eds.): *Consciousness: From Perception to Reflection in the History of Philosophy*, 267–285.
- Zahavi, D. (2007b). Perception of duration presupposes duration of perception – or does it? Husserl and Dainton on time. *International Journal of Philosophical Studies*, 15 (3), 453–471.
- Zahavi, D. (2009). Is the self a social construct? *Inquiry*, 52 (6), 551–573.
- Zahavi, D. (2010). Inner (time-)consciousness. In: D. Lohmar & I. Yamaguchi (eds.): *On Time – New Contributions to the Husserlian Phenomenology of Time*. Phaenomenologica 197. Springer: Dordrecht, 319–339.
- Zahavi, D. (2011a). Empathy and direct social perception: A phenomenological proposal. *Review of Philosophy and Psychology*, 2 (3), 541–558.
- Zahavi, D. (2011b). The experiential self: Objections and clarifications. In: M. Siderits, E. Thompson & D. Zahavi (eds.): *Self, no self? Perspectives from Analytical, Phenomenological & Indian Traditions*. Oxford: Oxford University Press, 56–78.

- Zahavi, D. (2014a). *Self and Other. Exploring Subjectivity, Empathy, and Shame*. Oxford: Oxford University Press.
- Zahavi, D. (2014b). Empathy and other-directed intentionality. *Topoi*, 33, 129–142.
- Zahavi, D. (2017a). Ownership, memory, attention: Commentary on Ganeri. *Australasian Philosophical Review*, 1 (4), 406–415.
- Zahavi D (2017b). Thin, thinner, thinnest: defining the minimal self. In: C. Durt, T. Fuchs & C. Tewes (eds): *Embodiment, Enaction, and Culture: Investigating the Constitution of the Shared World*. MIT Press, Cambridge: MIT Press, 193–200.
- Zahavi, D. (2018). Consciousness, self-consciousness, selfhood: A reply to some critics. *Review of Philosophy and Psychology*, 9 (3), 703–718.
- Zahavi, D. (2020a). Reflexivity, transparency, and illusionism: Engaging Garfield. *ProtoSociology – An International Journal and Interdisciplinary Project*, 36, 142–156.
- Zahavi, D. (2020b). Consciousness and selfhood: Getting clearer on for-me-ness and mineness. In: U. Kriegel (ed.): *The Oxford Handbook of the Philosophy of Consciousness*. Oxford: Oxford University Press, 635–653.
- Zahavi, D. & Kriegel, U. (2016). For-me-ness: What it is and what it is not. In: D. Dahlstrom, A. Elpidorou & W. Hopp (eds.): *Philosophy of Mind and Phenomenology*. New York a. Oxon: Routledge, 36–53.
- Ziegler, T. (1893). *Das Gefühl. Eine psychologische Untersuchung*. Stuttgart: G. J. Göschen'sche Verlagshandlung.
- Zippel, N. (2014): The hyletic time-consciousness and the embodied subject. In: K. Novotný, P. Rodrigo, J. Slatman & S. Stoller (eds.): *Corporeity and Affectivity. Dedicated to Maurice Merleau-Ponty*. Leiden: Brill, 35–48.