Body, Brain, and Situation in Sensorimotor Enactivism

María Jimena Clavel Vázquez

University of
St Andrews

This thesis is submitted in partial fulfilment for the degree of

Doctor of Philosophy (PhD)

at the University of St Andrews

August 2020

Candidate's declaration

I, María Jimena Clavel Vázquez, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 80,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree.

I was admitted as a research student at the University of St Andrews in September 2016.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

October 15, 2020

Date Signature of candidate

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

October 15, 2020

Date Signature of supervisor

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, María Jimena Clavel Vázquez, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

No embargo on print copy.

Electronic copy

No embargo on electronic copy.

October 15, 2020

Date Signature of candidate

October 15, 2020

Date Signature of supervisor

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, María Jimena Clavel Vázquez, hereby certify that no requirements to deposit original research data or digital outputs apply to this thesis and that, where appropriate, secondary data used have been referenced in the full text of my thesis.

October 15, 2020

Date Signature of candidate

Abstract

This thesis tackles issues that arise around Sensorimotor Enactivism's (SMEn) conception of embodiment. SMEn is a theory that claims that perception is something perceivers do, that requires the possession and execution of practical knowledge about the way sensory information changes after an interaction. One of the key commitments of this view is to the idea that this knowledge—sensorimotor knowledge—depends on details of the body of the perceiver.

The thesis is divided in four parts, each addressing different challenges. Part I sets the stage for the problems that are discussed in the rest of the thesis. I examine SMEn's claims that perception is active and knowledgeable, as well as its position regarding representations and externalism.

In Part II, I discuss the concern that SMEn is chauvinistic in that it claims that only creatures with bodies like ours can have experiences like ours. I argue that SMEn cannot be said to be chauvinistic because, even if its embodiment claim involves a commitment to an irreducibly perspectival aspect of perceptual experience, Sensorimotor Contingencies (SMCs) are structural descriptions of perceptual systems that already involve information about objects' invariant features.

In Part III, I defend a predictive approach to SMEn to address the concern that, while it claims that perception is embodied, it leaves out one crucial element of the material basis of perceptual experience: the brain. Although the predictive framework is standardly taken to involve a representational and internalist profile, I argue in favour of the compatibility between SMEn and the Free Energy Approach (FEA).

Part IV addresses a concern that is based on the phenomenological analysis of the body, according to which the body is marked by someone's social and historical circumstances. I argue that the predictive version of SMEn makes plausible the idea that perceptual experience is social all the way down.

Para Godi y

para Rosita.

General acknowledgements

Several sections throughout the thesis, in particular in chapters 1, 2, 5, and 6, are taken word for word from the paper "A match made in heaven: predictive approaches to (an unorthodox) sensorimotor enactivism" (Clavel Vázquez, 2019), published in open access online in *Phenomenology and the Cognitive Sciences*. This paper is licensed under the Creative Commons Attribution 4.0 International License. This license permits unrestricted use, distribution, modification, and reproduction in any medium of the material, provided that appropriate acknowledgement is given to the original work and publication source, and that a link to and notice of the license are included, indicating too if changes were made. A copy of the Creative Commons license can be found here: http://creativecommons.org/licenses/by/4.0/. Although most of the text in question has been left untouched, in some cases I have made modifications to adjust the arguments and writing to reflect the position defended in the thesis, as well as to reflect changes in the debate.

For the phenomenological analysis that appears in Chapters 7 and 8, as well as for the conclusion, I have occasionally drawn on the manuscript "Embodying Imagination" co-authored with Adriana Clavel-Vázquez, currently under review.

Funding

This work was supported by Mexico's National Council of Science and Technology (Consejo Nacional de Ciencia y Tecnología, CONACYT) under PhD scholarship 295962/440722; by a SASP (St Andrews and Stirling Philosophy Graduate Programme) Maintenance Studentship; and by Mexico's Secretariat of Education (SEP, Secretaría de Educación Pública).

Acknowledgements

This thesis and research project would not have been possible without the continuous support, invaluable advice, and encouragement from Prof Michael Wheeler. His work and feedback have shaped my philosophical path. He has truly been the ideal supervisor.

I also want to thank the incredible support and feedback from Dr Alisa Mandrigin. Her careful reading and discussion were essential to the development of many important arguments in this thesis.

I want to thank Andrés Hernández for our long, fun, and relentless philosophical discussions. For his invaluable feedback, for always listening to me with unending patience, helping me figure out what it is that I want to say. Thanks for all the encouragement, support, and love, without which I would not have been able to complete this thesis. Thanks for sharing with me this incredible adventure.

Thanks to Adriana Clavel-Vázquez for sharing her passion for philosophy with me, and for the privilege of being her colleague. This thesis would not have been possible without her incessant support, encouragement, and incredible feedback. Thank you for believing in me.

Thanks to Ígor Clavel Herrera for all the love and support, and for always being there. Thanks to Tadzio Clavel Herrera for all his love and encouragement, and for cheering for me.

Thanks to César Palacios-González for the support, guidance, and invaluable advice.

Thanks to my family, *los Vázquez*, *los Clavel*, *los Hernández Villarreal*, for caring for me. And to little Leonora, Octavio, José Pablo, Santi, Diego, and Oli for sharing with me all their joy and curiosity.

I also want to thank my incredible friends, who keep cheering for me, Fabiola Vallin, Paulina Guendulain, Sebastián Perlín, Guillermo Ruz, Anna Domínguez, Leonardo García, Maricarmen Garcés.

Thanks to the St Andrews and Stirling Philosophy Graduate Programme, to the Stirling Philosophy department, and to the St Andrews Philosophy department. This incredibly lively community has enriched and given shape to my research. Thanks to Dr Derek Ball, Dr Simon Prosser, and Dr

Colin Johnston for their invaluable feedback on my research. Thanks to Dr Philip Ebert for his incredible support. Thanks to Prof Katherine Hawley, Dr Alison Kerr, Prof Sarah Broadie, Dr Justin Snedegar, Dr Alex Douglas, Dr Walter Pedriali, Prof James Harris, Lucie Randal, Audrey Horsburgh, Virginia Respinger, and Claire Exley for their support, particularly, in projects related to our Minorities and Philosophy chapter.

I also want to thank Dr Ezequiel Di Paolo and the IAS-Research Centre for Life, Mind and Society, of the Universidad del País Vasco, for allowing me to spend a couple of weeks with them and benefiting from their incredible feedback.

Thanks to all my philosophy friends: Paolo Savino, Lisa Bâstian, Quan Nguyen, Clotilde Torregrossa, Eline Gerritsen, Camille Fouché, Deb Marber, Lorenzo Lazzarini, Ashley Watkins, Saranga Sudarshan, Miguel Egler, Katharina Bernhard, Ethan Landes, Sophia Rommel, Jakob Hinze, Savvas Ioannou, Amr Salih, Paul Conlan, Indrek Lobus, Moritz Baron, Ásgeir Matthiason, Deryn Thomas, Matt Clark, Lara Jost, Emilia Wilson, Nick Allen, James Skinner, Lixiao Lin, Joe Bowen, and Ravi Thakral. Thanks to Xintong Wei for the friendship, the incredible conversations, and for her support. It was great sharing this experience with her.

Finally, and most importantly, I want to thank Rosa Adriana Vázquez Gómez, my mother, for her unending love and encouragement. For pushing me to be better, for being supportive, and for always listening with joy when I talk about philosophy. Thanks for always believing in me, for having an understanding ear and the kindest and wisest words. Without her example and guidance none of this would be possible.

Contents

Abstract IV

Acknowledgements VII

Introduction 1

I. Sensorimotor Enactivism 7

Chapter 1. A theory of perceptual experience 8

1.1 An alternative account of perceptual experience 9

1.2 Perception: an active and knowledgeable phenomenon 16

Chapter 2. Externalism and Representations in Sensorimotor Enactivism 32

2.1 Representation-friendly SMEn 32

2.2 Explanatory externalism 44

II. The theory of Sensorimotor Contingencies, embodiment and some concerns about chauvinism 53

Chapter 3. The theory of Sensorimotor Contingencies 54

3.1 Accounting for the character of experience 55

3.2 The theory of Sensorimotor Contingencies 65

Chapter 4. Chauvinism and embodiment 75

4.1 Content chauvinism 77

4.2 Sense modality chauvinism 85

III. The Predictive Approach to Sensorimotor Enactivism 98

Chapter 5. Predictive Processing and the Free Energy Approach 99

5.1 The predictive framework: Predictive Processing and the Free Energy Approach 100

5.2 Representations in the predictive framework 111

Chapter 6. Predictive Sensorimotor Enactivism 123

6.1 Neuro-talk in Sensorimotor Enactivism 123

6.2 The predictive approach to Sensorimotor Enactivism 125

6.3 Addressing the concerns about the compatibility of Sensorimotor Enactivism and the predictive framework 132

6.4 An overview of the predictive approach to Sensorimotor Enactivism 143

IV. Predictive Sensorimotor Enactivism and Situated Embodiment 146

Chapter 7. Situated embodiment 147

7.1 From implementational embodiment to vital embodiment 148

7.2 Situated embodiment 158

7.3 The biosocial body 163

Chapter 8. Situating perceptual experience in the predictive approach to Sensorimotor Enactivism 169

8.1 Gender and the situated body 170

8.2 The question about the social character of perception 178

8.3 Situating perceptual experience 183

Conclusion 190

Bibliography 199

Introduction

What could it mean to say that perception is embodied? On the one hand, there is something obvious and trivial in that claim. Consider, for instance, Rafael Núñez's (1999) conception of weak embodiment, according to which a cognitive process, broadly construed, is embodied if it is said to be directly supported by biological structures. This is, of course, the case with perception, since it is a process in which we obtain information from the world by means of our senses. But, is this all there is to claiming that perception is embodied? Sensorimotor Enactivism (SMEn) claims something stronger. For instance, when discussing the idea that the body is always present in the background of our experience, Alva Noë asks us to:

> [c]onsider that organic, somatic field of sensations that forms the ever-present background to our lived achievement of the environment's presence. I mean the tightness in the shoulders or the thrumming in the ears, the aches, tingles, twinges, shivers, vibrations, trembles, chills, tickles, flashes, strains and tautnesses that are, in a way, never absent. I mean that felt constancy of the earth's gravity, the sun's warmth, the night's chill. I mean the felt stretching of the muscles as we rotate the eyes in their orbits (Noë, 2012, p. 12).

Although there is more to unpack here, I want to note that, from this perspective, the conception of the body seems to be richer than just the involvement of the senses conceived as bodily channels of information. For SMEn, perception is a process attributed to a whole embodied agent. Accordingly, the body plays a key role in fixing the character of perception. Perceptual experience is taken to consist in a deployment of embodied skills that is marked by different aspects of the body. SMEn's conception of the embodied character of perceptual experience constitutes the guiding thread of this thesis.

Before introducing the main research problems examined in the thesis, it is worth saying more about the theoretical space where SMEn is located. SMEn belongs to the wave of 4EA (embodied, embedded, extended, enactive, and affective)[1] cognition that has populated the philosophy of

[1] Although not included in the four Es that form the acronym, ecological psychology—the research framework developed by (Gibson, 2015)—is often considered within the 4EA cognition wave. In addition, some proposals within the predictive processing framework have been shaped in a way that places predictive processing in the vicinity of 4EA approaches of cognition. See e.g. (Clark, 2016).

cognitive science and the philosophy of mind since the early 1990s.[2] These approaches are, in part, grouped together in virtue of their rejection of, or at least a strongly revisionary attitude towards, certain assumptions within more traditional approaches in cognitive science.[3] For instance, Albert Newen et al. characterize 4EA approaches as "opposed to the internalist, brain-centered views" that pertain to classical cognitive science (Newen et al., 2018, p. 4). In contrast, for these approaches, cognition is shaped in non-trivial ways by the body and elements external to it. Let me, very briefly, outline these theses:

- The embedded approach to cognition emphasizes the causal relevance of elements external to the cognitive system in the system's problem-solving routines (e.g., (Agre & Chapman, 1987)).
- The thesis of the extended mind argues in favour of the possibility that, in at least some of these problem-solving routines, the material base of the cognitive system itself goes beyond the brain and the skin. In contrast to the embedded approach, external elements are constitutive to the cognitive process in question, rather than just causally relevant (e.g., (Clark & Chalmers, 1998; Wheeler, 2010a)).
- The thesis of the embodied mind holds that cognition, broadly construed, is shaped and constrained by facts about the cognitive system's body or material basis (Ward & Stapleton, 2012, p. 98). Some proponents of this thesis aim at showing that, in at least some cases of problem-solving tasks, the cognitive system relies on internal representations whose content is agent-centred and context-sensitive. Other proponents aim at showing that there are sensorimotor strategies available to tackle the relevant cognitive task. See e.g. (Gallagher, 2005; Varela et al., 1991).
- Enactivism has been discussed hand in hand with the thesis of the embodied mind. It emphasizes that cognition depends on the interactions of an embodied organism. I say more

[2] Albert Newen et al. (2018), for instance, mark as the beginning of this new wave of cognitive science the work of Flor & Hutchins (1991) and of Varela et al. (1991).

[3] 4EA approaches reject certain core assumptions within the classical cognitivist and connectionist frameworks, which can be characterized by their commitment to at least two theses: the computational view of the mind, roughly the idea that cognition consists in information processing; and the representational theory of the mind, roughly, the idea that this processing is carried out by inner representational states. See, for discussion, (Wheeler, 2005, Chapter 1).

about this approach in the next paragraphs, since this is the framework with which SMEn is mostly associated (e.g. (E. A. Di Paolo, 2005; Hurley, 1998; Thompson, 2007)).

- Finally, the thesis of affective cognition claims that cognition is partly shaped by the value assigned by an organism to its surroundings (e.g. (Colombetti, 2013)).

Even when there are many disagreements between the proponents of these theses, they share a few common ideas. Michael Wheeler summarizes the following shared claims: i. these approaches prioritize occurrences of intelligent behaviour that are deployed in the system's real time interaction with the environment; ii. this behaviour is generated "through complex causal interactions" between brain, body, and world; iii. these approaches emphasize that, in relation to organic cognitive systems, the fact that these are biological systems matters for our explanations; and iv. these approaches find a natural ally in the dynamical systems perspective since this approach provides a way of modelling the complex interactions that generate intelligent behaviour (Wheeler, 2005, pp. 11–14).[4]

Within these approaches, SMEn belongs to the enactivist framework, which holds the view that cognition is brought about and enacted by the interactions of an embodied agent. Nevertheless, within the enactivist framework, there are three different positions: radical enactivism, autopoietic enactivism, and sensorimotor enactivism. Radical enactivism aims at showing that there are basic forms of cognition that are not supported by contentful internal states. In this way, radical enactivism denies that representations are necessary for cognition (Hutto & Myin, 2012, 2017).

Autopoietic enactivism, in turn, builds on the view advanced by Varela et al. (1991) to claim that the cognitive agent is a biological system that brings forth its own boundaries in its interactions with its surroundings. Neither the system, nor the world are specified or given prior to their interactions. Instead, both are enacted by the system's autonomous agency (Thompson, 2005, p. 241). The system enacts its own boundaries and conditions of constitution, as well as its domain of concern. Cognition arises from these interactions. A biological system's intelligent behaviour

[4] It is worth distinguishing between the dynamical systems perspective and the dynamical hypothesis. The former is an approach that provides a mathematical model of the way complex systems change over time. These models are expressed by a set of differential equations and describe the states that can be visited by systems over time. See e.g., (Beer, 2000). The dynamical hypothesis, in turn, refers to the idea that the mind is best understood as a complex system rather than as an information processor. See e.g. (Chemero, 2000).

is one of the ways in which the system attunes to the environment (Ward & Stapleton, 2012, p. 91).

Autopoietic enactivism emphasizes the continuity between life and mind, and takes cognition to exhibit, in a particular way, the features that pertain to the adaptive behaviour of the biological system. In contrast, SMEn focuses on perception, and holds that perceptual experience arises from and is constituted by the skilful interactions of the perceiver with the world. From this approach, perceptual experience depends upon the agent's grasp of the sensorimotor regularities that govern these interactions.[5]

Reviewing this context is crucial because, while SMEn is typically associated with the approaches discussed above, its exact place within this literature is often questioned (Buhrmann et al., 2013). As we will see, SMEn takes a critical stance towards traditional approaches when it comes to the nature of the mind, perception, and cognition. However, its position towards representations has been ambiguous. In virtue of this ambiguity, SMEn might be said to stand at the border between orthodox and unorthodox approaches to cognition. When interpreted as a representational approach, SMEn could be said to enrich the orthodoxy in claiming that perception is a process in which we build a representation of the world (Vernazzani, 2019, pp. 4528–4529).

In order to answer the question about SMEn's position with respect to the 4EA literature, though, another matter needs to be addressed related to its causal story and the mechanisms that underlie perception. While SMEn provides an outline in their theory of the sensorimotor regularities that govern a person's or organism's sensorimotor interactions, the causal story remains ambiguous (Buhrmann et al., 2013). By saying more about SMEn's causal story, it is possible to fill in the gaps to understand where the view stands with respect to the other frameworks discussed. Making sense, thus, of SMEn's position towards the nature and constitution of the cognitive system.

But how does this relate to the question regarding the body that was posed at the beginning of the introduction? According to SMEn, perception is constituted by embodied skills. Sensorimotor interactions are governed by perceivers' grasp of the sensory information that is made available

[5] In the vicinity of SMEn is the sensorimotor approach developed by (Ward et al., 2011). According to this approach, the content of perceptual experience is determined by perceivers' grasp of the relation between the possibilities of action offered by the objects of our perception—affordances—and the invariant features of these objects (Ward et al., 2011, p. 379).

given certain details of their bodies. SMEn holds that the vehicles of perceptual experience span over brain, body, and world. The causal story advanced by the proponents of SMEn is, thus, one that accounts for the interactions between the perceiver and the world as they are constitutive of perceptual experience. Moreover, this causal story includes details about the perceiver's body. In line with this idea, Noë claims that our causal story might need to mention "hands and eyes" (Noë, 2004, p. 25).

As I mentioned, the guiding thread of the thesis is SMEn's conception of embodiment. I address the different aspects of the embodied character of perceptual experience. The thesis is divided in four parts, each addressing different challenges that arise from SMEn's conception of embodiment. Part I presents SMEn's main features, and thus sets the stage for the problems that are discussed in the rest of the thesis. Chapter 1 examines SMEn's claims that perception is active and knowledgeable, and Chapter 2 examines its position in regard to representations and explanatory externalism, the thesis that the vehicles of perceptual experience must be specified in virtue of the elements that are doing the explanatory work.

In Part II, I discuss the dependence of perception on the body and the concern that it is chauvinistic in that it claims that only creatures with bodies like ours can have experiences like ours. Against this claim, I argue that even if SMEn's embodiment claim involves a commitment to an irreducibly perspectival aspect of perceptual experience, Sensorimotor Contingencies (SMCs) are structural descriptions of perceptual systems that already involve information about objects' invariant features. Chapter 3 examines the theory of SMCs, the systematic patterns of co-variations between sensory information and possible interactions between the perceiver and the object(s) of perception. In light of SMCs' dependence on the body, Chapter 4 responds to the concerns that SMEn is chauvinistic.

Part III argues in favour of a predictive approach to SMEn to address the concern that, while it claims that perception is embodied, it leaves out one crucial element of the material basis of perceptual experience: the brain. Chapter 5 defends an externalist version of the Free Energy Approach (FEA), according to which the process of surprise minimization spans over brain, body, and world. Although the predictive framework is often taken to involve a representational profile, Chapter 6 argues in favour of the compatibility between SMEn and the FEA. I show, firstly, that the representations advanced by the FEA comply with SMEn's constraints discussed in Chapter 2. Secondly, I show that the FEA is compatible too with SMEn's explanatory externalism.

In Part IV, I address a concern that is inspired and based on the phenomenological analysis of the body, according to which it is marked by someone's social and historical circumstances. I argue that the predictive version of SMEn provides a framework to make sense of how the situated body shapes perceptual experience. Chapter 7 examines the conception of embodiment at play in the FEA and shows that it can be interpreted from the phenomenological conception of situated embodiment. Chapter 8 argues that because certain aspects of our situation contribute to the way we participate in social practices and to the expectations others have from us, our routines of surprise minimization depend, in an important way, on aspects of our situation.

I. Sensorimotor Enactivism

Chapter 1. A theory of perceptual experience

In the XIX century, Louis-Jacques-Mandé Daguerre, based on previous work with Joseph-Nicéphore Niépce, invented the daguerreotype, a photographic process that involved chemically treating a polished silver plate to make it photosensitive. This treated plate was then placed in the camera, where it was exposed to light. Daguerreotypes are unlike the pictures that result from other photographic processes. Given that the image is marked in a mirror-like surface, observing it can be difficult. To get a glimpse of all the details of the image, one has to move it around, place it in different lighting conditions and in front of different backgrounds and surfaces, one has to adjust the angle, and even adjust one's position. Daguerreotypes require an active participation of the observer to discover the image. The observer has to explore the object to see the image in all its richness. When looking at a daguerreotype, one does not have the image all at once. Instead, one has to look for it and for the details that appeared once but are now out of sight.

The experience when observing daguerreotypes might seem atypical. We might think that, typically, when we observe a scene, a picture, or an object, there is not much exploration happening. Alternatively, we might think that perhaps *some* exploration is taking place but only in cases in which, for instance, a scene outstrips the scope of our visual field. Still, unless something is disrupting our view or we are under unfavourable conditions (e.g. it happens to be an extremely foggy day), we tend to think of visual perceptual experience as one in which we get to see all there is to see at once.

Sensorimotor Enactivism (SMEn)[6] offers a view that puts pressure on this model not only of visual perceptual experience but of perceptual experience in general.[7] SMEn takes perceptual experience

---

[6] I refer to the view as it was firstly articulated by Kevin O'Regan & Alva Noë (2001), based on previous work by Susan Hurley (1998). Further developments of the view can be found in e.g. Noë (2004; 2012) and O'Regan (2011). The work on SMEn has taken two different avenues. On the one hand, e.g. Ezequiel Di Paolo et al. (2017) have developed a version of the sensorimotor approach in line with autopoietic enactivism. On the other hand, e.g. Anil Seth (Seth, 2014) has developed a version of the approach in line with the more recent predictive approaches to the mind. In Part III, I discuss the predictive approach to SMEn, since the view I defend in this thesis goes in line with Seth's proposal.

[7] Throughout this thesis, I mostly refer to the case of vision. SMEn has been applied, however, to other sense modalities. Moreover, its proponents turn to touch as a paradigm for perceptual experience. As Becky Millar (2020, p. 4) notes, Noë, for instance, briefly applies the view to audition (Noë, 2004, pp. 160–161). For a sensorimotor approach to the chemical senses, see Millar (2020). For a sensorimotor approach to audition, see Cooke & Myin (2011).

to be closer to the case of daguerreotypes, a case in which one has to actively search for the image. Moreover, under the proposal of SMEn, vision is not the paradigmatic sense modality. Instead, they claim that perception should be modelled having touch as the paradigmatic sense modality. Perception is, for SMEn, akin to an exploratory activity in which we probe the environment as it offers possibilities for action.[8] For SMEn, perception is something we *do* mediated by something we know (O'Regan & Noë, 2001, p. 970).

In this chapter, I begin by reviewing two views rejected by SMEn. First, the sandwich view of the mind, a view about the structure of the mind, and the relation between perception, cognition, and action. The second, the snapshot conception of perception, is concerned with perception specifically, and it is a view of the visual field that has served as a model of perception. I then discuss one of the features of perception SMEn is interested in accounting for, namely, the phenomenon of perceptual presence. In the second section of this chapter, I turn to the main claims of SMEn, namely the claim that perception is active and the claim that it involves knowledge.

1.1 An alternative account of perceptual experience

1.1.1 Leaving some assumptions behind

SMEn starts off from the rejection of certain core assumptions about what the mind is. SMEn is advanced as an alternative approach that might do better in explaining certain mental phenomena, in this case perception. To better understand SMEn, I briefly discuss two views from which SMEn distances itself: the sandwich view of the mind, and the snapshot conception of the visual field. While the latter refers specifically to a view that concerns perception, the former refers to a position regarding the structure of the mind and what constitutes the core of cognition.

[8] Throughout this and the next chapter, I use perceptual experience and perception interchangeably. In chapter 4, I address briefly the question of whether there is unconscious perception and the position of SMEn in regard to this question. It is worth mentioning here that O'Regan & Noë (2001), for instance, are willing to bite the bullet and claim that unconscious sensorimotor processing does not amount to perception. For the purposes of chapters 1 and 2, the distinction will not be relevant.

The sandwich view of the mind

Let us start with the sandwich view of the mind. SMEn belongs to the views that resist what Hurley (2001) called the sandwich view of cognition (broadly construed). These are views in which action and perception are peripheral, and cognition is central. Cognition is, as it were, the "filling" of the sandwich (Hurley, 2001). The sandwich views of cognition are views according to which the mind can be characterized by (some of) the following aspects:

1. Perception, cognition, and action are three separate mental phenomena.

2. While perception and action are peripheral, cognition interfaces between them. Moreover, cognition is the core of the mind. Which means that all of the most relevant aspects of our mental lives, for instance, our rationality, our intelligent interactions with the environment, our capacity to interact with each other, are all explained mainly by this core. In turn, perception provides information about the outside world that will inform these intelligent responses, while our actions are just the result of our intelligent deliberation.

3. Finally, mental phenomena are to be explained classically, i.e. "involving symbols and recombinant syntactic structure" (Hurley, 2001, p. 3).

At the basis of this view of cognition rest a couple of assumptions that need to be challenged. The first of these is the assumption that (a) the relevant causal flow to explain cognition is linear, i.e. that there is an ordered relation between firstly perception, then cognition and then action, instead of dynamic feedback loops between sensorimotor processes and cognitive processes. Secondly, it is assumed that (b) the relations between action and perception are instrumental, i.e. that perception provides information for e.g. deliberation, and that deliberation, in turn, serves for action.

The sandwich view of the mind can be challenged by rejecting either or both of these assumptions. Whether this is the right conception of the mind is, of course, a matter that can only be decided empirically (Hurley, 2001). However, according to Hurley, there are methodological reasons to reject these assumptions. Rejecting them is conceptually more liberal. Empirical research benefits from having more conceptual resources at hand, since it offers different possibilities to investigate.

SMEn is among the views that seek to put pressure on these assumptions by showing the deep interplay of action and perception. As mentioned before, perception is for SMEn, after all, something we do. Moreover, the view of perception defended by SMEn is one according to which perception involves knowledge. As I explain throughout this section, for SMEn, perception is an active and knowledgeable phenomenon.

The snapshot conception of perception

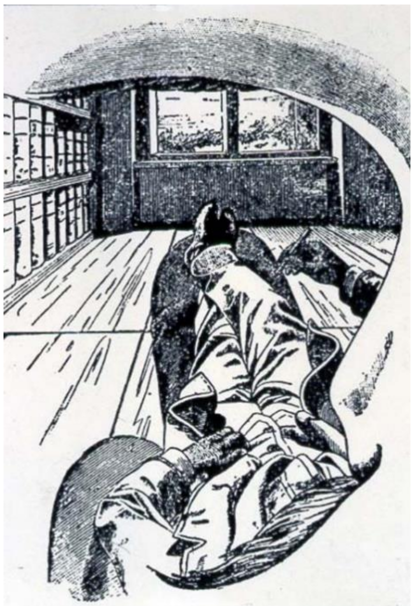

Figure 1. Ernst Mach depiction of the visual field ((Mach, 1959) as referred by (Noë, 2004, p. 35)).

The second view that SMEn aims at leaving behind is the snapshot conception of perception, which is directly related to perception, rather than to a view of the mind. For instance, according to Noë, the view that permeates our understanding of perception is one according to which, e.g. to see is to "undergo snapshot-like experiences of the scene" (Noë, 2004, p. 35). Noë illustrates this idea with Ernst Mach's (1959) illustration of the visual field (see figure 1). This understanding of the visual field invites a view of perception according to which what we get in perceptual experience is a detailed depiction of the object of perception or scene perceived at once. In other words, by means of perception, we build a picture of what the world is like. When conceiving perceptual experience in this way, taking perception to be the process by which we build representations of the objects perceived and of their properties comes straightforwardly.

The assumptions underlying the snapshot conception of perception are that: (a) perceptual experience is detailed, picture-like; (b) perception is a process by means of which an internal representation of the scene perceived is created; and (c) the presence of a picture-like mechanism (e.g. the cortical map in the case of vision) leads to picture-like content (O'Regan, 2011, p. 63).

Importantly, what is rejected by SMEn is the idea that this should be the model to pursue research on perception. When working under the assumption that this is what it is like to see, in this case, how a room appears to our eyes while we lie down—e.g. as detailed, uniform, and focused—we might risk leaving an important aspect of perceptual experience out, namely, the fact that there is evidence that suggests that we do not have all the details of a scene perceived at hand all at once. I return to this point in the following section. What I want to note here is that the rejection of the snapshot conception is the rejection of a model of perceptual experience that might be misleading.

1.1.2 Addressing the sense of presence

The limitations of the visual system

When thinking of perceptual experience—particularly of vision—, there are two aspects that need to be reconciled. While our sensory systems process more limited information than expected, our perceptual experience does not reflect these limitations, or at least not *as* limitations. In other words, even when our visual system is limited, our experience is of a detailed scene.[9] This, of course, must be addressed by a theory of perception, namely that our experience is not one that we expect would result from a limited system, it is not 'gappy', so to speak. Instead, when seeing a scene, the scene indeed appears rich and detailed.

Consider, for instance, O'Regan's (1992) discussion of the limitations of the visual system. He notes, for instance, that before reaching "photosensitive rods and cones", light must traverse a web of blood vessels and neurons that should obscure the photosensitive layer of the eye (O'Regan, 1992, p. 462). The way the axons of these neurons and vessels are placed leave an area with no photosensitive cells, namely the blind spot. In addition to this, the way cones are spread out in the retina is not uniform; in the outer regions of the retina, the distance between cones increases. As there is no uniformity in the distribution of cones across the retina, O'Regan also notes that colour vision is not uniform. Finally, he notes that saccadic movements displace, on some occasions, the retinal image. For O'Regan, the real mystery that should concern all vision scientists is "why we can see so well with such a terrible visual apparatus" (O'Regan, 1992, p. 463). The limitations of the visual system have also been noted by Zenon Pylyshyn (2000) who focuses, instead, on the little information that is encoded as the eyes probe a scene.

Change blindness

In line with this, one important phenomenon concerns cases in which we fail to note otherwise salient features of certain scenes (e.g., a friend waving at us), or we fail to notice dramatic and radical changes in the scene that should be evident (Simons & Chabris, 1999). This is the phenomenon of change blindness, a phenomenon in which a drastic change goes unnoticed because it occurs at the same time as some other disruption. The disruption can be something as

[9] In Part IV, I return to this in light of Anne Jacobson's (2012) paper where she discusses the social aspect of vision.

simple as an eye movement, a flicker or a blink in the scene, or the scene being briefly covered; and the change can go unnoticed even when the disruption does not cover the change completely or at all (O'Regan et al., 1999, p. 34). An example of this is the experiment developed by Daniel Simons & Christopher Chabris (1999) in which a group of individuals are shown a video of two teams passing a ball in which a highly salient and unexpected event occurs (either a woman with an umbrella or someone wearing a gorilla suit crosses the scene). Their task is to follow and make a mental count of the passes of one of the teams. Simons & Chabris found that around half the observers failed to notice the unexpected event.[10]

One way to explain the disparity between the richness of our perceptual experience and the limitations of the visual system can be to posit an inner model or representation of the scene perceived. The reason why the huge gap in the visual field caused by the blind spot is not experienced is that a model or representation of the scene perceived has been built. The idea is, roughly, that the representation fulfils this role of filling in the gap, or of providing the information that corresponds to the area covered by blind spot.[11] Change blindness puts pressure on a representational account: if indeed a representation of the scene perceived is built, why do we fail to detect those dramatic changes? It is important to note that while this puts pressure on representational accounts, it is not sufficient to disprove them. This is precisely the way the problem is posed by Simons & Chabris. They aim at addressing the disparity between "the richness of our experience and the details of our *representation*" (Simons & Chabris, 1999, p. 1059, emphasis added). The phenomenon of change blindness can be addressed by a representational proposal. The proposal would have to explain, though, that even when in perception we built a detailed representation of the scene, some unexpected details are left out. And it is against these proposals that SMEn is advanced: as an alternative, competing account of perceptual experience.

Let us take stock. There are some features that an account of perception should seek to reconcile: our experience of perceiving a detailed and rich scene, with the limitations of our visual system, and that with the evidence that we fail to notice drastic and salient changes and features of a scene. One way out of this problem is to dismiss our experience. In other words, to claim that that we are

[10] Simons & Chabris were specifically interested in the relation between attention and conscious perception, and argued that there is no conscious perception without attention (Simons & Chabris, 1999, p. 1071).
[11] Note that, under certain conditions, we can in fact 'see' or note the visual blind spot. For an exercise to see the blind spot, see e.g. (Lohner, 2020).

being fooled by our experience: the experience of a detailed and rich scene is nothing more than an illusion.[12]

An account of the sense of presence

SMEn provides an alternative position that does not require the positing of representations, and that does not claim that we are fooled by our experience. Instead of the details experienced being available as an inner model of the scene perceived, SMEn claims that those details are *virtually* available. The scene is experienced as detailed because that detail can be made available by probing and exploring the scene. Think of the example of the daguerreotypes. Even when we cannot see the picture all at once, we can inspect the object, interact with it, and position it and ourselves in a way that will allow us to access the photographed image. Of course, in the case of daguerreotypes, the experience of the object is not detailed. Daguerreotypes are, themselves, a case of visual experience in which the object is not perceived with richness and in detail due to features that pertain to the object, not to our visual system. However, it serves to illustrate the idea that the details can be made available if we were to move ourselves and the object around in certain ways. In this way, Noë introduces an interplay between presence and absence that takes place in perceptual experience. The experienced detail is *virtually* present not because of an inner model, but because the details are out there in the world. For Noë, "[t]he world outstrips what we can take in at a glance; but we are not confined to what is available in a glance" (Noë, 2008, p. 660).

There is one last point that requires attention. Even when certain occurrences go unnoticed by us, the scene still appears as rich as always. Moreover, even details to which we pay no or little attention, or that lie beyond our gaze, make an appearance in our perceptual experience. One example that has served to illustrate this is the wall of Marilyns, originally due to Daniel Dennett (1991, p. 354 ss). Imagine that you enter a room in which one wall has been covered by a wallpaper of several identical pictures of Marilyn Monroe's face, in the style of a Warhol artwork. It would be accurate to say that one experiences the wallpaper filled with Marilyns even though—given the limits of the visual system—it is impossible to look at them all at once in one glance and in clear focus (Noë, 2004). Moreover, to experience the wallpaper as filled with Marilyns' faces, it is not

[12] See e.g. Dennett's position, who claims that in our experience there is nothing beyond *judging* that a scene appears in a certain way (Dennett, 1991, p. 362 ss).

necessary to have looked at each section of the wall. Elements that are not currently perceived and that have not been perceived make an appearance in our experience: one does not need to fix one's gaze on each of the Marilyns to experience the wallpaper.

There is a prevalent phenomenon in perceptual experience in which elements that are absent enjoy a certain degree of presence, as in the case of the wall covered in Marilyn's wallpaper. This is a puzzling aspect of our perceptual experience, since even when we are not attending to a feature of a scene or object perceived, or even when an object is partially occluded, this feature appears in our experience as being absent or as being out of sight (Noë, 2004, p. 59). The case that is often used to illustrate this phenomenon is that of a cat sitting motionless behind a fence (Hurley & Noë, 2003; Noë, 2004). Even when we do not see the whole cat, our perceptual experience is that of seeing a whole cat, rather than bits of cat. Or consider the experience we have when seeing, e.g., a tomato: despite there being one sense in which we see only the side of the tomato facing us, Noë points out that the tomato appears in our experience as a three-dimensional, round object. The side of the tomato that is hidden also appears in our experience, even when it does not appear in the same way as the visible side of the tomato, but instead appears as absent (Noë, 2004, p. 60). Noë refers to this as the problem of perceptual presence. The problem arises whenever "[w]e have a sense of presence of that which, strictly speaking, we do not perceive" (Noë, 2004). This problem relates to the phenomena that psychologists refer to as amodal completion.[13] Amodal completion is prevalent not only in visual experience, but in other sense modalities as well. For instance, in the case of touch, we perceive the whole bottle even when our hands are in touch with a limited surface. And in audition we continue to perceive a tune or someone's voice even when interrupted by a loud sound (Nanay, 2018, p. 1).

SMEn is at attempt at doing justice to these interplays of presence and absence in perceptual experience. To account for these, SMEn argues that to perceive an object is to master our interactions with the world perceived. As mentioned before, the details of the perceived scenes, the faces of Marilyn not currently being seen, the back of the tomato or the bottle, are all virtually

[13] A distinction has been drawn between modal and amodal completion. While amodal completion refers to cases in which an occluded object is experienced as present (e.g. the cat behind the fence), modal completion refers to cases in which an object is experienced as occluding the inducers (e.g. the Kanizsa triangle). According to Nanay (2018, p. 2), both are cases of shape completion and should be supported by the same mechanisms. He argues that amodal completion is explained by mental imagery. Note here that I am not following Nanay's proposal for addressing this phenomenon.

available to the perceiver at the turn of the hand or her head. Perceptual presence is explained by the perceiver's mastery of the knowledge of the way these details can be accessed. SMEn takes visual perception as an exploratory activity mediated by knowledge: "vision is a mode of exploration of the world that is mediated by knowledge, on the part of the perceiver, of what we call sensorimotor contingencies" (O'Regan & Noë, 2001, p. 940). The sense of presence of perceptual experience does not come from a detailed picture of a scene, but rather from the fact that the details are there, present not in our mind, but in the world. In that sense, SMEn follows Rodney Brooks' claim that "the world is its own best model" (Brooks, 1991, p. 15). The sense of presence is due to the availability of the world itself, and our capacity to bring about those details in our interactions.

As seen from the discussion, SMEn characterizes perceptual experience as active and skilful, since to perceive something is, according to this view, to master a certain kind of practical knowledge that allows the perceiver to access sensory information. In the following section, I discuss in detail these two aspects—the active and the knowledgeable—of perceptual experience. I begin by discussing the claim that perception is something we do, namely that it is active.

1.2 Perception: an active and knowledgeable phenomenon

1.2.1 Perception is something we do

One of the first features of perception that becomes noticeable in SMEn's view is its active character. As mentioned, perception is defined as something we do. Zoe Drayson (2017), for instance, interprets SMEn's view as one according to which perceivers are necessarily agents.[14] Another way to articulate this claim is to say that, for SMEn, to be a perceiver is necessarily to be a skillful agent, one that interacts with the world in such a way that she can pick specific kinds of sensory information from the world. From this perspective, perception is *constitutively* an action. Drayson's interpretation of SMEn's position is on the right track. However, unless properly understood, this view faces many objections and seems implausible.

[14] Drayson (2017) notes that the claim about perceivers being necessarily skilful agents is a claim about nomological necessity, rather than metaphysical. Perceivers are necessarily skilful agents in virtue of certain details of the actual world.

When discussing the claim that perception is active, it is important to consider that there is no consistency across the literature on what precisely is the reference of the 'active' character of perception. What I mean here is that it is not clear which movements and/or actions are taken to be relevant for perceptual experience. This will become evident in the objections presented below. The active character of perception, at times, seems to refer to any movement either of the perceiver or of the scene or object of perception, including e.g. saccadic eye movements (O'Regan & Noë, 2001). While, in other occasions, the claim refers instead to doings—intended or not—of the perceiver (Noë, 2004). Moreover, these doings might not only be of, or might not directly relate to, the relevant sensory system, but they might be movements of the whole agent. Furthermore, as I discuss in the following section, there is one last twist in SMEn's view. It is not only that perception is active, but that it is skillful: i.e. it involves the *capacity* to act with sensitivity to the way sensory stimuli changes after an interaction. And, once again, this might mean that this capacity relates specifically to the relevant sensory system, to the whole agent, or to both. For instance, the view might be that what is relevant is the capacity to move ones' eyes in a certain fashion or at a certain speed, or the capacity to position oneself in a certain way.

Moreover, the claim that perception is active has received a lot of critical attention. Particularly when claiming that perceivers are *necessarily* agents. One immediate problem that arises relates to cases of perception that involve no action initiated by the perceiver. Although these cases do not lead to the rejection of the claim that perception is kinetic—in the sense that it involves movement—they raise questions about perception being something done *by the perceiver*. Perhaps perception is kinetic, but it does not require self-initiated action (regardless of whether the action is intended or not) and, in consequence, it seems unjustified to claim that perceivers are necessarily agents.

Another possibility is to read SMEn's claim, instead, as stating that perception is kinetic, because it involves movements of the perceiver or of the object perceived. However, this claim is equally problematic. Consider cases of perception in which no movement was involved, cases of what Noë calls inert perceivers (Noë, 2004, p. 12). Even when we might want to claim that the former are uncommon—e.g. saccadic movements are central to visual perception, these cases are conceivable.

So, given that inert perception is in principle possible, reading SMEn's claim as perception being necessarily kinetic is implausible.[15]

Moreover, the relation between action and perception becomes even more troublesome when we think of phenomena that are phenomenally similar to perception, for instance hallucinations or some cases of imagination (e.g. sensory imagery). If perceptual experience is due to the active engagement of the perceiver with her surroundings, some work is needed to account for these cases.

As can be seen, there are many things happening here, from a view that links perception and movement, to a view that instead links it to agency, to a view that claims that this link is necessary. In the following paragraphs, I consider three possible interpretations of SMEn's view that perception is active.

Movement is a contextual factor of perception

To simplify the discussion, let me formulate the position in question as one according to which perception is active just in the sense that it is kinetic. This is just the claim that perception involves movement. I begin from this claim because it will be easier to see how weak or how strong SMEn's position about the active character of perception can be. To do so, let us draw on a distinction introduced by Hanne De Jaegher et al. (2010) and applied by Fernando Bermejo et al. (2020) to this matter. The terminology was introduced by De Jaegher et al. with the aim of clarifying the different causal roles social interactions might be fulfilling in cognition.[16] De Jaegher et al. distinguish between *contextual factors*, *enabling conditions,* and *constitutive elements*. Each of these plays a different role in the explanation of the relevant phenomena. The following is an adaptation of their definition to the role of action in perception. Given any occurrence of perceptual experience P and any movement M,

(i) "[M] is a *contextual factor* if variations in [M] produce variations in [P],

(ii) [M] is an *enabling condition* if the absence of [M] prevents [P] from occurring, and

[15] See (Noë, 2010) for a discussion on the active character of perception.
[16] In chapter 7, the distinction will be again relevant in relation to the social aspect of perception. In this section, I am only concerned with the relation between action and perception.

(iii) [M] is a *constitutive element* if [M] is part of the processes that produce [P].” (De Jaegher et al., 2010, p. 443)

Claim (i) merely states that M should be considered a *contextual factor* if there is co-variation between movement and perceptual experience. This would be the weakest version of SMEn’s claim about the kinetic character of perception. According to it, a difference in movement corresponds to a difference in perceptual experience. This is compatible with the claim that perceptual experience can occur without movement. Bermejo et al. (Bermejo et al., 2020) argue, in connection to this claim, that in this case movement would alter or change “the manifestation of the phenomenon, but not whether it is manifested or not” (Bermejo et al., 2020). And given that it alters the manifestation of the phenomenon, it should be part of the explanation. In this weakest version of the claim, SMEn would be committed to a claim according to which perception is kinetic insofar as movements (either of the perceiver or of the object perceived) are relevant to the explanation of perceptual experience. However, movements are not necessary for perceptual experience. If SMEn was committed to this claim, the worry about inert perceivers would not arise.[17] Nonetheless, it seems that SMEn is committed to a stronger version of the claim about the character of perception. In the next section, I discuss this more robust position.

Movement is an enabling condition of perception

Consider O’Regan & Noë’s (2001) discussion of the case of inert perceivers. They accept that perceptual experience is possible without movement, although they claim that these cases are exceptions. Furthermore, they refer to a series of experiments by Tatjana Nazir & O’Regan (1990) to argue that, even when perceptual experience occurs without movement, performance in certain perceptual tasks diminishes without movement.

In these experiments, Nazir & O’Regan studied the performance of subjects in tasks that involved the identification of a previously unknown symbol. The symbols in the study were patterns of dots and squares that resemble Chinese characters. The experiments tested whether subjects were able to recognize a target symbol previously projected to a fixed retinal location when it was projected to a new retinal location. One of the conditions of the study was that eye movement was controlled.

[17] Some proposals within the framework of ecological psychology seem to defend versions of this claim. See, for discussion, (Bermejo et al., 2020).

Subjects were prevented from moving to see the target. O’Regan & Noë describe this condition as follows: “the eyes were fixated at the middle of the symbol, but as soon as the eyes moved, the symbol would disappear” (O’Regan & Noë, 2001, p. 948). The study found that subjects’ performance in recognising the symbol decreased under these conditions. According to Nazir & O’Regan, one possible explanation for the decreased performance during this task could be the following: the recognition of an image depends on the association of different patterns of retinal activations with the same stimulus (Nazir & O’Regan, 1990, pp. 97–98). This pattern changes depending on the position of the target. Nazir & O’Regan’s series of studies point to the thesis that the recognition of an image improves with the accumulation of experience that comes from learnt patterns that correspond to different retinal positions.

Nazir & O’Regan’s study allows me to jump to the second claim, namely (ii) the claim that movements are *enabling* conditions of perceptual experience. According to De Jaegher et al. (2010), enabling conditions are necessary for the phenomenon to occur, regardless of whether these are contemporaneous or not. This claim goes beyond merely claiming that movement is relevant for an explanation of perceptual experience because the latter co-varies with movement. The claim that movement is an enabling condition of perceptual experience can be supported by Nazir & O’Regan’s experiments. While perceptual experience without movement is possible, movement seems to be relevant for the performance in some cognitive tasks. After all, subjects’ poor performance in these experiments can be explained by referring to the lack of movement that prevented subjects from learning the relevant patterns of retinal activation.

The well-known kitten experiment by Richard Held & Alan Hein (1963) becomes relevant in connection with the claim that movement is an enabling condition of perceptual experience. The authors aimed at investigating whether impeding self-initiated movement (or self-locomotion) hinders the development of visually guided behavior. In the experiment, a pair of kittens were harnessed to a gondola while being exposed to certain visual stimuli. One of them—the active kitten—was free to move around, while the second—the passive kitten—was restrained. Although the locomotion of the passive kitten was restricted, it was able to execute other self-initiated movements, e.g. it was able to turn its head. While moving, the active kitten pulled the second kitten who, given the design of the gondola, was exposed to the same visual stimuli as the active kitten. The kittens were exposed to these conditions a few hours a day, for several weeks.

After ensuring that active kittens were able to successfully perform certain tasks of visually guided behavior, e.g. that they responded adequately in the paw-placement test in which they test their reactions when carried towards a horizontal surface, kittens were tested on the avoidance of a modified visual cliff. This test consists of a modified platform that drops three inches to a plate of glass. The plate has two sides, one deep and one superficial. On the deep side, a patterned surface located 30 inches below the plate is visible. On the superficial side, the same patterned surface is placed immediately below (Held & Hein, 1963, p. 874). The experiment showed that, after being exposed to these conditions, passive kittens failed to discriminate between the deep and the narrow sides of the plate. The authors took their experiment to show that self-locomotion was necessary for the typical development of visually guided behavior.

The kitten experiment was among several studies performed by Held and other researchers to support the view that active interactions are required for the development of certain perceptual skills. Unlike the claim I have discussed so far about the relevance of movement for a causal explanation, Held was committed to a stronger view, namely the view that "self-produced movements are necessary for developing a perceptual ability" (Bermejo et al., 2020). However, these results have been contested. Bermejo et al. mention, for instance, that in an attempt at replicating Held & Hein's results, Walk et al. (1988) attributed the poor performance of passive kittens not to a lack of self-locomotion, but to a lack of attention. With the support of other experiments, they argued that the conditions of the passive kittens led to a lack of attention to their environment which, in turn, led to the inability to discriminate depth (Bermejo et al., 2020; Froese & Ortiz-Garin, 2020). So, Held & Hein's results are consistent with an alternative hypothesis that would not lend support to the claim that self-locomotion is an enabling condition of, in this case, depth perception.

In contrast with this, a recent study by Alexander Attinger et al. (2017) advanced similar results as those obtained by Held & Hein. In this study, a pair of mice were placed in a "virtual reality setup" that resembled Held & Hein's gondola (Froese & Ortiz-Garin, 2020, p. 3). The setup consisted of a display and a trackball on top of which mice were placed. The display would change in accordance with the movements of an active mouse, while the locomotion of the other passive mouse was restrained. Both mice were exposed to identical visual stimuli. However, only one of them was able to influence what appeared on the display. On the basis of neuroimaging, the study showed a difference in functional development between the two mice.

Finally, in a recent paper by Tom Froese & Guillermo Ortiz-Garin (2020), the authors designed an experiment using the enactive torch.[18] The authors describe this device as "a hand-held sensory substitution device (…) which translates infrared-based measures of distance to nearby objects into intensity of vibrotactile feedback in the user's hand" (Froese & Ortiz-Garin, 2020, p. 1).[19] As in the previous cases, this experiment was related specifically to self-locomotion and its relevance in subjects' performance in specific tasks. In Froese & Ortiz-Garin's experiment, a pair of subjects were asked to move the enactive torch from side to side. The torch measured the width of the relevant object and then translated this information into vibrotactile stimuli, which are received by the subjects. While one of the subjects was able to move the torch, the other remained passive while her hand was moved, mirroring the stimuli received by the active subject. In this case, the authors tested whether passive and active subjects performed differently when identifying which of the objects was wider. The results of the experiment showed no statistically significant difference between the performance of each group of subjects.

As can be seen from the aforementioned studies, the results of experiments that aim at showing the relevance of self-locomotion and movement to subjects' performance in different perceptual tasks are mixed. Moreover, it seems that the relevance of locomotion or self-locomotion for the development of a perceptual ability and for the performance in a perceptual task cannot be generalized. One would expect a difference in relevance depending on whether it relates to visually guided behaviors or visual recognition tasks. While a certain kind of movement might be an enabling condition for visually guided behavior, in that it is a requirement of typical development, this might not be the case when it comes to another capacity.

[18] Froese & Ortiz-Garin were partly interested in designing an experiment similar to Held & Hein's that involved specifically human subjects, since they question whether the results obtained in Held & Hein's and Attinger et al.'s experiments are transferable to human perceptual experience.

[19] Sensory substitution is the phenomenon in which information coming from one sense modality is translated to another sense modality. This phenomenon will become relevant in section 3.2.

Perception is constitutively active

Let us take stock here. In the previous section, I examined the plausibility of the claim that movement, in whichever terms it is constructed, namely as locomotion, self-locomotion or as movement simpliciter, is an enabling condition of perceptual experience. One way to support this claim is to show that there is evidence of diminished performance in relevant perceptual tasks. The claim would be, then, that restrained movement hinders the performance and/or development of subjects in relevant perceptual tasks or of relevant perceptual capacities. As shown by the previous discussion, the evidence for this claim is mixed. On the face of that evidence, one might think that it is not worth considering the last, more robust, claim (iii) regarding perception being *constitutively* active. After all, if there is insufficient evidence to claim that movement is an enabling condition of perception, there seems to be much less evidence to claim that it is constitutive.

Let us examine De Jaegher et al.'s claim regarding constitutive elements. According to them, "[M] is a *constitutive element* if [M] is part of the processes that produce [P]" (De Jaegher et al., 2010, p. 443). They add that constitutive processes or elements are "present at the same time of the phenomenon", since they are among the processes or elements that pertain to the relevant phenomenon (De Jaegher et al., 2010, p. 443). In the case of perceptual experience, we would have to formulate this claim in the following way: insofar as movement is part of the processes or elements that conform perception, whenever there is perceptual experience, there is movement. Of course, this claim immediately runs into the objection that perceptual experience is in principle possible without movement.

It is important to note here that De Jaegher et al. are interested in showing the different *causal roles* a certain element fulfils in the manifestation of a given phenomenon. However, there is another way to read the constitutive claim. Instead of taking it as a claim about the causal role of a given element, we can draw on John McDowell's (1994) distinction between *constitutive* and *enabling* features of a phenomenon. While the latter are features that "play a causal role" in perception, the former are *defining features* of perception: these are just what perception is. De

Jaegher et al.'s distinction concerns the causal explanation of a phenomenon.[20] In that respect, an element might fulfil the roles indicated by the authors. However, when it comes to SMEn's claim that perception is constitutively active, we are not (only) referring to the role movement plays in perceptual experience, instead this is a claim about what perception *is*. Under this reading of the claim, I do not mean to say that movement is a constitutive element of perception. As indicated before, that claim is untenable.

To make sense of the claim that perception involves movement, let us start by considering Wheeler's definition of online intelligence: "[a] creature displays online intelligence just when it produces a suite of fluid and flexible real-time adaptive responses to incoming sensory stimuli" (Wheeler, 2005, p. 12). In the case of perception, however, its online character should be understood in the following terms: as (a) a temporally rich phenomenon, that (b) involves real-time anticipation and adaptation. SMEn's claim is about a process that unfolds over time and that constitutes an interaction between the agent and the world. Even when the claim that perception involves movements—in different ways—is relevant, SMEn's claim is not about movement as such.

Instead, SMEn aims at characterizing perception as something that is *done* by the perceiver. Perceivers possess and exercise knowledge of the "sensory effects of movement" (Noë, 2010, p. 250). In line with this idea, for instance, Erik Myin claims that we should place perception among other doings in which a person or an organism might engage, such as eating, walking, tracking, etc. (Myin, 2016, p. 88). What becomes relevant here is the unfolding sensory engagement of the whole person with its surroundings. As I discuss in section 3.1, this unfolding sensory engagement is determinant of the nature and content of perceptual experience. Drayson's (2017) claim that, for SMEn, perceivers are necessarily agents starts making more sense. However, it remains incomplete unless we consider SMEn's other key claim that perception involves knowledge. Because it is not only that in perceiving something someone is executing an action, but also that in this action the person deploys a set of abilities.

[20] In what concerns the enabling features of a phenomenon, I am not claiming here that De Jaegher et al. are not right in distinguishing between contextual factors, enabling conditions and constitutive elements. This distinction will become relevant again in chapter 7.

According to Myin, taking SMEn's claim in the terms just discussed allows the view to respond to an objection mentioned at the beginning of this section. This concern relates to offline phenomena that share phenomenal features with perceptual experience. The worry is that if we take perceptual experience to be a doing, we will run into trouble with offline phenomena that share phenomenal features with those experiences, e.g. hallucinations. On the face of this, SMEn could opt to defend disjunctivism and deny that these experiences really have the same nature. However, another way to tackle these offline phenomena would be to claim that the abilities that are deployed in perceptual experience are somehow deployed in these cases as well. Although further discussion is required to tackle this objection, for my purposes, this should do.

Let me summarize what I have claimed in this section. The third claim according to which perception is necessarily kinetic, particularly if this is understood as a claim about the necessary involvement of movement in perception, has been ruled out. It is too strong to be tenable. Although we can claim that SMEn is committed to the first claim that movement is a contextual factor of perceptual experience, it seems that something stronger is needed to capture SMEn's claim about the active character of perception. Moreover, given a commitment to a different kind of claim, namely that perception is constitutively active, SMEn would be committed, to an extent, to a more robust claim about the causal role of movement in perception. This was our second claim, namely that movement is an enabling condition of perceptual experience. Let me put this differently. Given SMEn's constitutive account of perception as a doing, the view would predict that movement, e.g. self-locomotion, is in some important cases an enabling condition of perceptual experience insofar as a lack of it hinders either the development of a perceptual capacity or the performance of subjects in a perceptual task.

There is one interpretation of SMEn's claim about the active character of perception that I am leaving out. This concerns the claim that perception is active in the sense that it is *for action*. The claim arises from SMEn's commitment to the dependence of SMCs on movement. The claim that perception involves the possession and exercise of practical knowledge about the way sensory input changes after an interaction has been interpreted, at times, as the claim that perception *serves* to guide action (e.g., (Pylyshyn, 2001)). This is the claim that the relation between action and perception is instrumental. The distinction between the dorsal and the ventral visual system has been used against such claim. In section 4.1, I discuss the two visual system theory. Here, I just want to note that SMEn is not committed to such a view, namely that perception serves to guide action (see (Noë, 2010).[21]

In the following section, I discuss the second claim defended by SMEn, namely that perception involves some kind of practical knowledge. The claim that perception is active and the claim that perception is knowledgeable go hand in hand. For SMEn, perceiving is not only something that is done by someone, but something that is done masterfully.

1.2.2 Perception is mediated by knowledge

Sensorimotor knowledge

As discussed previously, SMEn aims at challenging the sandwich view of perception. One way of doing this is by showing that there is not a clear-cut distinction between perception, action and cognition. According to this SMEn, perception consists, at least in part, in the possession and skillful execution of practical knowledge. An example that can be used to illustrate this is the following. Consider, for instance, how one spontaneously squints the eyes when an image looks blurry (Noë, 2004, p. 1). When we do this, the image appears clearer, more focussed. Squinting our eyes in this case suggests that we know how the image would change after doing this. SMEn claims that perception is constituted by our handle on the way sensory input systematically changes after an interaction.

This is sensorimotor knowledge, an implicit and practical knowledge of the lawful set of regularities that govern the systematic changes of input that follow an interaction. These regularities are, in turn, what is called Sensorimotor Contingencies (SMCs). In chapter 3, I discuss in detail the theory of SMCs. For now, it is sufficient to say that SMCs are covariations between movement either of the perceiver or of the object perceived, and its sensory outcome (Buhrmann et al., 2013; Noë, 2004; O’Regan & Noë, 2001). Sensorimotor knowledge consists, then, in our capacity or ability to act with sensitivity to SMCs or to the way sensory input varies given an interaction (see Silverman, 2016a, 2016b, 2018).

[21] This is related to Hurley’s (1998) claim that one possible avenue to reject the sandwich view of perception is to challenge the instrumental relation between action and perception.

Sensorimotor knowledge is advanced to account for the sense of presence that is prevalent in perception. Just as in the case of the blurry scene, one knows how the scene would change *if* one were to squint the eyes, something similar explains the case of e.g. the cat behind the fence. Even when one does not see the whole cat, one knows that by moving from side to side, one would see the bits of the cat that remain hidden behind the fence. The whole cat is present insofar as the agent knows how sensory information would change after certain interactions. When advancing an objection to SMEn, for instance, Daniel Hutto & Myin (2012) reconstruct the notion of sensorimotor knowledge in a way that can come in handy here. They claim that sensorimotor knowledge is meant to account "for the expectations that perceivers have concerning how things will appear in the light of possible actions" (Hutto & Myin, 2012, p. 25). Sensorimotor knowledge can be glossed counterfactually as the knowledge of how sensory information would change after an interaction between the perceiver and an object. For instance, the perceiver possesses the knowledge of how the waist of the cat would no longer be occluded by a picket, if she was to move to the right.

For perceptual experience to obtain, two conditions need to be satisfied (O'Regan & Noë, 2001). Firstly, perceptual experience requires sensorimotor interactions to be ruled by SMCs. This means that the interactions of the perceiver are governed by the lawful ways in which sensory input co-varies with the interactions of the agent given certain features of the perceiver as well as certain features of the object of perception or scene perceived. I discuss these two aspects of SMCs in more detail in section 3.2. Secondly, perceptual experience requires the *mastery* of sensorimotor laws. It is not only that the interactions of the perceiver correspond to or fit with a pattern of SMCs (which corresponds to the first condition), but that the perceiver *actively interacts masterfully* with the object in a way that is governed by these laws. As O'Regan & Noë put it, the perceiver "must be actively exercising its mastery of these laws" (O'Regan & Noë, 2001, p. 943). Consider again the case of the cat behind a fence. Perceiving the cat is governed by the practical knowledge of the way the hidden bits of the cat would appear if one were to move. It is not only that this can be described under e.g. certain laws of occlusion. For SMEn, the perceiver's experience of the whole cat already manifests that she has a handle on the relevant SMCs.

By introducing sensorimotor knowledge, SMEn is able to account for perceptual experience as a skillful interaction. The perceiver abstracts away certain sensorimotor lawlike regularities that are “available for them to recall” (O’Regan & Noë, 2001, p. 945). Perceptual experience consists, in part, in the mastery of SMCs in so far as the agent possesses knowledge of the rules that govern sensorimotor interactions. This, in turn, allows SMEn to account for the phenomenal qualities of perceptual experience. These are explained by the agent’s ability or capacity to access e.g. the hidden side of the tomato or the occluded bits of the cat. Moreover, as I explain in chapter 3 and 4, this allows SMEn to explain the differences between sense modalities. In short, for SMEn, sensorimotor knowledge fulfils two explanatory roles. On the one hand, it is meant to provide an account of perception as an interactive and knowledgeable phenomenon by showing that perceptual experience consists in the execution of embodied skills. And, on the other hand, it is meant to account for the fact that some of the phenomenal features of perceptual experience are determined by the systematic sensory consequences of non-actualized interactions (Silverman, 2018).

The problems of sensorimotor knowledge

As mentioned before, sensorimotor knowledge can be glossed counterfactually as knowledge about the resulting input of an interaction, e.g. *if* the perceiver were to move to the side, the waist of the cat would become visible. Up to this point, I have emphasized the active character of perceptual experience. In line with this, sensorimotor knowledge is described as implicit practical knowledge, a kind of know-how or embodied skill that is executed by the agent (Flament-Fultot, 2016; Loughlin, 2018; Silverman, 2016a, 2018). However, there is another interpretation available. This interpretation takes at face value the counterfactual formulation introduced at this section. According to this reading, sensorimotor knowledge is propositional knowledge about the way input would change given certain interactions, i.e. it is implicit propositional knowledge of counterfactuals involved in perceptual experience. Note here that this interpretation is explicitly rejected by O’Regan & Noë.

Neither of these readings comes without problems. On the one hand, the reading of sensorimotor knowledge as know-how faces an important challenge. Taking sensorimotor knowledge as know-how cannot fulfil the two explanatory roles mentioned above, namely accounting for the way in which perceptual experience is itself a skillful interaction, and for the qualitative features of perceptual experience. Sensorimotor knowledge understood in these terms can easily explain that perception consists in the execution of a skill. However, it might struggle to explain the qualitative features of perceptual experience that depend on non-actualized interactions, e.g. perceptual presence. The tension faced by SMEn is that it struggles to explain how perceptual experience consists in the execution of sensorimotor knowledge and, at the same time, is mediated by it in such a way that it could account for the expectations that explain the sense of presence (Hutto & Myin, 2012). For instance, when we see the tomato, we see it as a round object even when we do not interact with it in a way that would make it obvious that it is a round object: we do not need to hold the tomato, nor do we need to be disposed to turn it around to perceive it as a round object. The qualitative features of perceptual experience rely on the mastery of sensorimotor patterns of interactions that are not currently being actualized. To account for this, the proponent of SMEn must posit a second-order practical knowledge (Flament-Fultot, 2016; Silverman, 2018).

On the other hand, although the counterfactual reading allows for a more straightforward account of how SMEn fulfils the two explanatory roles mentioned above, it relies on the positing of internal representations to account for this (Seth, 2014). On this reading, perceptual experience consists in the "deployment" of internal representations of SMCs (Silverman, 2018, p. 158). Thus, this reading inherits the challenges that come with the positing of internal representations, such as: i. struggling to account for their content (see e.g., (Hutto & Myin, 2012, 2017)); and ii. arguing convincingly that the functional role attributed to putative representational states is, in fact, of a representational nature (see, e.g., (Ramsey, 2007)).[22] In addition to this, the practical aspect of sensorimotor knowledge becomes less straightforward.

As Silverman notes, SMEn seems to face a dilemma: either it takes sensorimotor knowledge as know-how and faces the drawbacks mentioned above, or it takes sensorimotor knowledge to be counterfactual and risks losing its practical aspect. To resolve this dilemma, David Silverman (2018) advances a proposal in which sensorimotor knowledge fulfils the two explanatory roles that it is meant to fulfil, while maintaining the practical character of sensorimotor knowledge. His proposal is to take perception as consisting of embodied skills that are, in turn, grounded on

[22] I am following Daniel Williams' (2018) summary of the challenges to representational accounts of cognition. He notes that it is possible to find in the literature a third challenge to representationalism. The cognitive functional challenge consists of two steps: (i) accepting that cognition should be subsumed to action or understood as an active phenomenon; and (ii) in consonance with (i), the positing of representations should be avoided partially or altogether (Williams, 2018, pp. 143–149). I do not address this point here because I do not consider it to be a challenge to representational accounts as such, but an explanatory strategy. Moved by the idea that cognition is an active phenomenon, some opt not to prioritize explanations that rely on internal representations. This strategy is similar to the explanatory strategy followed by SMEn. In section 2.2, I present this strategy.

sensorimotor knowledge. Silverman defends the view that sensorimotor knowledge is, in turn, an ability that grounds other embodied abilities. In that sense, it is a condition of possibility or a requirement to execute other embodied skills. In perception, agents manifest the possession of sensorimotor knowledge by exhibiting sensitivity to non-current interations. This knowledge grounds the execution of the embodied skills that constitute perception. Importantly, this is compatible with accepting that there are multiple ways in which sensorimotor knowledge might be causally enabled, including representational states.

As mentioned in the introduction, in this thesis I defend a predictive version of SMEn. The counterfactual reading of sensorimotor knowledge is one way to open the door to the notion of sensorimotor knowledge that is at play in predictive approaches to SMEn. Whatever the brain is doing, it should support the agent's sensitivity to non-current sensorimotor interactions. In chapter 6, I specifically discuss Seth's (2014) commitment to a counterfactual reading of sensorimotor knowledge. With the aim of specifying the neural implementation of sensorimotor knowledge, Seth exploits the aspects of predictive processing that are relevant for SMEn. More specifically, following Seth, my claim will be that predictive processing can accommodate the counterfactual aspect of sensorimotor knowledge.

But, what about all the problems that come with the counterfactual reading of SMEn and representations? These concerns can be deflated if we consider that whether and how these challenges are met depends not only on the previous claims by SMEn, but also on the causal account we offer of perceptual experience. Something that becomes relevant at this point relates to SMEn's commitment to naturalism. Although SMEn is an account of perceptual experience at the agential level (particularly, Noë's version of the view), it is committed to the claim that there is no clear-cut distinction between a theory of perception and a theory of the causal mechanisms that enable perception. This is because the properties that are found at the personal level, or better yet, that are attributed to the perceiver emerge from the complex sensorimotor interactions between the perceiver and the environment (Hurley, 1998). I discuss this in more detail in section 2.2.

In this chapter, I introduced SMEn as a theory of perception. I began by briefly discussing the two views from which SMEn walks away, namely the sandwich view of the mind and the snapshot conception of perception. SMEn offers an alternative account of perceptual experience that aims at addressing the tension between (a) the sense of richness and detail of perceptual experience, and (b) the limitations of sensory systems and the evidence that we sometimes miss quite significant events that occur right in front of us. SMEn's claim is that the detail and richness of perceptual experience is not due to us having a richly detailed model of the scene perceived. Instead, it is due to the fact that we can access all that detail. Perceivers are thought of, then, as skillful sensorimotor agents, agents that possess and execute practical knowledge of the way sensory input changes as one interacts with the world.

In the next chapter, I address two important matters within SMEn. Both relate to SMEn's commitments to externalism. As can be seen from the discussion in this chapter, perception is something done by the whole organism. For SMEn, referring solely to what happens in the brain, whatever it might be, would be insufficient to account for perceptual experience. Instead, it is necessary to refer to the interactions of the embodied perceiver. This claim can be, of course, associated with a radical position in what concerns representations. These two theses—a thesis about representations and a thesis about the vehicles of perceptual experience—are independent. However, they intertwine in interesting ways. If we think that referring to neural activations is insufficient to account for perceptual experience, then we might be less inclined to rely on internal representations to account for it. Regardless, committing to an externalism about perceptual experience is compatible with there being all kinds of representations in the brain. As we will see in the next chapter, this is the position defended by SMEn.

Chapter 2. Externalism and Representations in Sensorimotor Enactivism

As opposed to the view that perceptual experience is something that occurs in our head, Sensorimotor Enactivism (SMEn) claims, instead, that perception is something done by the whole organism or person. In line with this, SMEn rejects the idea that perceptual experience supervenes solely on neural states, in favor of a commitment to an externalism regarding perceptual experience. This is in the form of what Hurley (2010) calls *quality-externalism*. According to this thesis, perceptual experience supervenes on the brain-body-world system.

In this chapter, I discuss SMEn's specific brand of externalism. In relation to this, in this chapter, I address two matters. Firstly, I begin by discussing SMEn's position towards representations. As discussed, SMEn denies that the sense of presence of perceptual experience is, for instance, explained by positing representations. Still, SMEn does not deny that there might be internal representations involved in perceptual experience. In that sense, SMEn is best described as a representation-friendly view.

Nevertheless, despite the fact that the view might accept the positing of some representations, there is a sense in which SMEn is a *non-representationalist* view. By this I mean that SMEn follows an explanatory strategy that prioritizes the interactions of the perceiver with her surroundings, over a reliance on the internal structure of the perceiver. Moreover, it is by means of an argument to the best explanation that SMEn supports the commitment to *quality-externalism*, i.e. an externalism regarding perceptual experience. I discuss this in the second section of this chapter. Let us now turn to SMEn's position towards representations.

2.1 Representation-friendly SMEn

As discussed in the previous chapter, SMEn emphatically rejects the claim that perception consists in a process of building a model or representation of the scene perceived. Naturally, the view is associated with the anti-representationalist wing of cognitive science. However, among its proponents, there is no consensus on whether SMEn really is anti-representationalist, or whether SMEn's position should be characterised differently instead, since some proponents seem to be willing to accept at least some kind of representations.[23] In this section, I argue that SMEn's position is best characterised as representation-friendly. By that I mean that, while SMEn constrains both the *content* of the representations that might be involved and the *roles* these might fulfil, it is compatible with the claim that some representations might be involved in perceptual experience. In the first part of this section, I discuss the way SMEn constrains both the role and the content of the representations that might be involved in perception.

2.1.1 Constraining representations

SMEn is an alternative account of perceptual experience that opposes approaches that aim at addressing the problem of perceptual presence by positing representations of the scene perceived. Nonetheless, the view does not simply reject the thought that there is any use for the notion of representations when studying perception and the mind. Let me start this section by reconstructing some points made by O'Regan (2011) in regards to this matter.

O'Regan begins by noting that the use of representations is "ubiquitous in cognitive science" (O'Regan, 2011, p. 62). Despite this, theorists might have reasons to prefer alternative explanations and explanatory tools to representationalist approaches. This is the situation in the case of SMEn's proposal. As discussed, while representationalism can offer an account of, e.g., change blindness, the alternative account advanced by SMEn does not require the positing of compensatory mechanisms to explain why one fails to notice a salient change in the scene. In this way, it is argued, the account fares better than at least some of the representationalist alternatives. For O'Regan, one cannot just reject representations *tout court*, since positing representations might prove to be helpful in explaining other phenomena.[24] Representationalism is rejected in this specific case because there is a more feasible alternative. What led him (and other proponents of SMEn) to the rejection of representations in the specific case of perceptual experience are the

[23] Noë (2004), Noë & O'Regan (2001), O'Regan (2011), Silverman (2018), for instance, can be taken to argue for a weaker perspective on the positing of representations. However, Anthony Chemero (2009), Hutto & Myin (2012, 2017), Noë (2012), O'Regan & Degenaar (2014), Myin & Victor Loughlin (2018), among others, offer a more radical perspective on their rejection of representationalism.

[24] It is worth noting here that some approaches do take the strategy of rejecting representations altogether. Chemero (2009), for instance, advances his radical perspective as a competing theory to representational alternatives. Again, his aim is to provide alternatives to representational explanations. For him, the radical, non-representational, stance can better explain the same phenomena explained by positing representations. Hutto & Myin (Hutto & Myin, 2012, 2017), on the other hand, argue against the possibility of providing a naturalised account of content. Given that representations are contentful internal states, we should avoid positing them.

confusions to which the supposition that perception is in the business of building a representation has led. For instance, O'Regan claims, certain features of the retinal image—e.g. that it is inverted—led scientists to attribute this same feature to our experience—i.e. led them to claim that "we should also see the world upside down" (O'Regan, 2011, p. 63). Following Dennett, O'Regan refers to this mistake as the content-vehicle confusion: features of the vehicle are attributed to the content.

O'Regan notes that simply denying that there are any kinds of representations involved in perception might be untenable in light of not only the case of the retinal image, but also two other cases: first, evidence from the performance of subjects in mental rotation tasks indicating that mental images "really do behave like pictures" (O'Regan, 2011, p. 64); and second, the case of the cortical map located in area V1 of the neural cortex, where information is represented in terms of its location in space. In both cases, the anti-representationalist might insist that one would have to show that these two structures truly deserve the name of 'representations'. But, although some work is needed to respond to the worries that come from anti-representationalism, O'Regan's claim is that there are some cases that are good candidates to be labelled as 'representations'. Moreover, if the involvement of representations in perceptual experience is justified, SMEn must be able to accommodate them.

In light of these cases, it appears to be beneficial to articulate SMEn as a view that can be made compatible with the positing of at least some kind of representations. So, let me turn first to a conditional question, namely *if* SMEn were to posit representations, what kind of representations would these be? I begin with this question because, even when some proponents of SMEn might be willing to admit that it is possible that representations play a role in perception, they are still committed to limiting the kind of representations that should be posited in light of the proposed understanding of perception. In other words, those who claim that SMEn does not reject representations *per se*, claim instead that the theory constrains the positing of representations. Here, I distinguish between two different ways in which the positing of representations is constrained. Firstly, SMEn suggests that the *role* of representations that can be posited should be constrained (Noë, 2004; O'Regan, 1992, 2011). And, secondly, SMEn is committed to the claim that the *content* of these representations should be constrained as well.

Constraining the role of representations

Let us review the first constraint, namely the constraint placed on the *role* representations play in perception. See, for instance, the following claim by Noë:

> The claim is not that there are no representations in vision (…) The claim rather is that the role of representations in perceptual theory needs to be reconsidered (…) It is a mistake to suppose that vision just is a process whereby an internal world-model is built up, and that the task-level characterization of vision (…) should treat vision as a process whereby a unified internal model of the world is generated. This is compatible with there being all sorts of representations in the brain, and indeed, with the presence of such representations being necessary for perception (Noë, 2004, p. 22).

This position involves, firstly, the rejection of a view that takes perception to be in the business of building an internal model of the scene perceived. For SMEn, a model or a representation of the world should not be the *end-result of perception*. Even when the best way to articulate a causal account of perceptual processing is in terms of representations, perceptual experience does not consist of representations. The position expressed here by Noë is consistent with a view according to which perceptual experience is *enabled* by representations but is not *constitutively* representational. As argued in the previous chapter, the best way to articulate SMEn's claim that perception is active—that it is itself the execution of a skilful action—is as a *constitutive* claim. According to McDowell's (1994) proposal, a constitutive claim is concerned with what perception is. As opposed to this, an *enabling* claim is a claim about the mechanisms involved in perceptual experience. Even when representations might not be a defining feature of perception, they might still pertain to its causal story. This distinction will appear once more in the next section (section 2.1.2).

Constraining the content of the representations involved

In addition to rejecting the idea that the end-result of perception is a model or representation of the scene perceived, SMEn constrains the *content* these representations might have. SMEn rejects an account of perception that involves the generation of a mental model of the scene perceived that is neutral with respect to the actions and constitution of the perceiver. For instance, SMEn rejects internal allocentric representations of space, e.g. representations that preserve metric relations between the objects of the scene perceived. Consider, for instance, the following claims by O'Regan: "[v]iewers simply take the incoming information as it comes, and do not attempt to integrate it into a precise, metric-preserving internal representation, but only into a kind of non-metric, schematic mental framework, or structural description" (O'Regan, 1992, p. 471). And: "the passive sensation we get from the retina or some iconic derivative of the information upon it (…) is being used to supplement a mental schema we have about the results of the possible actions that we can undertake with our eyes (or heads or bodies)" (O'Regan, 1992, p. 472).

So, if SMEn is to posit some kind of representations, besides these not being the end-result of perception, these should meet the following requirements: they should be representations of the scene perceived that are indicative of further possible contextual interactions, and that are agent specific, thus, referring to the agent's constitution and history of interactions. This proposal can be articulated in terms of what has been called "action-oriented representations".[25] As opposed to more traditional conceptions of representations, action-oriented representations depend on and are meant to be specific to the context in which a system, animal or person might be engaging (Wheeler, 2005, pp. 196–197). Moreover, they refer specifically to the actions that might be undertaken by the system in that particular context. In other words, these representations emerge from the specific history of interactions of the system in question as "control-structures" for an action specific context of engagement.

Here, I have discussed the features of the representations that would be compatible with SMEn. In the following paragraphs, I argue that even when SMEn denies that representations are constitutive of perception, the view can accept the positing of representations.

2.1.2 Making room for representations

At the beginning of this section, I discussed O'Regan's comments on why representations should not be rejected *tout court* in discussions concerned with perception and perceptual experience. Two of the examples he sets forward relate to specific aspects of the visual system, i.e. the retinal image and the cortical map of cortical area V1. So, when pressing SMEn in connection to the need

[25] For a discussion, see (Cappuccio & Wheeler, 2012; Clark, 1997; Wheeler, 2005).

of advancing an account of perceptual experience that posits representations, one question that might arise is related to these structures: what is SMEn's proposal to make sense of them?

According to O'Regan, the *quid* of the matter is not in the fact that some information is available whenever the object that caused it is not, nor that this information is, e.g. isomorphic to the object perceived in some relevant sense. In other words, it is not problematic that we can make sense of these structures *as* representations. Instead, what is relevant is the way the perceiver *uses* this information or this representation. The evidence that there might be "all sorts of representations in the brain"—as Noë also puts it—should not worry the proponent of SMEn (Noë, 2004, p. 22). What explains that, e.g. one sees the whole cat, rather than bits of cat, is that one is able to deploy a certain set of embodied skills. When exercising these skills to access the relevant sensory information, representation-like structures might be activated.

Now, even when SMEn constrains the role and content of the representations that might be involved in perceptual experience, there is an aspect of SMEn's description of sensorimotor knowledge that seems to call for representations. In the next section, I discuss this aspect.

Sensorimotor knowledge as a representation-hungry domain

As mentioned earlier, sensorimotor knowledge is defined as our "capacity to respond (…) with sensitivity" to the systematic ways sensory input changes with movement (Silverman, 2016, p. 283, see also Silverman, 2018). When deploying sensorimotor knowledge, the perceiver invokes non-current states of affairs that are not readily available in the environment to make sense of, or to successfully engage with, the object of perception. Furthermore, this deployment is meant to explain the qualitative aspects of perceptual experience. The problem is that the description of sensorimotor knowledge as a capacity to respond to absent features seems to invite a reading according to which sensorimotor knowledge *requires* the positing of representations.[26] When thought of in this terms, sensorimotor knowledge seems to be a representation-hungry domain.

[26] In a similar vein, Downey (2017) points out that SMEn's description of notions like attention and sensorimotor knowledge is such that it seems like they should be explained in representational terms (Downey, 2017, pp. 2–3). Myin & Loughlin (2018), on the other hand, consider that although SMEn is committed to the rejection of representations for explaining perceptual experience, the view might posit representations for other theoretical purposes (Myin & Loughlin, 2018, p. 212).

Clark & Josefa Toribio (1994) offer a prominent description of problem domains that require representations. For them, a condition to consider a problem domain representation-hungry is the following: "The problem requires the agent to be selectively sensitive to parameters whose ambient physical manifestations are complex and unruly (for example, open-endedly disjunctive)". They add that a domain is representation-hungry if it requires sensitivity to "states of affairs whose manifestation in the sensory inputs is (…) attenuated" (Clark & Toribio, 1994, p. 419).

Let us go back briefly to sensorimotor knowledge. The latter fits this description of representation-hungry domains. When deploying sensorimotor knowledge, the perceiver invokes non-current states of affairs. For instance, the perceiver invokes the knowledge that, when one object occludes another, a change of location will enable the perceiver to see a section of the occluded object that was previously covered. The perceiver invokes counterfactual knowledge about the way input changes after a given interaction.

So, we might think that SMEn's proposal that perception requires the possession and execution of such knowledge, immediately places SMEn within the representationalist wing of cognitive science. We should not worry, of course, primarily about where SMEn belongs. The reason why this is relevant is that is shapes the kind of research one might engage in. If perception involves the possession and execution of sensorimotor knowledge, and sensorimotor knowledge requires representations, research efforts should be put in the identification of the relevant mechanisms and structures that fulfil such role. It is for that reason that, even when the claim that sensorimotor knowledge requires representations seems to follow straightforwardly from this description, it is worth paying attention to the challenges that have been set forward to the description of representation hungry domains as domains that *require* representations. I review these challenges in the next section.

Challenges to representation-hungry domains

Now, the idea that representation-hungry domains can only be accounted for by means of representations has been challenged in the literature. In the context of a discussion against anti-representationalism, Wheeler (2005) raises this issue and refers to Christian Scheier & Holf Pfeifer (1998), who examine category learning for autonomous agents, a problem domain that faces several challenges. For starters, category learning requires the identification of sensory stimuli that are relevant for an agent in an environment that is constantly and continuously changing. Secondly, it also requires identifying when different sensory stimuli are coming from the same object. Scheier & Pfeifer develop a strategy to deal with category learning that does not require representations. Instead, the authors exploit the sensorimotor interactions of the agent—in this case the Khepera robot—to categorize objects. Although this case is not directly related to perception, it allows us to show that domains that are described as representation-hungry can be explained using alternative strategies.

In their experiments, Scheier & Pfeifer use the Khepera robot, a compact mobile robot equipped with sensors and a processor. One possible way of categorizing the objects with which the robot is interacting is by mapping the statistical space of the readings of its sensors. However, implementing this strategy is complicated. The statistical regions occupied by inputs coming from the same object are, in many cases, far apart and overlap at times with those of other objects. So, categorizing objects from this data becomes difficult. This problem is not tractable, since an accurate categorisation is highly unlikely. For this reason, Scheier & Pfeifer locate category learning within what Clark & Thornton (1997) call type-2 problems, i.e. problems where regularities cannot be found within the data itself.

To make this kind of problem tractable, it is necessary to transform the data in a way that regularities become visible. One strategy to do this is by improving the processing of the input: either by using previously learned representations, or by incorporating constraints based on "a priori assumptions" about the data (Scheier & Pfeifer, 1998, pp. 33–34). However, there is another strategy available. Instead of refining the processing of the input, Scheier & Pfeifer show that data can be "actively structured", i.e. structured on the basis of the sensorimotor interactions of the mobile agent, thus making visible the regularities that transform this into a tractable problem. Instead of recruiting representations, Scheier & Pfeifer rely on the interactions of the agent to structure the data. Although positing representations would be one way to proceed, the authors show that there is this other available strategy.

In line with Scheier & Pfeifer, the notion of representation-hungry domains has also met resistance from the trenches of other enactivist approaches. Julian Kiverstein & Erik Rietveld (2018), for instance, argue that the notion of representation-hunger relies on the assumption that, in these domains, cognitive systems are *decoupled* from the environment. For this reason, representation-hungry domains become challenging for non-representational frameworks. If the system is characterized, in these cases, as decoupled from the environment, then it cannot but rely on its internal states to tackle higher-order cognitive tasks.[27]

In contrast, they defend what they call a “strict continuity” between higher order and basic cognition and aim to provide an alternative framework in which representation-hungry domains do not rely on representations (Kiverstein & Rietveld, 2018, p. 148). The authors deny that there is a divide between offline and online cognition. In consequence, they claim that their approach will not face this challenge because the cognitive agent, as it is described under this proposal, is never fully decoupled from her environment. What allows the agent to deal with the absent and the abstract is her skilful coordination with the environment and the affordances it offers. Current interactions delimit and define the absent and abstract features that might be invoked by the agent. These features are made available as she interacts with her surroundings.

In the approach developed by Kiverstein & Rietveld (2018), the environment consists of virtual conditions, i.e. the possible interactions that are available to the agent in a given situation. These conditions are enacted through the agent’s activities and are directed at the maintenance of the system. For them, these are meaningful conditions in that they are shaped by the agent’s interests and concerns. The interactions of the agent can be understood in terms of action-readiness to the affordances offered by these meaningful environments. They illustrate this with the case of the office environment: “When we enter an office, we ready ourselves for certain possibilities—the colleagues we are likely to meet, the activities we regularly perform while at work, the interaction with our boss, the route we need to take to reach our office and so on” (Kiverstein & Rietveld, 2018, p. 153). Agents are sensitive to different structures of affordances and they skilfully coordinate their interactions with each of them.

Kiverstein & Rietveld propose to understand cognition in these terms. Representation-hungry domains, in particular, should be explained on the basis of the complexity of the system. From their perspective, an agent becomes more complex the farther away from her metabolic needs her actions are. In other words, the agent becomes more complex the more she acts to fulfil the need to e.g. meet with friends rather than the need to regulate her body temperature. The more complex

[27] Kiverstein & Rietveld assume here a very tight link between representation and decoupleability. Against this assumption, see e.g. (Wheeler, 2005, p. 211 ss).

the agent is, the farther away from the current situation the relevant affordances are. The affordances to which the agent responds are less tied to current situations, or, in other words, to the here and now. Kiverstein & Rietveld claim that “[a]s the agent grows in complexity, it becomes sensitive to tendencies and trajectories in the evolution of its own states in relation to the environment that stretch steadily further through time” (Kiverstein & Rietveld, 2018, p. 152). The authors propose to account for representation-hungry domains in terms of action readiness. For instance, a case of episodic memory imagining is understood as the re-enactment of past experiences and actions. It is because the agent can skilfully coordinate their interactions with different environments that she can enact such an episode.

The arguments show as well that, when dealing with domains that require sensibility to regularities that are not readily available in the environment, there are strategies available that do not require us to draw on representations. Domains that can be described as representation-hungry domains are not explained solely by means of internal representations.

Before turning back to the case of sensorimotor knowledge, let me briefly review one last challenge advanced to the notion of representation-hungry domains. Jan Degenaar & Myin (2014) show that neither cognitive activities in domains that involve something *absent*, nor domains that involve something *abstract* necessitate representations.

In what concerns *absent* features, they offer a case in which removing an object with which an agent is interacting does not entail the requirement of representations. Think of the following case: when going from the coffee machine in her office to her desk, an agent follows a trajectory that requires her to go around a table that stands in the middle of the room.[28] One day, the table is removed, but the agent still follows the same trajectory. In this case, the nature of the agent’s behaviour seems to be the same before and after the removal of the table, i.e. the agent exhibits the same capacities. The mere removal of the table does not seem to call for a change in the nature of the agent’s capacities. So, if the behaviour was firstly characterised as representational, it seems that, after the removal of the table, it should still be characterised in this way. And the other way around, if it was firstly characterised as non-representational, after the removal of the table, it should still be characterised as non-representational. Accounting for the nature of the capacities

[28] The examples presented in this section are adapted from (Degenaar & Myin, 2014). For a similar case, see (Wheeler, 2005, pp. 215–216).

involved in this case requires an additional argument to claim that these are either representational or non-representational.

In what concerns *abstract* features, Degenaar & Myin (2018) start by showing that Clark & Toribio's (1994) argument depends on a distinction between unitary and diverse properties. For Clark & Toribio, it seems that to deal with diverse properties, the agent requires an internal representation that unifies these diverse properties. For Degenaar & Myin, the idea that is at play here is that, whenever there is a convergence of diverse stimuli that correlates with one neural correlate, this internal state is typically conferred the status of a representation.

However, the authors argue, it does not follow from the fact that diverse properties need to be unified in this way that internal representations perform that task (i.e. the task of unifying dissimilar sensory stimuli). To support this idea, the authors ask us to consider the many different cases in which causes with diverse physical manifestations have one "common physical effect" (Degenaar & Myin, 2014, p. 3646). For instance, a metal nail that can be moved by a magnet, one's hand, or by the wind. However, we would not conclude from this that the movement of the nail represents its different causes. Similarly, with cases that concern cognition. As in the previous case, claiming that the agent is dealing with the *abstract* does not call by itself for a representational story. The authors, of course, need to give an alternative story. However, here what they aim at emphasizing with this example is that to go from the idea that a domain has abstract properties to the idea that it requires representations, one step is missing. We need another argument.

Prima-facie representation-hungry domains

A point that is crucial about the previous arguments is that, although they show that representation-hungry domains do not *require* representations, these domains can still recruit a strategy that relies on the positing of representations. In these domains, representations can become useful theoretical tools. This is what Silverman (2018) describes as *prima facie* representation-hunger.

Relying on McDowell's (1994) aforementioned distinction between *constitutive* and *enabling* features of perception, Silverman introduces a distinction between *constitutive* and *enabling* representationalism. As discussed, while enabling features are those that "play a causal role" in perception, constitutive features are defining features of perception: they are just what perception is. Therefore, while constitutive representationalism claims that cognition is necessarily representational, enabling representationalism claims that representations are contingent: representations are only causally relevant to explain a domain.

Let us go back now to sensorimotor knowledge. At the beginning of this section, I claimed that, when deploying sensorimotor knowledge, the perceiver invokes non-current sensorimotor interactions. If one were to describe sensorimotor knowledge as a domain that *must* be explained representationally because it invokes absent aspects of the object of perception, one would be committed to constitutive representationalism, the view that perception is constituted by representations. Constitutive representationalism claims that cognition, broadly construed, "could not occur unless there was internal representation" (Silverman, 2018, p. 168). The problem is, as Silverman sees it, that constitutive representationalism goes one step too far. In the context of SMEn's proposal, the description of perception as a representation-hungry domain becomes problematic because it risks losing the practical aspect of sensorimotor knowledge and mandates the positing of representations. For Silverman, the description of sensorimotor knowledge advanced by SMEn has features that invite a representational reading, but do not require it. This is what Silverman calls enabling representationalism. This kind of representationalism accepts that there is some "theoretical work that could be done by internal representation" without claiming that perception just is the building of representations (Silverman, 2018, p. 168).

Silverman's distinction between constitutive and enabling representationalism becomes more plausible when we consider the arguments that have challenged the idea that representation-hungry domains *require* representations. The challenges presented above show that cognitive domains that are described as representation-hungry domains do not *require* representations. Domains that are *prima facie* representational-hungry can be explained following both representational and non-representational explanatory strategies. Positing representations is just one of the strategies available to account for problem domains that exhibit the features previously described. So, given the description of sensorimotor knowledge advanced by SMEn, the theory can, in principle, be described as *prima facie* representational, i.e. as a theory that accepts the positing of representations. In claiming that perception is a *prima facie* representation hungry domain, SMEn is not committed to constitutive representationalism but to an enabling kind of representationalism that does not *require* representations. According to this view, then, representations could be posited to account for perceptual processing, without these constituting perceptual experiences.

Therefore, in principle, SMEn can be described as representation-friendly. One possible strategy to explain that sensorimotor knowledge involves invoking non-current states of affairs is of a representational nature. When looking at the causal story of the mechanisms[29] that enable perceptual experience, we might encounter a scenario in which sensorimotor knowledge is enabled by structures that are best described as representations. However, this strategy is not mandatory. It might also be that, as the agent interacts with her surroundings, the absent features that are constitutive of perceptual experience become available.[30]

As mentioned earlier, whether we define sensorimotor knowledge as *requiring* or merely *inviting* the positing of representations influences what we are looking for (or where we are looking) when it comes to the mechanisms that are at play in perceptual experience. In other words, our accounts of what is constitutive of a phenomenon such as perceptual experience have an impact on the strategy undertaken when looking into the enabling story that explains this phenomenon and its features. Models that take perception to be the process by means of which we build representations of the scene perceived, to be used for other mental processes typically turn to the brain when looking for those representations. SMEn deploys a different strategy, one that emphasizes the relevance of the interactions of an embodied perceiver. This is the focus of the next section. Let me just reiterate, before closing this section, that this strategy is compatible with the possibility of there being some structures involved in perceptual experience that might be described as representations. This will become relevant in section 6.3, when discussing the predictive version of SMEn.

2.2 Explanatory externalism

One way to open the conceptual space to the possibility that the vehicle of perceptual experience is not just the brain is the view that perception is an action performed skilfully by an embodied agent in which the perceiver exercises her ability to access further sensory information, and that the content of experience is virtual, outstripping what we have all at once.[31] As mentioned in the

[29] Note here that I am using the term mechanism very loosely and independently from any reference to the mechanistic theories of the mind.
[30] It is worth noting that this strategy of distinguishing *enabling* from *constitutive* representationalism still faces the challenges generally advanced against representationalism in connection with content ascription (Hutto & Myin, 2012, 2017).
[31] In section 3.1, I discuss SMEn's position regarding the content of perception in more detail.

previous chapter, according to Hurley, empirical research benefits from conceptual liberality and from an abundance of conceptual resources. In line with this, Noë claims that SMEn's proposal "remove[s] the major theoretical obstacle to entertaining the possibility that experience might supervene not on the brain, but rather on brain-animal-world systems." Furthermore, Noë adds, *"[i]t is an empirical question whether our brains can do the work needed to enable us to enact our virtual worlds*" (Noë, 2004, p. 217, italics in original).

2.2.1 Naturalism in Sensorimotor Enactivism

I began this section by emphasizing the latter point made by Noë because I want to note here that SMEn is a naturalist position. SMEn is a theory of perception and perceptual experience that aims at addressing these phenomena from the point of view of the agent and her experience. What I have said up to this point concerns SMEn as an account of the constitutive features of perception, rather than of the mechanisms that enable this phenomenon. The naturalism to which SMEn is committed is one that, for starters, claims that a constitutive theory of perception is continuous with an enabling account. Whatever is said about the constitutive features of perception, this must be informed and supported by empirical research concerned with the phenomenon. Moreover, proponents of SMEn go one step further. Take, for instance, Noë's claim that a theory of perception must straddle both levels, the agential and the causal (Noë, 2004, pp. 31–32). In line with this, he rejects the idea that there is a clear-cut distinction between the personal or agential level and the subpersonal. Instead, the continuity between these two levels—and of the constitutive and enabling perspectives—is due to the fact that what is associated with the personal or agential level depends on the causal level. According to Hurley, features exhibited at the agential level emerge from the complex sensorimotor interactions between the agent and the environment. Hurley illustrates this continuity with the case of a circus performer whose complex movements are both a source of and dependent upon complex dynamic feedback loops between herself and the environment. For Hurley, in the case of the circus performer, the content of her intentions and perceptions depends on and emerges from this complex causal flow (Hurley, 1998).

The story offered by SMEn makes it possible that the causal story from which the content of perceptual experience emerges includes not just the brain, but the non-neural body and the environment. Moreover, when looking for the causal story that explains perceptual experience, the interactions of the embodied agent should be prioritized over a reliance on the inner structure. Why? Well, in at least some cases, this strategy might turn out to more efficient: rather than building an internal model of the scene perceived, the world can work as its own model. Nonetheless, as argued in the previous section, this should not preclude SMEn from accepting the possibility that other mechanisms are in place.

2.2.2 Sensorimotor Enactivism's explanatory strategy

Let us start by reviewing one example of SMEn's explanatory strategy. This strategy is made explicit in the theory of sensorimotor contingencies (SMCs) advanced by Di Paolo et al. (2017). Let us consider, for instance, one of the cases they advance: a case of tactile discrimination between two different shapes, e.g. discriminating between two light switches of different sizes but similar shapes, using only the tip of the finger (Di Paolo et al., 2017, p. 62).

The authors advance a model of this behaviour starting with a minimal cognition model (in this case, a network comprised of only two neurons), which uses agents with very simple sensory and motor structures. The coupling with the environment enabled by these structures becomes central to understanding the performance of the agent in a specific task.[32] By taking this explanatory strategy, the authors do not intend to deny the presence and influence of the elements that are internal to an agent. Instead, they aim at explaining the behaviour of an agent by emphasizing the determining role of the agent's interactions.

The explanatory strategy comprises two claims: (a) that the brain is insufficient to explain the behaviour of an agent; and (b) that the interactions of the embodied agent enjoy explanatory priority over any reference to the internal organization of the system.[33] In other words, as a strategy to understand and explain intelligent behaviour, the embodied agent's interactions with the environment should precede any reference to its inner structure. In that sense, an account of, e.g.

[32] The authors study the behaviour of minimal agents in discriminating narrow from wide shapes. The agents are meant to move away from the wider shape and approach the narrower one. The task cannot be undertaken by relying on direct sensory information because the shapes are randomly positioned at each trial. After several trials, agents that performed better, i.e. that distinguished wide from narrow shapes across different positions, were selected. The authors, then, model agents' sensorimotor patterns in order to identify and predict the trajectories visited by the agents, their long-term tendencies and regularities, and the differences between behaviours. The analysis also models the agent's internal structure. See Di Paolo et al. (2017).

[33] This strategy is implicit also in e.g. Di Paolo (2014) and Loughlin (2018).

neural dynamics, is subordinated to, and should be understood within, the context of the interactions of the embodied agent.

This same strategy is present in the work of Noë and O'Regan. See, for instance, this claim:

> But the brain's activation does not in itself constitute the seeing. In partner dancing, specifying the bodily configuration or brain state of the dancer is not sufficient to specify the dance (because we need additionally to know how the partner is currently interacting). Likewise, in seeing, specifying the brain state is not sufficient to determine the sensory experience, because we need to know how the visual apparatus and the environment are currently interacting (O'Regan & Noë, 2001, p. 966).

And the following by Noë:

> On the enactive approach, brain, body, and world work together to make consciousness happen (Thompson & Varela, 2001) (…) Experience is not caused by and realized in the brain, although it depends causally on the brain. Experience is realized in the active life of the skilful animal (Noë, 2004, p. 227).

As can be seen in these quotes, the claim for the explanatory priority of the interactions of the embodied agent is accompanied by remarks about vehicles of perceptual experience. So, the claim goes: an account of the interactions of the embodied agent is required to account for perception because perception is not realized solely in the brain. Rather, the realizing base of perception extends beyond the inner (neural) structure of the agent. This commitment is also formulated in terms of the causal/constitutive distinction: an account of the interactions of the embodied agent is required because perception is, constitutively, an interaction between the agent and the world. The brain either enjoys a 'mere' causal relevance or it is only partly constitutive. More needs to be said to explain the relation between the ontological and the explanatory claim, and to make the case for the ontological commitments at play. I discuss this in the next section. Here, it is sufficient to note that the explanatory claim is connected to an ontological claim, that is, the claim about the realizing base of perceptual experience. For SMEn, given that perception is constituted by the interactions of the embodied agent (or given that the agent-environment system realizes perception), accounting for perception by referring to the agent's interactions is required.

2.2.3 Explanatory externalism and the case for quality-externalism

Let me start by summarizing what was discussed in the previous section. For SMEn, the interactions of the embodied agent enjoy explanatory priority over the inner structure of the agent. This priority should be understood as an explanatory strategy that comprises two claims: (a) the claim for the insufficiency of the neural mechanisms in an account of perception; and (b) the claim that the structure and dynamics of the brain should be understood in virtue of the interactions of the embodied agent. As discussed earlier, this strategy is at times accompanied by a claim about the vehicles of perceptual experience. So, how does SMEn get there? As I discuss in this section, the claim does not follow straightforwardly from SMEn's account of perceptual experience.

It is helpful to recall here that one of the claims of SMEn is that perception consists in a skilful action of an embodied agent. The perceiver, from this perspective, deploys a set of bodily skills. Moreover, this description opens the door to the conceptual possibility that the vehicles of perceptual experience extend to the non-neural body and the world. In light of that possibility, SMEn pursues an explanatory strategy that prioritizes the interactions of the embodied agent. However, nothing in SMEn's story justifies a claim about the vehicles of perceptual experience. If SMEn were to rely on its claim that perceptual experience constitutively consists in the interactions between the embodied agent and the environment, the theory would be vulnerable to an objection from the internalist. This objection, formulated by Clark (2009), is a version of the causal-constitution fallacy objection advanced by Adams and Aizawa, according to which SMEn would be conflating factors that are only causally relevant, with factors that are constitutive to the phenomenon in question (Adams & Aizawa, 2010; Aizawa & Adams, 2009). According to Clark, SMEn confuses evidence for the role of the interactions of the embodied agent in training and tuning the system in question, with evidence for their role as vehicles of perceptual experience (Clark, 2009, p. 970).[34] So, surely, the interactions of the embodied agent are causally relevant for their perceptual experience. However, their role is not as vehicles of the experience.

To justify the stronger claim that perceptual experience is realized in "the active life of the skilful animal", in Noë's words, SMEn requires an argument for the best explanation. For starters, opting for its explanatory strategy is justified because the best way to account for perceptual experience

[34] For an overview of the debate, see Kirchhoff & Kiverstein (2019a, Chapter 2 and 3).

is by referring to the interactions of the agent. Furthermore, this is the best way to make sense of the agent's inner structure. Here, the claim is not only that external elements are constitutive of perceptual experience, but that taking this perspective provides the best explanation (Clark, 2009).

This is what Noë (2007) calls explanatory externalism. He claims that "the nervous system modifies experience by *keeping track of our relation to things around us*" (Noë, 2007, p. 465, italics in original). To make sense of the nervous system and of neural activity, it is necessary to pay attention to the interactions of the agent. Moreover, for Noë, the claim about the vehicles of perceptual experience is justified on the basis of the explanatory strategy, and not the other way around. Perception is constituted by the interactions of the agent (or, as he puts it, by the world) because these are the explanatory substrates, i.e. the realizers that are required to explain perception. He adds that "[n]eural systems are essential for experience, but they are not all that is essential, and so we cannot understand experience in neural terms alone" (Noë, 2007, p. 467).

Hurley (2010) makes a similar point in favour of what she calls *quality-externalism* or externalism about perceptual experience, namely, the view according to which the realizing base of perceptual experience encompasses, in at least some cases, the agent's dynamic interactions with the world. To make her case, Hurley examines cases of internal simulations that mimic sensory events and distinguishes between offline and online cases of simulation. While the former are cases where the system mimics sensory events but is not currently interacting with the environment, the latter, online simulation cases, are those where the system is currently interacting with the environment (Hurley, 2010, pp. 139–141).[35] Although in both cases it is undeniable that internal simulations are necessary to account for the simulated sensory event, it is not obvious that the agent's interactions are necessary. Hurley supports her claim that interactions are indeed necessary to account for these cases by arguing for the derivative character of simulations. In online simulations, the simulations' parameters are continuously updated by means of the agent's interactions. The interactions of the embodied agent provide "ongoing tuning and maintenance" (Hurley, 2010, p. 142). In the case of offline simulations, on the other hand, the role of the feedback loops that are mimicked results from the agent's developmental history and history of interactions. In consequence, to make sense of both cases of simulations, it is necessary to make sense of the agent's (history of) interactions. Furthermore, Hurley's claims that it is in light of the elements

[35] Hurley's (2010) description of these cases is reminiscent of proposals of predictive processing.

that are necessary for our explanations that we should fix the boundaries of what constitutes the system. The system is individuated based on our explanations. Similar arguments are raised by Kirchhoff & Kiverstein (2019a). They claim that, when the relevance of the generative process in grounding perceptual experience is recognised, the dependence of simulations on the agent's interactions and the primacy of the latter become evident (Kirchhoff & Kiverstein, 2019a, p. 60).

There are a couple of arguments that are worth considering at this point, both advanced by Clark (2009). If the claim is that it is necessary to refer to the interactions of the embodied agent with the environment to make sense of the contributions of the brain to perceptual experience, Clark worries that the very same strategy might be advanced by the internalist. Standard internalist views might argue that a mental state is neurally supported in many different ways and that, to distinguish between neural contributions, considerations about external elements might be relevant. There is nothing about the argument that supports externalism (Clark, 2009, p. 971).

Clark himself notes that SMEn's argument is slightly different: SMEn provides an account that offers an additional explanatory advantage. It is not only that referring to external elements allows a distinction between neural structures that are involved in perceptual experience, a distinction that would otherwise be ignored. It is also that SMEn provides an answer as to *why* this pattern of interactions supports e.g. visual rather than auditory experience (Clark, 2009, p. 972).[36] But, why should we think that the internalist does not have the resources to explain this, and do so without conceding that the vehicles of perceptual experience extend beyond the boundaries of the brain? For Clark, we should not underestimate the internalist, because she might be able to provide an answer to that question.

Clark might be right that, just like SMEn, some internalist views might be able to provide an answer to this question. Nonetheless, it becomes a matter of providing the best possible explanation. It is no longer a matter that is solved by referring to a previously accepted claim about the vehicles of perceptual experience. SMEn's claim is that, in domains in which the contributions of internal and external factors are entangled, the best possible explanation will be provided by an account that encompasses both. In connection with this, Hurley claims that in dynamical explanations, the causal-constitutive distinction is not explanatorily transparent (Hurley, 2010, p.

[36] In section 4.2, I further develop this point.

126).[37] The claim would be that to understand perceptual experience, the best explanation available will refer to the interactions of the embodied agent.

At this point, the claim about the vehicles of perceptual experience has almost become otiose: distinguishing between the causal and the constitutive is not explanatorily informative. It is perhaps for this reason that Hurley briefly entertains the possibility of dropping the distinction altogether. She claims that: "it isn't clear how causal-constitutive talk can be mapped onto complex dynamical explanation, or even what work a criterion of the constitutive is supposed to do in this context. We don't have such a criterion here, and it isn't clear that the cognitive sciences need one in order to provide good explanations" (Hurley, 2010, p. 126).[38]

In the context of SMEn, the worry might be that dropping the distinction misses the point of SMEn. It is not only that paying attention to the contributions of the world allows for a better account of perceptual experience. It is, instead, that the claim that the interactions of the perceiver constitute and determine perceptual experience becomes evident when my explanation cannot do without referring to them. We come to know about what constitutes the vehicles of perceptual experience by means of the role they play in our explanation, because otherwise we would not be able to account for perceptual experience.

In this chapter, I began by discussing SMEn's position towards representations. Even when perception is not constitutively a process by means of which the perceiver builds a representation or model of the scene perceived, this does not rule out that representations are somehow involved in perception. SMEn is a view that can be described as representation-friendly. Representations are not constitutive of perceptual experience, however, they might pertain to our causal story. Moreover, given SMEn's position about perception being a skillful action, the view recommends an explanatory strategy that prioritizes the interactions of the agent. This strategy follows from SMEn's claim about what constitutes perception. However, it does not support a more robust claim about the vehicles of perception. To justify the latter claim, SMEn relies on the explanatory

[37] In a similar vein, Kirchhoff & Kiverstein (2019a) challenge the standard causal-constitutive distinction. They argue for a notion of constitution as a diachronic relation. In this way, the distinction between the causal and the constitutive becomes blurry. See (Kirchhoff & Kiverstein, 2019a, Chapter 6).

[38] For another approach to the causal-constitutive distinction that is in the vicinity, see (Loughlin, 2018a; Loughlin & Zahidi, 2017).

externalism thesis, according to which the vehicles of perceptual experience include the non-neural body and the world because these are the explanatory substrates of perceptual experience.

In Part II, I revise the theory of sensorimotor contingencies (SMCs). According to SMEn, SMCs play a crucial role in accounting for perceptual experience, since they fix, at least in part, its character. Moreover, SMCs provide the bridge between SMEn as a theory of perception and a proposal about the mechanisms that enable perception. In chapter 3, I focus on the features of this theory. I begin by reviewing the theory of content advanced by SMEn and move on to discuss SMCs. These depend in an important way on the body of the perceiver. This dependence has led to the concern that SMEn is a chauvinistic theory of perception. This is the concern I address in chapter 4.

II. The theory of Sensorimotor Contingencies, embodiment and some concerns about chauvinism

Chapter 3. The theory of Sensorimotor Contingencies

In Part I, I discussed the main features of Sensorimotor Enactivism (SMEn), the view that perception is a skilful action undertaken by the whole person or organism. In chapter 1, I focused on the active and knowledgeable aspects of perceptual experience, as well as on the motivations that drive this conception of perceptual experience. In chapter 2, I focused on SMEn's commitments towards externalism and its compatibility with the acceptance of at least some representations.

SMEn is advanced as an alternative to representational theories of perception. The view can more easily account for puzzling aspects of perceptual experience, such as the phenomenon of perceptual presence, in which what we get in experience outstrips what is strictly presented to our senses at a time. For instance, when seeing a cat behind a fence, the content of the experience is a cat, rather than bits of the cat. If someone were to report the sighting of a cat behind a fence in terms of seeing bits of cat, that person would be misconstruing her experience (or, at least, one would hope so). What is puzzling is that one sees a whole cat even when one only receives stimuli that would correspond to the sections of the cat that are visible. As explained in previous chapters, instead of invoking an internal representation of the scene perceived, SMEn claims that it is an implicit practical knowledge of the way sensory input changes after an interaction (i.e. sensorimotor contingencies or SMCs) that explains the perceptual experience of seeing a cat behind a fence.

SMCs are central to SMEn's position. They play a crucial role in explaining that perceptual experience has the content and quality it has. Recall that, according to this view, for perceptual experience to occur, two conditions need to be satisfied. Firstly, the interactions have to be governed by SMCs, i.e. by the lawful ways in which sensory input changes after an interaction. Secondly, the perceiver must be in possession of and exercise practical knowledge of SMCs. It is by means of these sensorimotor interactions that the perceiver enacts perceptual experience. As we should see later, the character of this experience is partly fixed by SMCs.

In this chapter, I discuss the main features of the theory of SMCs. However, there is one more point I would like to note before moving on to the first section. Recall that, as discussed in section 2.2, SMEn is a naturalist view of perception. The view is meant to be continuous with a scientific account of the mechanisms of perception. SMCs provide the bridge between a theory of perception at what would be, under other considerations, the personal or agential level, and a theory about the subpersonal level.

To understand more fully the explanatory role played by SMCs, I begin this chapter by discussing how SMEn accounts for the character of perceptual experience. Section 3.1 says a bit more about the place of experience and the relevance of addressing the explanatory gap in the project of SMEn. I, then, discuss SMEn's view on the content of perception. As I explain, SMCs contribute to fixing the character of perceptual experience, i.e. its content and quality. In section 3.2, I discuss SMEn's theory of SMCs. These are the systematic patterns of co-variation between sensory information and movements of the perceiver or the object(s) perceived. SMCs are sensitive to physiological features of the perceiver. It is in virtue of this claim that a concern arises that SMEn might be a chauvinistic theory of perception, in that the view holds a commitment to the idea that for a system to be a perceiver like us, it needs to have a body like ours. In other words, the view restricts the attribution of e.g. visual perception in virtue of details of the perceiver's body. I address this problem in chapter 4. For now, let us turn to the character of perceptual experience and the relevance of addressing the explanatory gap.

3.1 Accounting for the character of experience

3.1.1 Bridging the explanatory gap

Think back to the phenomenon of perceptual presence, in which experience outstrips what is strictly present to our senses at a time. SMEn aims at doing justice to this puzzling feature. Of course, accounting for the richness of perceptual experience can be achieved in many different ways. What matters is that, according to SMEn, these features—e.g. that our experience is rich and detailed—should not be overlooked. Furthermore, SMEn's view is that accounting for this can be more easily attained by referring to the embodied skills deployed by a perceiver. In SMEn, experience is placed at the centre of the theory of perception. One of the reasons SMEn moves away from other approaches is its commitment to providing an account that is *phenomenally apt*, i.e., that does justice to perceptual experience. For SMEn, perception is not a mental process that provides information required for thought and action, and that happens to have unexplained qualitative features. Rather, it is a skilful and sensitive engagement of which certain qualitative features are constitutive. Ultimately, to provide a satisfactory account of perceptual experience, it is necessary to provide an account of these qualitative features (or the lack thereof). This includes a commitment to addressing the explanatory gap. According to its proponents, one of SMEn's virtues is that, by proposing that perception is a way to engage and interact with the environment (Myin & Loughlin, 2018), it can overcome the explanatory gap.

It is worth distinguishing the different problems related to the explanatory gap. For starters, it is important to distinguish between relative and absolute explanatory gaps (Hurley & Noë, 2003). On the one hand, the problem of the absolute explanatory gap is concerned with the question of why perceptual experience, in our case, is accompanied by a phenomenal feel at all (Hurley & Noë, 2003).[39] On the other hand, there are a couple of related questions that arise not absolutely, but when putting one mental state or process against each other. The question, then, is not why perceptual experience is accompanied by a phenomenal feel at all, but why it is accompanied by *this* quality rather than *that*. These are the comparative explanatory gaps and it is possible to further distinguish between an *intermodal* and an *intramodal* explanatory gap. The intermodal comparative explanatory gap asks "[w]hy does [a] certain [process] give rise to visual rather than auditory experience?" (Hurley & Noë, 2003). The intramodal comparative explanatory gap relates instead to the question of "[w]hy does [a] certain [process] give rise to experience of red, say, rather than experience as of green?" (Hurley & Noë, 2003).

Addressing the explanatory gap, in its different modalities, is one of the central concerns of proponents of SMEn. Moreover, the view is that this is done by referring to the patterns of sensorimotor interactions that are characteristic of different perceptual interactions (Myin, 2016, p. 88). Hurley & Noë (2003), for instance, aim at showing how vision, hearing, and touch can be distinguished by referring to the difference in patterns of sensorimotor relations between an agent with the environment. Myin & Loughlin (2018), in turn, argue that SMEn is committed to the idea that experience is identical with the interactions of the agent and the environment. For them, having e.g. a visual sensation is identical with the sensitivity or attunement to and interaction of the agent

[39] Hurley & Noë articulate these questions in terms of neural processes. For example, why neural processes are accompanied by a certain phenomenal quality. I have formulated the question differently because the same question can be asked to a theory about the neural correlates of consciousness or experience, as well as to a theory such as SMEn that aims at claiming that the vehicles of perceptual experience are distributed across brain, body, and world.

with the object of perception. According to Myin & Loughlin, the merit of SMEn is that it provides a way of analysing and accounting for "any qualitative experience".

Myin (2016) runs a similar argument. He rehearses Prinz's (2006, 2009) worry that referring to the interactions of the whole organism fails at addressing the explanatory gap in the same way as referring to neural processes does. After all, regardless of whether we refer to an internal or to an extended process, in both cases there is an unbridgeable gap between those objective processes and the subjective character of consciousness. But, for Myin, not accepting that the explanatory gap can be bridged is the remnant of a commitment to a dualist view of the mind, a view that is rejected by SMEn (Myin, 2016, p. 86). The explanatory gap is not an ontological problem. Instead, it should be understood as an epistemological problem, one that can be tackled more plausibly by referring not only to the stable patterns of neural correlates of "phenomenal feels", but by referring to the richer and complex dynamic interactions between the agent and the environment. That a certain neural profile can be accompanied by some other phenomenal feel or by no phenomenal feel is conceivable because too much is missing from the story (Hurley, 1998, Chapter 8 as quoted by Myin, 2016).

I am not interested in discussing whether SMEn succeeds in bridging the explanatory gap. Instead, I am interested in noting that the patterns of interaction between the perceiver and the environment are what are supposed to be doing the work required to bridge the explanatory gap. The thesis is that, by examining and modelling the interactions of the perceiver, it becomes clear that some sensorimotor interactions have the phenomenal quality they have. As I discuss in the next section, there is a weaker version of this claim. Rather than claiming that SMCs can fully explain the character of perceptual experience—bridging thus the explanatory gap—I will commit to a weaker claim, namely that SMCs can partly explain the character of perceptual experience.

3.1.2 Enacting content

As mentioned, one of the puzzling features of perception is that the experience we have somehow outstrips the stimuli that our senses receive at a time. Moreover, perceptual experience is such that whatever we perceive we do so, necessarily, from a specific angle, and at a given time and location. Regardless, by means of this process we come to know something about the world. From how objects appear to us in perceptual experience, we learn about certain properties they have. For instance, from the size a tree appears to have, we learn about its actual size. Susanna Schellenberg puts this as follows: "one always perceives objects from a particular location, but none the less can perceive their intrinsic properties" (Schellenberg, 2007, p. 605). So, how is it that in perceptual experience we go from properties that depend upon the perceiver's perspective to properties that pertain to the object regardless of its relation to other objects, what Schellenberg calls intrinsic properties?

In the previous chapters, I rehearsed SMEn's response to this question. According to this view, we come to know about objects' intrinsic properties on the basis of our possession and execution of sensorimotor knowledge. In section 1.2, I claimed that sensorimotor knowledge was knowledge of the way sensory input changes after an interaction. But, how are we meant to understand this claim? To get us started, let us say that sensorimotor knowledge is knowledge of the way an object's appearance changes after an interaction. This constitutes SMEn's core thesis: the character of perceptual experience—its quality and content—is partly fixed by the lawful covariations between movements and sensory stimuli, as well as by perceivers' practical knowledge of these covariations, namely SMCs.

SMCs fix the character of perception. According to the strong version of the view, SMCs and the knowledge thereof can fully account for the explanatory gaps. As I mentioned, we can weaken this claim. The weak version aims at defending, instead, that the pattern of SMCs and the knowledge thereof is partly determinant of the intentional content of perceptual experience (Kiverstein, 2010). Accordingly, SMCs are part of what fixes the character of perception. That a perceptual episode has a certain character depends on SMCs, since these describe the way in which sensations unfold over the course of an interaction, bringing about a given perceptual experience. This is the first requirement mentioned above. It is necessary, but not sufficient, to account for perceptual experience. In addition to this, perceptual experience depends on perceivers' expectations manifested in the features of the experience. This is the second requirement mentioned above. Furthermore, in connection to this, it is important to note that perceivers' possession and execution of this practical knowledge depends on those perceivers' history of perceptual learning and habituation. This latter point will become important in section 3.2, when we discuss the varieties of SMCs. Let us turn, for now, to a more detailed discussion of SMEn's theory of the content of experience.

The dual character of perceptual experience

So, what is it that we perceive? When Noë (2004) discusses this point, he begins by stepping away from the sense data theory. So, let me start there. Very roughly, according to the sense data theory, what is given in experience is not strictly the world. For instance, we do not perceive the tilted coin held in front of our face, but instead an image or representation that captures the appearance of the coin, e.g. its oval shape. This view receives support from the argument from illusion. Given the possibility that we might be imagining, dreaming or hallucinating that coin, we cannot conclude that what is given in perception truly is the coin. Since there are other non-veridical experiences that share the content of that very same perceptual experience, if we wish to take experience at face value, we just cannot claim that in perception we experience the coin. Instead, we must claim that we perceptually experience the image of a coin. To claim otherwise is to misrepresent what is given in perceptual experience.[40]

The claim that experience is here taken "at face value" can, of course, be contested. For instance, Noë refers to P. F. Strawson's (1979) objection to the sense data theory. For Strawson, the sense data theory takes a step back from experience. Instead, taking experience "at face value" is to claim that our perceptual experience is of the tilted coin, the coin that is out there in the world, held right in front of our face. So, it is sense data theorists who misrepresent experience by claiming that we experience an *image* of the coin.

Noë's (2004) proposal stands in between these two views. Of course, he agrees with Strawson in that it is the world that is presented to us in experience, rather than some construct. However, there is something right about the view advanced by sense data theorists. The core idea of this theory is that "perceiving is a way of finding out how things are from how they look or sound or, more generally, appear" (Noë, 2004, p. 81). For Noë, "appearances are perceptually basic". The best way to make sense of this is to claim that perception has two aspects or moments. One moment corresponds to what the sense data theory identifies, namely appearances. But the second moment corresponds to what is identified by those who reject this view, namely the fact that in perceptual experience we are presented with the world. According to Noë, "[w]e find out about appearances,

[40] As Noë notes, the argument from illusion depends on the assumption that non-veridical experiences and veridical experiences of e.g. a tomato are indistinguishable. In connection to this point, Farid Masrour (2019) for instance has argued that whether there are hallucinations that are indistinguishable from perceptual experiences is strictly speaking an empirical question, one that remains open, but that is not currently supported by experimental research.

and in finding out about appearances we find out about (come into contact with) the world" (Noë, 2004, p. 82).

That perceptual experience has a dual character has been noted by e.g. Christopher Peacocke (1983), who distinguishes between the representational content of perception that consists of that which is presented in experience—e.g. the coin—and the qualitative features of the experience—i.e. the oval shape of the tilted coin. According to this view, appearances are not part of the representational content. Instead, they are disassociated from it. They cannot both belong to the content of experience because they are conflicting properties. The content of our experience can either represent the coin as circular or as oval, but not as both. Moreover, given that perceptual experience can be assessed in terms of its accuracy (either it is veridical and presents the world as it is, or it does not), the content of this experience should give us the conditions of this accuracy and refer to the way the coin really is (and it really is circular).

Noë agrees with Peacocke in that the character of perceptual experience has a dual character. However, he rejects the claim that the two aspects can be disassociated in this way. Both constitute the character of our experience. Disassociating them is not necessary because these are not conflicting properties. The oval shape of a circular coin is not its shape in the same way as its circularity is. While the former is its apparent or perspectival shape, the latter is its actual shape. Let me say more about perspectival properties.

P-properties

Perspectival properties (or p-properties) are a function of the location from which the object is perceived and the physical properties of the object. They have two dimensions: (a) a factual dimension; and (b) a perspectival dimension, that is, how the aspect of the object would vary if the perceiver or the object perceived were to move.

P-properties are mind-independent. They are neither sensations nor feelings, nor do they depend on them (Noë, 2004, p. 83). Noë claims that "P-properties are properties of the environment" which "are only available (visible, audible, etc.) to creatures with the right kind of sensorimotor apparatus (e.g., the right kind of bodies). Insofar as they are there *for us*, it is only owing to our natures as agents with sensorimotor skills" (Noë, 2004, p. 84). So, the coin has a certain set of physical features in virtue of which it appears the way it does. For instance, it has a certain shape in virtue of which it occupies a certain space; it is made out of metal and reflects light accordingly; the coin has also been engraved, so it has a certain texture and markings that are visible under certain light conditions. These are all intrinsic properties of the coin. We, as perceivers, have a certain kind of body, we stand erect on our legs and grasp objects with our hands. Given our typical size and the typical size of a coin, we can easily see one side of the coin in a glance. That a coin appears oval is not only a property of the object, but a relational property of the coin and the perceiver. Noë illustrates this with the case of a table. Its rectangular shape becomes apparent as one moves around it, as it appears trapezoidal from one location and rectangular from another. In line with this, O'Regan & Noë (2001) claim that these are not facts about the perceiver's "raw feelings", but facts about the interaction of the perceiver with her environment: "[t]hey pertain, rather, to the person's (or animal's) active engagement with the world [s]he or it inhabits" (O'Regan & Noë, 2001, p. 965).

For Noë, P-properties are variant, relational properties of objects. We find out about the invariant features of the objects of perception by the way these variant properties unfold. These properties vary in relation to movements as one interacts with the world. As Kiverstein (2010) notes, Noë describes p-properties "in terms of two-dimensional colored surfaces". But this cannot be all, since we do not just experience objects' spatiality in terms of a two-dimensional perspective. Unlike Noë, Kiverstein proposes to describe P-properties more broadly as relational properties that depend on a given spatial point of view or perspective, as well as on the intrinsic features of the objects of perception. Furthermore, neither p-properties nor invariant properties take part of an inferential process in which we go from the former to the latter. One does not infer that the table is rectangular given its p-properties or that the coin is circular given its p-property of appearing oval. Instead, they are just part of our experience. Experience so described acquires content in virtue of the unfolding interactions with the table and the coin. Sensorimotor skills are deployed in these interactions. Sensorimotor knowledge would, then, be knowledge of the lawful way p-properties vary or unfold over time and how they correlate to objects' invariant properties. SMCs are these lawful variations.

The role of sensorimotor contingencies in fixing the character of perceptual experience

According to Kiverstein (2010), SMCs pertain to the machinery of perceptual experience. However, sensory information that constitutes SMCs does not refer only to proximal stimuli. This is, of course, part of it. Kiverstein draws on James J. Gibson's (1972) distinction between optical stimulation and optical information. As perceivers familiarize with the way their senses are impacted by stimuli and how this correlates with different objects perceived under different conditions, the information received by perceivers goes from mere optical stimulation to optical information. When it comes to optical information, proximal stimuli say something about perceivers' surroundings in virtue of the way stimuli vary in connection to other sense modalities and to movement.

To illustrate this, think of an example presented by Noë of what he calls experiential blindness (Noë, 2004, p. 5). These are cases in which the subject reports not seeing anything, without there being any "recognizable sensory deficit" (Gangopadhyay, 2010). The example concerns patients who live with blindness due to cataracts, which obstruct the passage of light to the retina. Noë notes that one would expect the removal of the cataract to be similar to removing a blind or cover; that, after the surgery, patients should be able to see without a problem. Instead, patients report receiving stimuli—they see a blur, and they detect light, colour and movement—without being able to make sense of what they see. When describing these patients, Gregory & Wallace (1963) observe that patients keep searching eye movements to the minimum and their saccades are atypical ((Gregory & Wallace, 1963) as quoted by Gangopadhyay (2010, p. 400)). In this way, cataract patients do not manifest possessing sensorimotor knowledge that would allow them to skilfully explore the environment.[41]

Not only do SMCs pertain to the machinery of perceptual experience, they contribute as well to its character in that they might be manifested in the form of sensorimotor expectations. Practical knowledge of SMCs not only enables the skillful exploration of the environment; it also plays a role in accounting for perceptual presence. To explain how, let me turn to Kiverstein's (2010) treatment of the foreground and the background of perceptual experience.

[41] This point relates to the distinction between perception, perceptual experience and sensation. In the next chapter, I discuss the problem of unconscious perception as it arises in connection to the ventral and dorsal systems. Here, I just want to note that although I have used the terms perception and perceptual experience interchangeably (and will continue to do so), this might be problematic precisely because of cases of unconscious perception. According to the case of experiential blindness, there are cases of sensation or sensory stimuli that are not accompanied by experience. These cases, we might argue, are not strictly cases of perception. Not every case that involves sensory stimuli amounts to perception. However, there are cases in which, even when no experience is reported, subjects' behavior is consistent with the processing of e.g. visual information. In this case, it becomes more difficult to claim that, given the lack of experience, this does not amount to perception.

The foreground of experience is comprised by “whatever objects and properties are, at the time, the targets of a perceiver’s focal attention” (Kiverstein, 2010). In the foreground of experience, objects’ intrinsic properties are presented from a perspective. P-properties are just the way intrinsic properties appear in the foreground of experience. Think back to the Marilyns wallpaper. For starters, given certain features of our body: our height, position, etc., the wall will likely be far larger than we are. So, we can only see a certain section of the wall, depending on the distance at which we are standing from it. Moreover, we only attend to certain sections of it at a time. And, given the limitations of our the visual system, that to which we attend will depend on what falls within the foveal area (Kiverstein, 2010). That all of Marilyn’s faces are of the same size appears in my experience in terms of p-properties. But, why?

To answer this, let me now turn to the background of experience. As discussed before, we do not have everything that pertains to a scene all at once, at a glance. Let me put this in terms of what we had called the foreground of experience: the elements that pertain to the foreground of experience are not all that appears in experience. That is precisely the puzzling feature of perceptual experience to which we have referred time and again. Many other details and aspects of a scene—that are not currently the objects of my focal attention—enjoy a certain degree of presence: e.g. the other faces of Marilyn do not just disappear from our experience whenever we are not directing our glance at them. This is what, according to Kiverstein, constitutes the background of experience. The sense of presence of this background is due to perceivers’ practical knowledge that enable potentially accessing these objects. And, as we have discussed before, this knowledge is shaped by the lawful patterns in which sensory information follows an interaction, namely SMCs.

In the same way as the background of experience appears in virtue of our sensorimotor expectations, the intrinsic properties of the objects that appear in the foreground are explained by our sensorimotor expectations. The intrinsic properties attributed to an object depend on its p-properties. P-properties serve as the basis for the attribution of intrinsic properties. According to Kiverstein, intrinsic properties are similar to the objects that pertain to the background of experience in that “they exceed what we can perceive in a single encounter” (Kiverstein, 2010). And, for that reason, they also depend on sensorimotor expectations.

Direct realism

One last point worth noting about SMEn's theory of perceptual content relates to its commitment to direct realism (Beaton, 2016), according to which perceptual experience is constituted by an unmediated subject-object relation. As can be seen from the previous discussions, one worry that constantly looms over SMEn is the concern that it might fall into representationalism. SMEn belongs to what have been called relationalist accounts of perceptual experience.[42] Unlike some versions of representationalism regarding perceptual experience, according to which perceptual experiences are just representations, relationalism claims that perceptual experiences are, instead, the relation between the perceiver and the object of perception (Nanay, 2015, p. 321).

To avoid collapsing into a representationalist view, according to which the content of perception are representations of, let us say, counterfactual sensorimotor possibilities, SMEn must be ready to claim that both the object and the specific perspective of the perceiver are constituents of perception (Campbell, 2008).[43] It is important to note that I take direct realism to be compatible with the claim that representations might enable perceptual experience. The claim would be that even when the perceiver is in direct contact with the world, this contact might be enabled by mechanisms that can be best described as representations. What becomes relevant here is again the distinction between a causal account of the mechanisms that enable perceptual experience, and the view of what constitutes perceptual experience and its content.

Let me briefly summarize here what I have discussed so far. SMEn aims at doing justice to perceptual experience. The best way to do so is by referring to perceivers' engagement with the world. Perception acquires content by means of these interactions. Moreover, sensory stimuli start making sense because they unfold in a patterned way depending on different movements and certain aspects of the position undertaken by the perceiver. This is the way SMCs fix, at least in

[42] For a discussion, see (Berger & Nanay, 2016; Nanay, 2015).

[43] The main challenge that remains is, of course, cases of non-veridical experiences that are akin to perceptual experience: e.g. cases of hallucination, imagination, dreams, etc. Beaton, for instance, proposes to claim that, in these cases, subjects are disposed to deploy certain skills associated with the object (Beaton, 2016, p. 271). In relation to this, Masrour (2019), for instance, argues that there is no conclusive evidence for the existence of hallucinations that are just perceptual experiences without mind-independent objects. Although not defending direct realism directly, Masrour aims at blocking the argument from hallucination by claiming that there is no conclusive evidence for the existence of hallucinations without objects.

part, the content of perceptual experience. SMCs are just the patterned ways in which sensory information unfolds, knowledge of which enables perceptual experience. Despite the crucial role SMCs play in SMEn, more is needed to understand exactly what these lawful patterns truly are.

3.2 The theory of Sensorimotor Contingencies

As discussed at the beginning of this chapter, one of the goals of SMEn is to offer an account that can bridge the explanatory gap or that can, at least, provide an account of the way sensorimotor interactions fix the character of perception. The strong version of SMEn is committed to the view that referring to sensorimotor interactions permits fully accounting for the character of perception. However, as I mentioned before, I am interested in the weaker version of SMEn, the claim that referring to these interactions is necessary to an account of the character of perception. Moreover, the view is not just that the interactions between the perceiver and the environment are necessary. Instead, the view is that what plays this important explanatory role are the lawlike patterns of sensorimotor interactions: SMCs.

3.2.1 Sensorimotor Contingencies

SMCs are the *systematic patterns of co-variation between sensory information and possible interactions between the perceiver and the object(s) of perception*.[44] Perhaps the best way to understand what the theory refers to as SMCs is the following example, which will become relevant in chapter 4, when discussing the chauvinism concerns.

O'Regan & Noë (2001) take the case of the visual system, in particular, the patterns of stimulation of the retina, to illustrate their view. These patterns of stimulation depend in part on the shape and size of the eye, and on the shape of the retina itself. For instance, when the eye is presented with a straight horizontal line, the stimulation is such that what is reflected in the retina is an arch that is distorted as the eye moves. In a similar fashion, given the different concentrations of photoreceptors in the retina, the information received changes as the eye moves. In the presence of a specific kind of object and the movement of the eye, a pattern is established given the resulting

[44] SMCs are also linked to the phenomenality plot, a model that allows the analysis of different experiential mental states based on certain features of the input that is related to neural activity associated with these states (O'Regan et al., 2004, 2005a, 2005b).

retinal stimulation. In line with this, O'Regan et al. claim that "[t]he incoming data concerning a fixated patch of color depend on eye position. Because of non-uniformities in macular pigment [a pigmented spot in the retina] and retinal cone distributions, eye movements provoke different patterns of change in sensory input" (O'Regan, Myin, & Noë, 2005b, p. 60).

According to this example, the notion of SMCs refers to the variation in proximal stimuli that follows the movements of the perceiver. There is a distinction that is important to introduce at this point: the distinction between (i) movement induced regularities and (ii) object induced regularities. The former are those regularities described by the previous example. The pattern of interactions derives from the structure of the visual apparatus: the perceiver explores the object of perception by means of this very specific organ. The relevance of the structure of the visual apparatus becomes central for the specification of the exploration that is available to the agent and the sensory information that the agent will obtain from it. The latter, object SMCs, are the pattern of sensorimotor interactions determined by the intrinsic properties of the objects of perception. O'Regan & Noë (2001) mention, for instance, the shape, texture and size of the object. These attributes determine the pattern of interactions available to the agent. The pattern of interactions available for the perceiver is a function of the structure of the perceiver and the structure of the object. For this reason, the attributes of the object that are available to the perceiver depend, ultimately, on the kind of exploration that is enabled by the perceiver's own physiological structure and the structure of the object. We can illustrate this point when thinking of the differences between visual and tactual perception. Vision and touch allow us to explore similar aspects of the objects of perception: e.g. the shape of an object. However, they are different ways of sampling the shape of the object. Importantly, this does not mean that the difference is radical. As the discussion on sensory substitution devices in section 3.2.2 will show, the patterns of SMCs of vision and touch do not differ radically.

Something important to note here relates to the question of whether SMCs describe changes between proximal stimuli or distal stimuli and movements. As I mentioned before, as perceivers familiarize themselves with changes in stimuli, over time, SMCs are enriched. SMCs describe, then, not just the patterns of correlation between, e.g. retinal image and movement, but patterns between distal stimuli and movement. In relation to this, O'Regan & Noë distinguish between SMCs that depend on the structure of the e.g. the visual system and those that depend on the features of objects and the environment. About the latter, they claim that SMCs are the “[result] of categorisation of objects and events in the environment” (O’Regan & Noë, 2001, p. 943). These are SMCs associated to perception. On the other hand, the former, SMCs that depend on the structure of e.g. the visual system, are “independent from any categorisation” of objects and events. These are SMCs associated, in turn, with sensation. On the basis of this distinction, we can explain too that it is possible to have cases in which someone processes sensory stimuli without this amounting to perception. This was the case of experiential blindness introduced before. In this case, according to Noë, while subjects might possess SMCs associated with sensation, what they are missing are sensorimotor skills.

In addition to this, the theory of SMCs is meant to explain how it is that one sensory modality is accompanied by a certain phenomenal feel and not by any other. In this way, SMEn aims at bridging what we called the intermodal comparative explanatory gap. In the next section, I introduce the case of sensory substitution to illustrate this explanatory role fulfilled by SMCs.

3.2.2 The explanatory power of Sensorimotor Contingencies: Sensory Substitution and sense modalities

SMEn is taken to explain sensory substitution, i.e. a phenomenon in which stimuli from one sense modality (e.g. tactile stimuli) is interpreted (and, according to some testimonies, experienced) as stimuli coming from another modality (e.g. visual stimuli). One example of this phenomenon are sensory substitution devices. These devices allow patients who have lost sensory capacities to substitute them “by means of another sensory modality” (Auvray & Myin, 2009). For instance, some devices allow patients with blindness to “see” or to identify visual features of the object they are interacting with. This case is of particular interest to SMEn because, unlike other cases of sensory substitution, it does not involve neural rerouting. Rather it involves altering the external relations of the patient and the objects (Hurley & Noë, 2003).

Sensory substitution devices receive input in one sense modality, e.g. vision, and translate it into the kind of input associated with another sensory modality, e.g. tactual input. The user receives stimuli coming from this other modality. For instance, Tactual Visual sensory substitution systems (TVSS) consist of a camera and vibrators or electrodes. The camera is connected to the vibrators which are placed on the subject’s back, thigh, abdomen, forehead, fingertip, or tongue. The device produces tactile stimuli that are applied on the patient's skin via the electrodes or the vibrator. Similar devices have been developed that convert visual information into auditory stimuli (Auvray & Myin, 2009). Some Auditory Visual Sensory Substitution devices (AVSS) are echolocation systems that work similarly to a sonar: they detect the location of a distant object with sound waves and transmit it to the patient with headphones. Other AVSS devices work more similarly to the TVSS. These devices convert an image into different sound frequencies (Auvray & Myin, 2009, p. 1039). Similar devices have been developed to allow patients with visual impairments to read.

Although at the beginning experience corresponds to the sense modality received by the user—in TVSS devices, for instance, the experience is tactual—the subject quickly adapts and reports receiving visual information (Hurley & Noë, 2003). Furthermore, subjects succeed at tasks that require, for instance, spatial information (e.g. localization tasks and shape recognition). Some devices even recreate visual illusions (Hurley & Noë, 2003). Importantly, in order for the subject to adapt and perceive tactual stimuli as visual, it is essential for the subject to be able to actively interact with the object, i.e. to be able to explore it. As reported by subjects, experience with TVSS devices is more like vision than it is like touch: "TVSS subjects report experiencing characteristically visual effects and illusions, such as parallax, perspective, looming, zooming and depth effects, and the waterfall illusion" (Bach-y-Rita, 1996; Bach-y-Rita et al., 1969 as referred to in Hurley & Noë, 2003, p. 144).

What explains these results? According to SMEn, in TVSS cases, SMCs associated with vision are preserved. An example advanced by Hurley & Noë is that in TVSS cases an object can be brought to view by moving in a way that is similar to visual cases: "both vision and TVSS-perception are governed by laws of occlusion for which there are no good analogues in touch or audition" (Hurley & Noë, 2003). What allows patients to perceive and detect visual features are SMCs that govern the interactions with the object. Given that TVSS devices, for instance, allow similar interactions with the object that are ordinarily characteristic of biological vision, they result in an experience that is visual in some sense: it enables the perceiver to detect certain invariant features of the object (e.g. its size), certain illusions are preserved, etc. A similar case can be made for AVSS devices. Some features of the auditory system can be exploited to obtain vision-like patterns of sensorimotor interactions because the auditory system is able to e.g. deal with complex environments, with rapid changes, and it is able to make fine discriminations (Auvray & Myin, 2009, p. 1039).

Sensory substitution devices support the thesis that SMCs fix, at least in part, the character of perceptual experience. In these cases, one sensory modality deploys SMCs in such a way as to give the agent the possibility of accessing information that pertains to another sensory modality.

So, let me bring together some of the ideas discussed so far. SMCs are descriptions of the correlation between sensory information and movement, broadly construed. They depend, on the one hand, on features of the relevant sensory system, e.g. the visual system; and, on the other hand, on features of the objects of perception, e.g. the space they occupy, and on features of the environment, e.g. the distance between two objects. The co-variation between these elements follows regular patterns that are due to the structure of sensory systems and the objects of perception. Sensorimotor knowledge is the implicit practical knowledge of the lawful patterns that constitute SMCs, and the possession and exercise of which is constitutive of perceptual experience. Sensorimotor knowledge is strengthened in virtue of certain regularities that the perceiver finds in the environment, in virtue of the interactions that are habitual for the perceiver, and in virtue of the perceiver's own interests (Di Paolo et al., 2014). Finally, this knowledge manifests flexibility. As can be seen in the case of sensory substitution devices, perceivers can incorporate certain patterns of interaction into their repertoire, adapting in this way to their changing environment, as well as to changing features of their own body. Of course, how quickly they can adapt and learn, and which modifications can be easily incorporated or can turn out to be more challenging, will depend on the features of each case. The quick adaptation between tactile stimuli and visual stimuli is, in part, explained by the similarity between the patterns of SMCs that pertain to each modality.

The theory of SMCs has required further refinement. This is, in part, due to the ambiguity in the treatment of SMCs in the literature (Vernazzani, 2019). Moreover, as I mentioned before, it seems that the patterns of interactions that are incorporated and executed by individuals depend not just on features of the objects with which they interact and features of their sensory systems. Instead, they also depend on perceivers' sensorimotor habits, and the goals pursued at each moment. Furthermore, the regularities that are described by SMCs depend on features and details of the perceiver's body: for instance, they depend on fine-grained details of the visual system, as well as on other details such as the different ways perceivers can position their bodies. However, it is not entirely clear what the details are that matter. Moreover, SMEn is an embodied theory of perception according to which perception is the deployment of embodied skills. According to this, the character of perceptual experience depends on the actions that can be skilfully deployed by a perceiver with a certain body. In the last section of this chapter, I say more about this claim. Before that, I briefly review the dynamic theory of SMCs advanced by Buhrmann et al. (2013). They propose to distinguish four kinds of SMCs that aim at remedying the ambiguity of previous proposals.

3.2.3 Varieties of SMCs

Burhmann et al. (2013) have developed a dynamic model that operationalizes the notion of SMCs.[45] This grants SMCs another explanatory role. SMCs allow the modelling of the agent-world system on the basis of which SMEn explains perception. Modelling the perceiver's sensorimotor interactions allows the prediction of future states of the agent-world system (Vernazzani, 2019). This proposal is advanced to remedy the ambiguity of the notion of SMCs as they were found in the literature. The authors distinguish between four kinds of SMCs: (i) the *sensorimotor environment*, (ii) the *sensorimotor habitat*, (iii) the *sensorimotor coordination*, and (iv) the *sensorimotor strategies*. These are all descriptions of sensorimotor regularities, i.e. of regular patterns of interactions. The different kinds of SMCs are captured by differential equations that consider different variables, such as sensory dynamics, dynamics internal to the system, and body dynamics. Each of these descriptions delivers a set of possible sensorimotor trajectories or interactions that are available to a system comprised by a perceiver with certain features (e.g. the possession of a visual system) and by an environment with certain other features (e.g. objects located at certain distance, or with certain textures, colours, shapes, etc.). In what follows I describe these different kinds of SMCs in more detail.

Sensorimotor Environment

Let us start with what would be the more general description. The *sensorimotor environment* takes into account (a) the movements of the system in question, and (b) the sensory consequences of

[45] See also (Di Paolo et al., 2017). While Buhrmann et al. (2013) claim that the dynamical model of SMCs they propose is neutral and can be made compatible with different frameworks, Di Paolo et al. incorporate the model to a proposal that subscribes to a version of autopoietic enactivism.

these movements. The description of the sensorimotor environment delivers a very general description of SMCs and is the notion of SMCs that appears more commonly in the literature. One example is the one introduced above about the visual system and the horizontal line. The model illustrates the possible interactions of the perceiver as a function of sensory values given motor states.

Here, the body is described, ultimately, in terms of motor possibilities afforded by the morphological structure of the agent. The sensorimotor environment delivers a set of possible sensorimotor interactions that are available to an agent in virtue of specific physical features (e.g. "rotations of the head [that] lead to a lawful change in the temporal asynchrony between sounds received in the left and right ear" (Buhrmann et al., 2013). To illustrate one of the trajectories that would be modelled by the sensorimotor environment, Buhrmann et al. advance the case of a sponge being squeezed and the perception of softness that results from this. The sensorimotor environment describes how different movements relate to the resulting sensory input felt by the agent, e.g. the pressure felt on the tip of the fingers squeezing the sponge, and the way this pressure increases as the grip closes and how it decreases as it opens (Buhrmann et al., 2013).

Sensorimotor Habitat

The *sensorimotor habitat*, in turn, takes into account the same variables as the sensorimotor environment, namely (a) the movements of the agent and (b) its sensory consequences, plus (c) the evolution of the internal dynamics of the agent. In consequence, the set of possible sensorimotor trajectories described by the sensorimotor habitat is more constrained than the set described by the sensorimotor environment. Importantly, the sensorimotor habitat is not simply a more specific version of the sensorimotor environment. Although the trajectories contained in the habitat of a system appear in the environment of said system, the sensorimotor habitat aims at capturing the active engagement of the system since it incorporates how a system interacts with a given surrounding (e.g. the trajectories that it tends to visit) and it considers how the internal states of the system evolve over time. While the sensorimotor environment presents all the possible trajectories available given the variables considered, the sensorimotor habitat takes a temporal perspective by considering the coupling of an agent with the environment. Taking into account (c) the internal dynamics of the perceiver implies that the model reflects regularities and constraints that did not appear as such in the sensorimotor environment.

Sensorimotor Coordination

The *sensorimotor coordination* considers the three variables included in the sensorimotor habitat, namely (a) the movements of the agent, (b) its sensory consequences, and (c) the internal dynamics of the agent, but it constrains (d) the trajectories in virtue of their stability and reliability given the performance of the system in a specific task or in light of a specific goal. In the case of the sensorimotor coordination, what becomes important is the functionality of a certain sensorimotor trajectory, that is, whether it contributes to the fulfilment of the task, e.g. discriminating a couple of sponges based on their hardness (Buhrmann et al., 2013). The sensorimotor trajectories that are modelled here are those that contribute to the fulfilment of the task: e.g. rhythmically pressing the sponge.

Sensorimotor Strategies

Finally, *sensorimotor strategies* or *schemes* take into account the variables considered by the sensorimotor coordination, namely (a) the movements of the agent, (b) its sensory consequences, (c) the internal dynamics of the agent, and (d) trajectories relevant to the performance of a task. However, the sensorimotor trajectories that are modelled in these schemes are constrained in light not only of a task and goal, but also in light of their effectivity, since these are the trajectories prioritized by the agent. In that way, this model incorporates a dimension of normativity. Sensorimotor strategies are the trajectories comprised in the model of sensorimotor coordination that have proven to be the most effective or that have been assessed as preferable by the perceiver (Buhrmann et al., 2013).

According to the authors, the notion of sensorimotor strategies is meant to account for the mastery of SMCs that is constitutive of perceptual experience. The authors illustrate this with the case of experiential blindness. That the patients who recently went through cataract surgery cannot perceive the face of the doctor is explained by a lack of mastery of the sensorimotor trajectories that would allow them to e.g. move their eyes at the right speed. The thought would be, then, that perceivers reinforce the trajectories that bring about the relevant information about the environment. It is by means of perceptual learning and by reinforcing certain sensorimotor trajectories that perceivers can move from proximal stimuli to distal stimuli, and from perspectival properties to invariant properties and expectations about the background of experience.

As I mentioned before, this is not something each of our sensory systems does in isolation. In traversing certain sensorimotor trajectories—those modelled by each of the interpretations introduced above—the whole body of the perceiver is involved: e.g. we move our heads from side to side and stretch our neck to get a glimpse of the cat behind the fence; we dispose to grab the round tomato with our hand ready to surround it with each of our fingers; we move our hand rhythmically to see how soft the stress ball is; we squeeze our eyes and lower our heads trying to get a better focus of the document we are editing in the computer.

3.2.4 SMEn's embodiment claim

As seen in the examples just introduced, the skills that constitute perception are *embodied* skills. These are skills that involve the whole agent. Moreover, the skills deployed by the perceiver are shaped by her body, since SMCs are agent-centred. SMCs are systematic patterns of co-variations between movements of the perceiver and sensory input. This is SMEn's *embodiment claim*. Think back to O'Regan & Noë's (2001) illustration of the retinal image and the horizontal line. In this case, the patterns of stimulation of the retina depend partly on the shape and size of the eye, and on the shape of the retina itself. In this case, the stimulation is such that what is reflected in the retina is an arch that is distorted as the eye moves. So, in the presence of a specific kind of object and the movement of the eye, a pattern is established given the resulting retinal stimulation. The body of the perceiver becomes a crucial determinant factor of the set of SMCs that are constitutive of e.g. seeing a horizontal line (O'Regan et al., 2005b, p. 60). Note, however, that if we consider the four kinds of SMCs introduced above, the physiological details that matter for SMCs are not just e.g. details of the sensory organs. SMCs depend on the position of the perceiver's body, on the movements and interactions prioritized by the perceiver and on the way these movements are performed.

Consequently, embodiment becomes a central feature of perception: the character of perception depends on the body of the perceiver. Noë (2004) makes this feature of SMEn more explicit. He claims that the skills, whose exercise is constitutive of perception, are bodily skills. And, for this reason, perceptual experience depends in part on the possession of the kind of body that could exercise those skills (Noë, 2004, p. 25).

The embodiment claim has, ultimately, a consequence for the model of explanation adopted by SMEn. Perception is embodied because an explanation of the character of perceptual experience—i.e. typically a phenomenon at the psychological level of explanation—incorporates details that concern the material basis of perceptual experience—i.e. typically a phenomenon at the implementational level of explanation. In other words, whichever theory is advanced, it cannot be formulated independently from a theory about the material constitution of the perceiver.

The dependence of the character of perceptual experience on details of the body has led to the concern that SMEn is a chauvinistic theory of perception. In connection to that worry, Noë claims that:

> In virtue of *what* can the visual systems of the human, the crab, and the bumblebee all be deemed *visual* systems? Is it because they all give rise to the same kind of experiences, qualitatively speaking? How far must we commit to accepting that? According to the enactive view, there is sufficient high-level, gross sensorimotor isomorphism between these different perceptual systems to count one and all as *visual* systems. There are also very many low-level sensorimotor differences, and these differences are sufficient to make the visual "experience" of the bee and the person different. But isn't this just the consequence we want? (Noë, 2004, p. 113)

According to Noë, SMEn achieves the following. Firstly, by means of the set of SMCs, SMEn gives an account of what makes two episodes of perceptual experience episodes of the same kind. We can add here that SMEn explains why sensory substitution devices succeed in enabling experiences that are akin to the target sensory modality (e.g. TVSS systems enable something akin to visual perception). Moreover, it offers an account of why the quality of our experiences might differ across individuals and species. For Noë, given this, if SMEn is chauvinistic, it is chauvinistic “to an acceptable degree” (Noë, 2004, p. 113). In the next chapter, I articulate and evaluate this worry.

Chapter 4. Chauvinism and embodiment

In chapter 3, I discussed the proposal that according to a weak version of Sensorimotor Enactivism (SMEn), sensorimotor contingencies (SMCs) contribute to the character of perceptual experience in two ways. Firstly, the possession and exercise of sensorimotor knowledge explains the sense of presence enjoyed by objects and features that pertain to the background of experience. Secondly, this knowledge explains too that we perceive the invariant features of the objects of perception that comprise the foreground, and that we do so starting off from their perspectival properties. In perception, we encounter the world directly and this encounter is enabled by implicit practical knowledge of the systematic ways in which appearances change after an interaction. Furthermore, by means of perceptual learning and by reinforcing certain sensorimotor trajectories we go from knowledge of the way proximal stimuli changes—e.g. the retinal image of the tree that enlarges as I move forward—to knowledge of the way distal stimuli changes after an interaction—e.g. the tree that is in front of me that will come closer when walking towards it.

The character of perceptual experience depends in part on SMCs, the systematic patterns of co-variation between sensory information and interactions between the perceiver and the object(s) of perception. SMCs are sensitive to the embodiment of the perceiver. They are agent-centred in that they are shaped, in part, by the physiological configuration of the perceiver. The latter partly determines the interactions that are possible for the perceiver, as well as constituting the point of view from which the perceiver receives sensory information. Consequently, the skills that constitute perception are embodied skills. It is in line with this idea that Noë claims that "[t]o perceive like us (…) you must have a body like ours" (Noë, 2004, p. 25).

This view, however, could have unpalatable consequences. We might wonder 'how much a creature needs to have a body like ours in order to, e.g., visually perceive in the way we do?'. The concern is that the dependence of the profile of SMCs on the body might invite a kind of chauvinism. By chauvinism, I mean simply that, under SMEn's account of perception, the dependence of the character of perception on the body of the perceiver is such that SMEn cannot extend the attributions of character and sense modality to all the cases of perceptual experience it would under other considerations. And, consequently, the theory of SMCs struggles to fulfil its explanatory purposes. Throughout this section, I finesse this definition. For my purposes, at this point, this should do to understand chapter’s 3 conclusion. There, I introduced a quote by Noë (2004) where he claimed that perhaps it is true that SMEn invites chauvinism, but it does so to an acceptable degree: i.e. to the degree that leads to the conclusion that two perceptual occurrences are of the same kind, e.g. bat vision and human vision are both instances of vision, while maintaining that they still enjoy important phenomenal differences.

In this chapter, I address the concern of chauvinism. I argue that two different chauvinism worries should be distinguished: firstly, the *content chauvinism* concern, a concern related to the dependence of SMCs on the body and their role in fixing the character of perceptual experience; and, secondly, the *sense modality chauvinism* concern, related, in turn, to the claim that sense modalities are individuated by referring to the associated SMCs. In section 4.1, I review the concern as it was firstly introduced by Clark (2002, 2008) and Clark & Toribio (2001). I then move on to reviewing David Milner & Melvyn Goodale’s model of the two visual systems, since Clark (2008) further develops the content chauvinism concern in light of this model. Milner & Goodale’s model conflicts in important respects with SMEn’s proposal, specifically their claim that visual consciousness does not have a perspectival aspect to it, at least not in the sense of ‘perspectival aspect’ articulated by SMEn. In this section, I argue that the proposal of SMCs developed in chapter 3 does not require a commitment to the chauvinistic claim, namely the claim that every low-level detail of the perceiver’s body makes a difference to the profile of SMCs. Moreover, there is recent evidence that might serve to support SMEn’s dual content view and that puts pressure on Milner & Goodale’s model.

In section 4.2, I review the *sense modality chauvinism* concern. In the literature, the *content chauvinism* concern has been addressed by referring to cases of sensory substitution devices. According to these responses, SMEn is not chauvinistic because it can identify that TVSS devices enable an experience that is visual or quasi visual. I argue that, when the *content chauvinism* concern is addressed in this way, there is a change of conversation since it concerns a different although related concern. This is what I call the concern of *sense modality chauvinism.* To tackle this concern, I argue that the four dynamical models of SMCs, introduced in the previous chapter, are structural descriptions of perceptual systems. Furthermore, one of these descriptions, one framed in terms of sensorimotor coordination, can help us tackle this concern.

4.1 Content chauvinism

When discussing the active character of perception in section 1.2, I mentioned that the active character of perception is often thought to imply an instrumental relation between action and perception, i.e. by taking perception to be active in the sense that perception is *for* action. Recall too that, according to Hurley (2001), the claim about the instrumental relation between action and perception is among the presuppositions of the sandwich view of the mind, one that SMEn rejects. SMEn's view is not that perception is *for* action (Noë, 2010, p. 249). Instead, as discussed in chapter 1, the active character of perception should be understood constitutively, i.e. perception is a doing, it is itself an interaction: it is the exercise of sensorimotor capacities, a temporally rich phenomenon, that involves real-time anticipation and adaptation. Regardless, SMCs are the systematic patterns of co-variation between sensory information and movement. Information about the world is parsed in terms of the movements that can be performed by an individual with specific physiological features.

I come back to this point because the close relation between action and perception has been put into question in light of Milner & Goodale's influential model of the function of the dorsal and the ventral visual processing systems (Milner & Goodale, 2006, 2008, 2010). Simply put, Milner & Goodale's view is that these are two independent systems. While one of these systems is for movement control (the one associated with the dorsal stream or neural pathway), the other is for perception and cognition (the one associated with the ventral stream or neural pathway). In the face of this model, an immediate concern arises for SMEn, since even when one of these systems is clearly primed for action, the other is not. Instead, the ventral system processes information about objects' invariant features, independently of how these features vary with movement and of any connection to the organism's perspective. It is in part in light of Milner & Goodale's model that Clark & Toribio (2001) introduce their concern about SMEn's claim that movements specified at low-level details contribute to fixing the character of perceptual experience.[46] So, let us start by discussing Clark & Toribio's concern, before moving on to discuss in more detail how it relates to the two-visual systems model.

[46] See also (Clark, 2002, 2008).

4.1.1 Chauvinism in Sensorimotor Enactivism

Recall SMEn's embodiment claim. According to this claim, SMCs are sensitive to the embodiment of the perceiver. SMCs are the systematic patterns of co-variation between movement and sensory information. Regarding the relevance of specific features of the perceiver's body we can ask how much the body of the perceiver matters.

Clark & Toribio's (2001) original concern arises when accepting one interpretation of the embodiment claim according to which SMCs are sensitive to *every low-level* physiological feature of the agent. Let us call this the *chauvinistic claim*. For example, think back to the case of the horizontal line and the retinal image. Given that the patterns of stimulation of the retina depend on the size and shape of the retina itself, a difference in, let us say, the shape of the retina should make a difference to the profile of SMCs. According to Clark and Toribio, if we accept the chauvinistic claim, it will not be possible to attribute the same character to two occurrences of perceptual experience that pertain to individuals who differ in low-level physiological details, e.g. the shape of their ocular globe is different. For Clark & Toribio, we have no a priori reason to accept the chauvinistic claim, i.e. the claim that *every* low-level detail makes a difference to the character of perceptual experience. For instance, there is no a priori reason to accept that the fact that someone's eyes saccade faster makes a difference to the way an object appears to them.[47] The matter is, of course, empirical.

In order to tackle Clark & Toribio's concern, SMEn can abandon the chauvinistic claim in favour of adopting a weaker position. According to this weaker position, only *some* low-level physiological features contribute to fixing the character of perceptual experience. According to SMEn, the character of perceptual experience is fixed by SMCs. SMCs are perspectival by definition: they are co-relations between the movement performed by a specific agent and sensory information. Accordingly, there is a perspectival aspect to the character of perceptual experience that seems irreducible.

Nevertheless, a different problem for the weaker position arises in light of Milner & Goodale's model. The model puts pressure on the perspectival element that, according to SMEn, is present

[47] One response to the specific example advanced by Clark & Toribio is that for SMEn, SMCs are the systematic profile of co-variations, it would seem that the speed of eye saccades would not make a difference to how a feature appears from a given location.

in perceptual experience. As we will see in the next section, according to this model, the character of perception is fixed by the system associated with the ventral stream, the output of which are the invariant features of objects. In short, according to this model, the character of perception does not have a perspectival element to it.

4.1.2 The two visual systems

The division of labour across the two visual systems

The ventral and the dorsal streams are two "independent channels" that carry information from the retina to area V1 in the cerebral cortex (Milner & Goodale, 2006). When first identified, the difference between the two streams was articulated as a difference in the nature of their *input*. While the dorsal stream was thought to process input about objects' location, the ventral stream was thought to process objects' intrinsic properties such as shape, size and colour. Milner & Goodale proposed instead to distinguish the streams based on their *output* and, consequently, on the way each stream processes received information.

According to Milner & Goodale, in non-human vertebrates, independent modules have evolved to support behaviours specific to each organism. There are modules that support visuomotor control and that require visual information that allows the execution of specific complex movements, for instance, grasping an object, catching prey or avoiding obstacles. To enable the coordination of movements required for these interactions, the information available must be coded egocentrically and in real time, i.e. as the behaviour unfolds. In addition to this, primates' cortex evolution reflects the development of modules connected to cognitive and perceptual processes that require information about objects' invariant features. The information required by a visual system that is geared towards movement control is different from that which is required by a system that supports cognitive and perceptual processes (Milner & Goodale, 2006, p. 42).

According to this model, while the dorsal system processes input and transforms it into information that serves for movement guidance and control, the ventral system transforms input into representations of objects' invariant properties, as well as objects' location articulated in terms of allocentric coordinates (Milner & Goodale, 2008, p. 774). Note that the information that results from the processing that occurs in each system can be conflicting. For instance, one system might be deceived by a visual illusion to which the other system might be resistant (Milner, 2008, p. 172). Another important feature is that, according to this model, the dorsal system is bottom-up, i.e. it is shaped solely by the sensory input received moment by moment. In contrast, the information coded by the ventral system is shaped also by top-down influences, i.e. the representations formed are influenced by higher-order cognitive processes (Milner, 2008).

In what concerns action, Milner & Goodale claim that labour is divided across both streams. Visual information can influence action in two different ways (Milner, 2008). While the ventral stream provides information about the world that is relevant to the selection of an action, the dorsal stream provides the information required for executing it. The latter is involved in the "detailed programming and real-time control at the level of elementary movements" (Milner & Goodale, 2008, p. 776). For this, the "high-level" information construed by the ventral stream is of little use. The dorsal system "instead relies on current bottom-up information from the retina to specify the required movement parameters" (Milner & Goodale, 2008, p. 776). Another way to describe the division of labour between the two systems is to say that while the ventral system provides the information required for action planning, the dorsal system provides information required to carry on that action, i.e. for controlling and monitoring it. Importantly, the latter includes motor programming (Milner & Goodale, 2008, p. 777).

This does not mean that there is no communication between the two streams, let alone collaboration. The ventral stream is involved in the selection of target objects and action types (e.g. catch the ball). The dorsal stream, in turn, is involved in the fine-grained planning and executing of that action. However, this interaction occurs at a higher level and does not require the existence of a common code (Clark, 2008b, p. 188).[48]

Milner & Goodale's theory finds support in evidence related to patients with bilateral optic ataxia caused by damage to the dorsal stream. Optic ataxia is a deficit that is characterized by subjects' inability to visually guide their actions to reach out for an object. For instance, these subjects fail to direct their movements towards the target object and show a deficit in calibrating their grip when reaching towards it (Milner & Goodale, 2006).

[48] For a refinement of Milner & Goodale's theory, see Jeannerod & Jacob (2005) and Jacob & Jeannerod (2003).

In addition to this, the authors support their theory on the case of patient D.F., a patient with visual form agnosia, i.e. an impairment that consists in subjects' failure to recognise objects. D.F. suffers from lesions in areas that pertain to the ventral stream. Milner & Goodale characterise D.F.'s case as a deficit of *visual perception*. D.F. fails at tests that involve the recognition of geometric forms, the identification of drawings of common objects, as well as the ability to copy these drawings (Milner & Goodale, 2006, p. 127). D.F. also fails at tasks that involve the recognition of objects' dimensions and orientations (Milner & Goodale, 2008, p. 776). Now, despite D.F.'s poor performance in tasks that require visual perception so understood, her visual functions related to motor control seem intact. She performs well in tasks that involve the detection of flashes of light and visual flickers (Milner & Goodale, 2006, p. 127). Moreover, based in part on reports from D.F. about not seeing e.g. the orientation of an object, Milner & Goodale conclude that the information that is consciously available corresponds to that of the ventral stream (Milner & Goodale, 2008, p. 776). According to them, information processed by dorsal processing cannot be, in principle, consciously accessed. In line with this, Milner claims that "*the visual products of dorsal stream processing are not available to conscious awareness*; (…) they exist only as evanescent raw materials to provide moment-to-moment sensory calibration of our movements, and then are gone" (Milner, 2008, p. 173 italics in original). We might be aware of the actions and movements performed, but not of the information processed for the calibration of these movements.

So, as can be seen, Milner & Goodale's model of the two visual systems conflicts, in some important respects, with SMEn's view of perception. Let me now briefly discuss how they compare.

Sensorimotor Enactivism in the face of the two visual systems

To begin, it is worth clarifying some terminological differences. Throughout the thesis, I have referred mostly to perceptual experience using either that term or just 'perception'. Under Milner & Goodale's model, perceptual experience is relegated to information processed by the ventral system, since only this system contributes to fixing the character of perceptual experience.

In what concerns cases of non-conscious or pre-conscious perceptual processing, the two approaches differ too. Although Milner & Goodale (2008) refer mostly to cases of perceptual experience, they claim that their model should be extended to non- and pre-conscious perceptual processing. By this they mean that, in cases of e.g. unconscious perception, the content and mechanisms of these states should be comparable to that of conscious perception. For instance, cases that involve percepts that represent the size of an object, even when the agent does not have conscious access to this information, should count as cases of perception.

In connection to non-conscious perceptual processing, SMEn accepts that there are cases in which sensory stimulation does not amount to perception. O'Regan & Noë put this in terms of "stimuli impinging on the senses". This would be the case in change blindness discussed in section 1.1. Even when there are stimuli, the agent shows no sensitivity to them (O'Regan & Noë, 2001, p. 964).

Their different conceptions of perceptual experience are crucial to distinguish SMEn's model from Milner & Goodale's. According to SMEn's proposal, awareness (visual consciousness, in Milner & Goodale's terms) comes in degrees and consists in perceivers' sensitivity to sensory stimuli. If a perceiver manifests sensitivity to SMCs or exhibits mastery thereof, there is a sense in which she is aware of the relevant stimuli. Perceptual experience is that mastery.[49] Of course, this leads to a radically different interpretation of what is happening in, e.g. D.F.'s case. According to O'Regan & Noë, D.F. shows sensitivity in that she performs actions that are guided by sensory information (O'Regan & Noë, 2001, p. 969). D.F. should be considered to be aware, up to a degree, of the relevant visual information about the objects with which she is interacting.

Perhaps the most significant differences relate, firstly, to the relevance given to the perspective or point of view of the agent, and, secondly, to the fact that the systems do not require information that shares the same code.[50] For Milner & Goodale, the information that contributes to fixing the content of perceptual experience is object-centred, allocentric information. So, even if we were to

[49] Regarding unconscious perception, while SMEn claims that mastery of SMCs is necessary for perceptual experience, it is not sufficient. In connection to this, consider O'Regan & Noë's (2001, p. 960) distinction between transitive consciousness—i.e. awareness of a feature or element of a scene—and visual consciousness—defined as a higher order capacity. According to them, the former requires that the perceiver attends to the current mastery of SMCs. The latter, instead, manifests the possession of sensorimotor knowledge. When describing this distinction, Clark reconstructs SMEn's notion of visual consciousness as the higher-order capacity in which mastery of SMCs is "invoked *in the service of reason, planning and judgment*" (Clark, 2008, p. 180).

[50] The models also differ slightly in the role granted to representations. In Milner & Goodale's model some representations are the end-result of perception. However, for now, let us say that these representations might be compatible with the role granted to them in SMEn, as discussed in section 2.1. O'Regan & Noë (2001, p. 969) note, for instance, that Milner & Goodale's model does not posit a unified allocentric representation of the scene perceived even when it posits the involvement of representations of objects from an allocentric perspective. Although, this requires further discussion, it is not relevant for my purposes here.

integrate SMCs into Milner & Goodale's model of the two systems, low-level SMCs—those that give rise to the chauvinism worry—seem to be adequate descriptions only of the kind of information that is at play in the dorsal system, i.e. the "actual sensorimotor engagements" (Clark, 2008, p. 191). And, even then, SMCs would contribute to fixing the content of representations that, according to Milner & Goodale, play no role in perceptual experience. According to Clark, what plays a role in perceptual experience, are the ""sensorimotor summarizing resources" whose computational form is geared to use in reason, planning, and action selection" (Clark, 2008, p. 191). In short, it is not just that low-level sensorimotor details are abstracted away in perceptual experience. It is also that top-down information about objects' invariant features is deployed for planning, action, and reasoning. Mastery of SMCs might well be taking place, but this is not what determines what appears to us.

Tackling content chauvinism

For O'Regan & Noë (2001), the difference between what they call object-centred and observer-centred information is not particularly problematic, since SMCs include regularities that are due to objects' invariant features. Recall that, according to SMEn, perceivers move from egocentric information to objects' invariant features and other information about the background. Object-centred information and the allocentric coordinates of objects are known in virtue of perspectival information. Moreover, O'Regan & Noë claim that the division of labour is consistent with the distinction between movement induced regularities and object induced regularities (see section 3.2). What O'Regan & Noë, nonetheless, seem to reject is that these are two independent codes.

So, could it be the same code that is doing the work? As mentioned, Milner & Goodale's model does not require it. But, does it *allow* it? One reason against this is advanced by Jacob & Jeannerod (2003). According to them, perceptual experience requires that perceivers keep track of objects and that they are able to re-identify these objects. However, this is not something that can be accomplished by means of perceptual representations. According to Jacob & Jeannerod, visual percepts do not code information egocentrically, since the egocentric perspective is only available "to thought and imagination". Moreover, for the authors, representing egocentrically the tertiary relation involved in seeing one object to e.g. the left of another object involves a shift from visually representing the locations of these objects to representing their location in thought (Jacob &

Jeannerod, 2003, p. 32).[51] So, in order to represent more than one object at a time, the ventral system, unlike the dorsal system, would require an allocentric frame of reference. This division in frames of reference, however, leads to a communication problem, for how is the dorsal system supposed to know that the target object for the mandated action is that object that is to the left of the perceiver? (Noë, 2010).

According to Kiverstein, the communication problem can be avoided if both systems use the same frame of reference. Moreover, this frame of reference can be egocentric. Kiverstein draws on José Luis Bermúdez's (2007) notion of relational perception in which an object is represented in the "context of a scene", centred on the perceiver (Kiverstein, 2010). In relational perception, objects are represented in relation to the perceiver and to other objects. Furthermore, the relative location of multiple objects can be represented in an egocentric frame of reference. Moreover, recall that SMEn's claim is not just that SMCs are comprised by systematic co-variations between fine-grained movements and sensory input. According to SMEn's theory of the dual character of perceptual experience, perceivers start off there and move to information about objects' invariant features. This, of course, means that SMEn's embodiment claim is not the chauvinistic claim according to which *every low-level* detail makes a difference to the character of perception. Low-level details are the starting point of more complex profiles of sensorimotor interactions. Furthermore, this means that the reinterpretation of SMCs advanced by Kiverstein—which I followed in chapter 3—already presupposes, in a way, what Clark calls sensorimotor summarizing. What is at play is a more complex system of co-variations of the way sensory information changes after interactions of the perceiver. This includes low-level details, but also sensorimotor summarizing.

Of course, as Kiverstein himself notes, this thesis would fail "if it should turn out that perceptual experiences have spatial content that is purely allocentric in the way proponents of the two-visual systems hypothesis have argued" (Kiverstein, 2010).

[51] Against the idea that the egocentric perspective is not available to perception see Wheeler (2005). According to Wheeler, a robot built by Nicolas Franceschini, Jean-Marc Pichon and Christian Blanes (1992), is an example of a system that builds and uses a map with egocentric coordinates to navigate a room, avoid obstacles and reach a goal. The robot builds a map that stores information about objects in relation to the context of action of the robot, and in a space that is defined egocentrically (Wheeler, 2005, pp. 196–197).

Let me address this last point before jumping to the next section. Recent studies by Jorge Morales et. al (2020) suggest that processing associated with the ventral system is not empty of perspectival information, and are consistent with SMEn's view of the dual character of perception. In these studies, the authors aimed at assessing whether the perspectival shape of an object—e.g. the oval appearance of a circular coin—is present in perceptual experience. The authors designed a series of experiments in which subjects were asked to identify an oval coin when compared to a circular coin which appeared, in some cases, tilted. Their results showed that there was a delay in successfully picking out the oval coin when it is presented in comparison to a tilted circular coin. These results suggest that there is something common to the representation of the oval and the circular coins, which, in turn, suggests for us that the perspectival shape of the object seems to be part of the information that is coded by the ventral stream. This is reflected in the fact that the perspectival shape of an object distracts subjects when performing tasks concerned with the identification of an object. Furthermore, this suggests—as required by Kiverstein—that the content of perceptual experiences is not purely allocentric. Lending support to the compatibility between the two visual systems theory and the model of perception of SMEn.

Let me take stock here. Clark & Toribio's (2001) concern was that, for SMEn, *every* low-level detail mattered for perceptual experience. But, as discussed SMEn is not committed to this, the chauvinistic claim. SMEn's commitment is, instead, to a different claim about the contribution of SMCs to the character of perceptual experience. This commitment meant that perceptual experience has an inescapable perspectival aspect that depends upon the body of the perceiver, it is agent-centred. This stands in stark opposition to Milner & Goodale's model. What I have shown here is that there is a way of making Milner & Goodale's model compatible with SMEn. SMEn is committed to the claim that perceptual experience is, in some important way, agent-centred. However, SMCs cannot just depend on low-level details, nor are just about the changes in sensory input. SMEn's theory of SMCs already involves sensorimotor summarizing.

4.2 Sense modality chauvinism

Is there any other reason why we should worry about chauvinism when it comes to SMEn's proposal? Recall, here, that by chauvinism I meant that, under SMEn's account of perception, the dependence of the character of perception on the body of the perceiver is such that SMEn cannot extend certain attributions to all the cases of perceptual experience it would under other considerations. In the previous section, the concern was about the *content* of perceptual experience. SMEn's commitment to the embodiment claim meant that it would fail to consider evidence about the workings of the ventral visual system that showed that perceptual experience is not agent-centred as predicted by SMEn. However, there is another problem that might arise in light of the embodiment claim. This is a concern about falling into *sense modality chauvinism*.

To see why this is a different problem, let us go back to Noë. According to him, one of the things SMEn gets right is that it can explain why the TVSS device and the organic human visual system "count (…) as visual systems" (Noë, 2004). Moreover, this treatment shows that SMEn is not chauvinistic. His strategy to tackle the concern advanced by Clark & Toribio (2001) has focused on showing that SMEn delivers something akin to multiple realizability (Noë, 2004, p. 26). The thought would be the following. SMEn's view entails that systems (i.e. animals, people, artificial systems) that are physically unlike each other in some respect do not have perceptual experiences that fully share the same character. Regardless, it can still show that these occurrences are experiences of the same kind (i.e. visual or quasi-visual).

By taking this approach, however, we have changed the conversation. We are no longer discussing the problem of chauvinism in connection with the character of perception, which is the problem addressed in the previous section. Although it is still a problem that arises in virtue of the embodiment claim, it is no longer concerned with the character of perceptual experience. Let me illustrate this, briefly, with the case of sensory substitution devices. Let me reconstruct what I take to be Noë's (2004, p. 26) treatment of these cases:

(a) SMCs are lawful co-variations of sensory information after an interaction of the perceiver: that the TVSS device is placed on e.g. the subject's forehead will affect the perspective from which she receives input and, in consequence, will affect the profile of SMCs associated with the TVSS device.

(b) SMCs fix the character of perceptual experience: the features of the experience of a TVSS device user are due to the sensorimotor regularities that govern interactions between the world and the perceiver.

(c) Differences in the body of the perceiver lead, in some cases, to differences in the character of perception.

(d) Regardless of these differences in character, in both cases, subjects report that their experience is characterised by e.g. occlusion effects.

(e) In consequence, in a sense, these are instances of the same kind: the experience enabled by the TVSS device and that enabled by the organic human visual system are both cases of vision.

According to SMEn, the character of the experiences of the TVSS device user and that supported by the organic visual system must differ in some respects, since this is fixed by the profile of SMCs (e.g. the perspectival shape might be different in each case). Recall that this is the claim that gets SMEn into trouble in what concerns content chauvinism. Regardless, these can still be categorized as (almost) being of the same kind, i.e. vision and quasi-vision. Accordingly, SMEn is committed to the claim that sense modalities, the kinds in question in cases of sensory substitution, are individuated by means of their sensorimotor profile. What I want to emphasize here is that, according to this treatment of chauvinism, one should not worry that SMEn is chauvinistic because the view rightly categorizes sense modalities. This is the concern I cover in this section, namely the worry of chauvinism as it arises in connection to sense modalities. A consequence that, presumably, SMEn successfully avoids.

4.2.1 Sense modality chauvinism

According to the treatment of TVSS devices introduced above, a concern about a theory being chauvinistic can be articulated as a concern about the taxonomy delivered by a theory. In connection to this, there are two ways in which our theories could go wrong: a theory could be either chauvinistic or liberal.[52] For my purposes here, I take *chauvinistic* theories to be those that

[52] *Liberal* theories would be those that are too loose in their categorization. I draw on Ned Block's (1978) distinction between liberal and chauvinistic theories. There are different ways in which a taxonomy could result in misattributions. For instance, our taxonomy could be such that we fail to attribute a mental state at all to an individual; or that we attribute the same type of mental state to two individuals whose reports about the phenomenological character of these differ. We could also think of these taxonomical mistakes as applying at different levels. For instance, we could be chauvinistic or liberal about: *mentality* (e.g. failing to attribute mentality to systems that, under other considerations, are attributed mentality), the *content* of a token mental state, its *phenomenological character*, and, finally, the *type* under which a token mental state falls.

misidentify token mental states that fall under a type because they are unjustifiably strict in the conditions that establish whether an occurrence of a mental state is an instance of a type. For instance, a theory might not identify, e.g. human and octopus' vision as being of the same kind, identifying instead two different sense modalities, e.g. *vision-h* and *vision-o*. A competing theory might claim that this theory is chauvinistic: what the theory takes to be vision is too restrictive. The conflict between the chauvinistic theory and the competing theory is, of course, a conflict in their definition of vision. For instance, the competing theory might argue that the chauvinistic theory is using a criterion that conflicts with evidence about human and octopus' vision. This worry is, ultimately, about the way a theory carves the world.

By misattributing mental states, chauvinistic theories fail to identify that the two systems share a property, e.g. enabling visual experience. This is the mistake that, according to Noë, SMEn *does not* make. Recall that, according to Noë, SMEn embraces chauvinism to the right degree. Perhaps one way to put this is to say that SMEn embraces content chauvinism, while avoiding sense modality chauvinism. Visual experience does depend on the body, but it does not lead to an unjustifiably strict account of vision. The view shows, rightly, that systems with very different physical configurations can be recognised to be similar in the right way: the experience enabled by TVSS devices is vision-like.

Yet, this response is uninformative. No one would deny that they are *similar* systems. For starters, they both support experiences with similar effects, e.g. the occlusion effect. Furthermore, they both allow that the perceiver identifies e.g. the shape of an object. SMEn's response is meant to ease worries related to intuitions that drive, for instance, multiple realizability arguments. These are intuitions about other minds, e.g., Martians and robots. According to these intuitions, we tend to think that systems with very different configurations can be attributed the *same* kind of mental state, and not just similar states. When considered from that perspective, the concern is that, by embracing sense modality chauvinism, SMEn would fail to do justice to these intuitions. Moreover, the concern is that SMEn cannot deliver multiple realizability, the thesis according to which psychology is concerned with properties and relations such that they can be instantiated by systems that belong to different physical kinds. So, the conclusion that TVSS devices enable *visual-like* experience, given that the latter is fixed by SMCs is, in this respect, unsatisfactory.

Regardless, the problem of chauvinism and the thesis of multiple realizability are orthogonal. Consider, for instance, functionalism, the view that mental states are individuated by means of their causal relations to systemic inputs, systemic outputs, and other mental states and processes. Functionalism delivers multiple realizability, since it allows us to identify that two systems that would pertain to different physical kinds—e.g., a human being and an octopus—can instantiate the same mental state—e.g., feeling pain. By individuating a mental state in terms of its functional profile, functionalism allows us to identify cases of e.g. pain even when these are instantiated by systems that pertain to different kinds. And, in consequence, it seems that functionalism extends the attributions of kinds of mental states to the cases we intuitively think deserve them. For instance, it attributes pain to both humans and octopuses. Having the possibility to deliver multiple realizability, then, seems to be a step in the right direction in what concerns chauvinism.

Now, the problem is that, even when functionalism delivers multiple realizability, chauvinism might find its way back in. As Block (1978) rightly notes, functionalism can still be chauvinistic, since the functional profile that defines that a mental state is an instance of pain can be specified in such a way that it only describes e.g. human pain leaving octopuses' pain aside. In the case of functionalism, whether the chauvinism concern is addressed depends on the way the functional profile of a mental state is specified.

So, since one might have a view that delivers multiple realizability without avoiding chauvinism, in this section, I leave the matter of multiple realizability aside and focus on the question of chauvinism.

In line with this discussion, SMEn will be chauvinistic in what concerns sense modalities if its specification of what makes an occurrence of perceptual experience an instance of one sense modality rather than of another is too fine-grained. Put otherwise, the question is: how fine-grained must the specification of SMCs be to claim that the sensorimotor profile of the TVSS device are sufficiently similar to that of the organic human visual system?

4.2.2 Alternatives to interpret TVSS devices

To address this concern, I briefly review two alternative interpretations of the case of TVSS devices. I refer to two systems: the perceiver whose experience is enabled by the organic human visual system, and the perceiver whose experience is enabled, instead, by the TVSS device. I call the former the visual system and the latter the TVSS system.

(a) The similarity grid.

One way to interpret the case of TVSS devices is to follow Noë's proposal that TVSS devices enable "visual or quasi-visual experience" (Noë, 2004, p. 28). For him, TVSS devices and the human organic visual system enable *similar* experiences in virtue of being homomorphic with respect to their pattern of SMCs. The systematic pattern of co-variations between movements of the TVSS system and of the visual system, respectively, and the received stimuli can be mapped onto each other.[53] For instance, recall that, according to Hurley & Noë (2003), laws of occlusion are preserved. Of course, the homomorphism is partial, given the dependence of SMCs on the body.

For Noë, this should be sufficient to address the charges of chauvinism: insofar as the TVSS system exhibits a similar pattern of SMCs as that of the visual system, the TVSS system can be said to be a case of quasi-vision. We can establish, based on this, a similarity grid to handle these cases: a system counts as a visual system depending on where it stands in this similarity grid. The TVSS system stands in such close relation to the visual system that it deserves the label quasi-visual. For two systems to be similar to each other in an interesting way (e.g., to deserve to be classified as visual or quasi-visual), the pattern of SMCs does not need to be the same all the way down, but only to a relevant respect.

(b) TVSS devices enable a new modality.

Another way to interpret the case of TVSS devices is by simply claiming that these are not cases of vision at all. Although TVSS devices might exhibit relevant features that make them like, e.g., vision, we should avoid labelling these systems in a way that would invite the thought that the TVSS system is an occurrence of visual perceptual experience. Along this line, Auvray and Myin (2009) argue that TVSS devices support a different perceptual modality altogether, since they enable novel perceptual interactions. TVSS devices are a case of perceptual enhancement because

[53] According to Auvray & Myin, under this interpretation of the TVSS system, Hurley & Noë are committed to what they call the deference thesis, according to which in TVSS systems "the substituting modality is "overruled" by the substituted modality" (Auvray & Myin, 2009, p. 1050).

they transform the perceptual capacities possessed by the user. Auvray & Myin argue that we should distinguish the end-result of perceptual experiences enabled by the TVSS device—e.g., that the TVSS user reports seeing the shape of an object—from the learning process and development of new sensorimotor skills that lead to such recognition—e.g. learning to move the device.

In what concerns the charge of chauvinism regarding sense modalities, neither of these options are promising. On the one hand, if we opt for (a), we must specify what the parameters of similarity are: are there paradigmatic cases of, e.g., vision? In other words, what is the criteria to establish that TVSS systems and visual systems are similar enough? Noë recognises the many low-level differences between the visual and the TVSS system. However, what he calls high-level similarities should override those low-level differences. The similarity grid delays the question we already face (Auvray & Myin, 2009, p. 1051).

On the other hand, (b) leaves other questions open, since we now need to explain why we cannot take TVSS devices to be a case of vision, but we can attribute visual perception to two systems that are unlike each other in other ways. In other words, it seems that if we follow Auvray and Myin and accept that TVSS devices result in a different perceptual modality, we must answer whether we are supposed to reach the same conclusion when it comes to purely physiological differences. Auvray & Myin themselves recognise that sensory organs are, in a way, also tools or add-ons. At a given point in our evolutionary history, these new organs started to develop, and over a process of learning and integration, this led to the "establishment of a sensory modality". Moreover, they add that they draw this conclusion "on the pain of becoming "historical chauvinist[s]"" (Auvray & Myin, 2009, p. 1053). They might be right. Regardless, something has to be said about the physical differences that are relevant to reach that conclusion. The problem of chauvinism regarding sense modalities arises when it is said that body differences amount to differences in the kind of perceptual process, and when this is done arbitrarily or unjustifiably. In other words, when we have no principled way of claiming that only organisms like so-and-so can have perceptual experiences of this kind. For Auvray & Myin, the criterion is related to the transformation of sensorimotor skills. The skills developed by the TVSS user are unlike those of someone who relies on organic human vision. The problem with their proposal is that it remains ambiguous: when are sensorimotor skills sufficiently unlike each other to be classified differently?

Here I do not mean to say that Auvray & Myin's proposal of denying that TVSS devices enable vision is not on the right track. Instead, my claim is that the only way of tackling the charge of chauvinism regarding sense modalities is to establish a systematic way to discern these ambiguous cases.

4.2.3 Tackling sense modality chauvinism

Let me briefly summarize where we are. As mentioned before, one of the reasons why chauvinism about sense modalities is worrisome is that it flies in the face of some of our intuitions about mentality, namely, that despite physical differences two systems can be attributed as having the same sense modality. The two proposals introduced above to interpret the case of TVSS devices lack a criterion that determines whether the experiences enabled by the TVSS system and by the visual system are cases of the same sense modality. This should be determined on the basis of the profile of SMCs, since these partly determine the character of perceptual experiences. The pattern of SMCs of the TVSS system must be similar enough to the profile of SMCs of the visual system in order to be classified as instances of the same sense modality. So, the question is: how to determine whether the profiles are similar enough or homomorphic? We can phrase this as a question about how fine-grained the specification of SMCs must be to claim that e.g. TVSS systems are sufficiently like cases of vision.

In what follows, I argue that interpreting SMCs as a *structural profile* can help us address the sense modality chauvinism worry. The regularities that comprise this profile, of course, depend upon the body of the perceiver. However, despite the relevance of physical details for the lawful patterns of co-variation, it is the profile of SMCs of a perceptual system that will allow us to determine whether two occurrences of perceptual experience pertain to the same sense modality.

Sensorimotor contingencies as structural profiles of perceptual systems

Recall that, in their most basic description, SMCs describe the pattern of co-variation between certain features of the system's physiological configuration in terms of possibilities of interaction and the resulting sensory information. Accordingly, SMCs are the description of a system that comprises both the perceiver and the objects of perception. Furthermore, SMCs describe the structure of this system: the relation between its pertaining elements, their arrangement, and the patterns of interactions between them. SMCs can be said to provide a description of the structural profiles of perceptual systems. It is on the basis of these structural descriptions that we can identify how the content of a perceptual state is fixed (Myin et al., 2014, p. 394).

This structural description depends, of course, upon some physiological features of the perceptual system, for instance, on the shape of the retina. Put differently, that the sensory information changes in such a way—e.g. that the horizontal line is drawn in that way on the retina—depends on certain features of the human eye and on certain features of, well, the line. However, it is conceivable that sensory information varies in the same way when a system that differs in some respects to the human eye is confronted with the horizontal line. These systems can be described as having structural similarities.

The latter is precisely the case of TVSS devices: both the visual system and the TVSS system can be described as having a homomorphic structure of SMCs. Let me illustrate this with a case of occlusion. What SMEn predicts is that, when a TVSS device user sees the cat sitting behind the fence, the user must interact similarly to someone who relies on their visual system. Their respective profiles of SMCs overlap in that they share patterns of co-variation between sensory information and movement. This is one way of making sense of the similarity grid proposal, since by comparing the structural profiles of the visual system and the TVSS system we can see that they are homomorphic.

However, establishing that SMCs describe the structural profile of perceptual systems only takes us so far. As mentioned, it is also necessary that we establish a criterion to determine whether the profile of SMCs of two perceptual systems is homomorphic enough to claim that they enable perceptual experiences of the same sense modality. In what follows, I discuss a proposal.

The relevance of the sensorimotor coordination to tackle chauvinism

Establishing a similarity criterion cannot be done unless the notion of SMCs is well-specified. As discussed in section 3.2, the ambiguity of the notion of SMCs has been addressed by the dynamical model advanced by Buhrmann et al. (2013) and Di Paolo et al. (2017). Recall that they develop a dynamical model that operationalizes SMCs and distinguish between four kinds of SMCs: (1) the *sensorimotor environment*, (2) the *sensorimotor habitat*, (3) the *sensorimotor coordination*, and (4) the *sensorimotor strategies*. The descriptions stand at different levels of generality and they take different variables into account. Each kind of SMCs is captured by differential equations that consider specific variables, among which we can find details about the internal structure of the system and about the evolution of the system over time. These descriptions provide us with a set of possible sensorimotor trajectories that the system in question might visit under different considerations. I argue that this dynamical model of SMCs allows the identification of regularities in the profile of SMCs that allow us to tackle the sense modality chauvinism worry.

What can these descriptions do for us? In section 4.1, I introduced (and denied that SMEn holds) the chauvinistic claim according to which *every low-level detail* matters. However, rejecting it is not sufficient. SMEn must be able to tell us which details matter. The more general description, the *sensorimotor environment*, can provide an answer to this. The *sensorimotor environment* already prioritizes physical details that make a difference to the motor possibilities of the system in question and its sensory consequences. The physical details that matter are those that contribute to the structural features that constitute the profile of SMCs. This description, then, provides a more detailed and systematic account of the way sensory information changes after movement, and how it correlates with the physiological configuration of the system in question.

More importantly for our purposes, the different kinds of SMCs can offer an answer as to whether the case of the TVSS device counts as an instance of vision or quasi-vision. Each of these models provides a more systematic description of the perceptual system in question. They allow us to identify the regularities that might be shared by the TVSS and the visual systems. In other words, they provide a description of the possible sensorimotor interactions that might be shared by these systems. The problem, one might think, is that it does not say much more, particularly in what concerns the individuation of sense modalities.

The notion of SMCs that might come in handy is that of *sensorimotor coordination*. This description considers, as a variable, whether certain sensorimotor trajectories contribute to the performance of a task. To illustrate this, consider a missile guiding system. One of the tasks that the missile guiding system must perform is to identify a target and track it.

Something similar could be done in the case of the TVSS system and the visual system. By describing the profile of sensorimotor coordination of these systems when performing a task in which the capacity of, e.g., depth perception is exhibited, it might be possible to identify a shared structural profile. A description of the sensorimotor coordination of these systems could allow us to see whether their profiles of SMCs overlap in what concerns e.g. a visual capacity. What we get from this is a way of identifying whether we have good reason to consider the two occurrences of perceptual experience as cases of the same sense modality, namely that they both succeed at a certain task.

Identifying the sense modality

There is still a sense in which this proposal could remain problematic. While the profile of SMCs might allow us to say whether two perceptual experiences are occurrences of the same sense modality, it does not tell us what sense modality it is. Whether a task relates to a capacity that is visual rather than auditory is determined independently of the profile of SMCs. Moreover, it seems that it should be determined in advance. The relevant parameter to determine whether two systems are instances of sense modality X is not a description of SMCs, but of the capacities to which the tasks are related, for instance, a task related to depth perception. Moreover, whether this is a task related to depth perception instead of a capacity related to other sense modalities is not something we find out by looking at the profile of SMCs.

Nevertheless, this concern only arises if we believe that we must rely on a single criterion for sense individuation and if we believe that sense modalities are discrete categories.[54] However, criteria of sense individuation are not meant to be mutually exclusive. As Macpherson argues, cases such as sensory substitution, echolocation, bee perception, snake infrared perception, challenge precisely this, the view that sense modalities are discrete kinds (Macpherson, 2011, 2015). For her, each of the criteria illuminates different dimensions of sense modalities. SMEn can rely on other criteria to determine whether the task selected for analysis exhibits the possession of a capacity that pertains to a given sense modality. Moreover, this proposal provides a way in which SMEn can avoid the charge of chauvinism regarding sense modalities. The descriptions of the sensorimotor coordination allow us to identify how a perceptual system deploys its sensorimotor skills to tackle a task.

[54] Paul Grice (1962) proposed to distinguish between four criteria: (1) the content criterion, (2) the qualia criterion, (3) the stimuli criterion, (4) the sense organ criterion. Auvray & Myin (2009) add three more: (5) behavioural equivalence, (6) dedication, and (7) sensorimotor equivalence. According to Auvray & Myin, SMEn defends (7) sensorimotor equivalence. As I have shown, however, to properly distinguish between sense modalities, SMEn needs more than just the profile of SMCs. It must commit to a view in which the different criteria illuminate different aspects of perceptual experience.

In this chapter, I have tackled two different versions of the chauvinism worry, both of which arise in virtue of the claim that the profile of SMCs depends on the body of the perceiver, and that this profile serves important explanatory roles. On the one hand, SMCs describe the sensorimotor interactions that give rise to perceptual experiences and their associated characters. On the other hand, SMCs serve as a criterion to determine the sense modality of an experience. The concern, generally speaking, is that the dependence on the body precludes SMEn from fulfilling these explanatory goals.

According to SMEn's embodiment claim, the profile of SMCs is sensitive to at least some low-level details of the body. In what concerns the content of perceptual experience, the idea is that SMCs fix, at least in part, the content of perceptual experience. But if this is so, and SMEn sticks to the embodiment claim, the theory would fail to do justice to evidence that the content of perceptual experience does not have that irreducible perspectival character that would be consistent with the idea that low-level details make at least some difference to experience. This concern finds support in Milner & Goodale's view. One way to summarize—and oversimplify—their view is the following: according to them, one is aware of the way things are, not of the way things appear. The system that has access to the way things appear does not contribute to your visual consciousness. I claimed, based on Morales et al. (2020), that appearances are pervasive in the content of perceptual experience. This lends support to SMEn's view that perspectival information, that might be dependent upon low-level details, is retained in perceptual experience.

Of course, the existence of the two systems is not put into question. Both comprise a typical visual system. Both contribute to our dynamic relations to things in the world. Moreover, both have an irreducible perspectival character. The content of our perceptual experience is brought about by those dynamic relations, whose lawful profile are SMCs. The different ways in which we come into sensory contact with the world depends on these SMCs. The best way to classify these different relations—the different sense modalities—is by referring to the profile of SMCs, a profile that is sensitive to details of perceivers' bodies. Here enters our second concern. The worry, here, is that, by committing to this claim, SMEn would fall into another kind of chauvinism, a chauvinism regarding sense modalities. By invoking the dynamic model of SMCs, however, SMEn can refer to specific capacities to discern whether two occurrences of perceptual experience justifiably deserve to be classified as instances of the same sense modality.

As can be seen, SMCs enjoy a key explanatory role in SMEn. Moreover, they are meant to provide a bridge between the theory of perception provided by SMEn and an account of the mechanisms that enable perceptual experience. However, some claim that more needs to be said about this mechanistic story. Part III explores one alternative.

III. The Predictive Approach to Sensorimotor Enactivism

Chapter 5. Predictive Processing and the Free Energy Approach

In section 2.2, I discussed that Sensorimotor Enactivism (SMEn) is a naturalistic theory of perception. Accordingly, it aims at being continuous with an account of the mechanisms that enable perceptual experience. More specifically, it aims at providing a naturalistic account of sensorimotor knowledge. A key element to achieve this continuity is the theory of sensorimotor contingencies (SMCs). As mentioned, one important problem within SMEn relates to the ambiguity in the notion of SMCs, particularly the account developed by Noë and O'Regan, whose arguments have served, for the most part, as a basis for the arguments discussed in this thesis. The lack of clarity in the notion of SMCs ultimately hinders SMEn's prospects of providing a causal story of the mechanisms that enable sensorimotor knowledge. Buhrmann et al.'s (2013) and Di Paolo et al.'s (2017) operational definitions of four kinds of SMCs, introduced in section 3.2, are crucial contributions towards this project. However, more needs to be said about the causal story that accounts for sensorimotor knowledge, and about the place of the brain in it.

In line with this, Seth claims that one challenge faced by SMEn is to explain: "at the level of neural mechanism what is meant by [SMCs] and their mastery" (Seth, 2014). Note here that Seth already frames this question as being directly connected to the neural mechanisms that enable perceptual experience. However, it is important to keep these two questions apart. Firstly, because this points to one further challenge for SMEn, namely that of accounting for the role of the brain in perceptual experience. And secondly, as discussed in section 2.2, because of SMEn's commitment to explanatory externalism according to which the brain is not the sole vehicle of perceptual experience. Instead, the vehicles of perception are distributed across brain, body, and world because only by referring to these can one provide the best explanation of certain features of perceptual experience. So, even when Seth is right in pointing out that SMEn is missing an account about the role of the brain in enabling sensorimotor knowledge, this is part of a greater omission.

These are the issues addressed in Part III. Following Seth (2014) and Adrian Downey (2017), I argue that one alternative to tackle this matter is to take a predictive approach to SMEn. More specifically, I propose to draw on the Free Energy Approach (FEA) developed mainly by Karl

Friston. This chapter focuses exclusively on predictive processing and on the FEA. [55] In chapter 6, I discuss the predictive approach to SMEn and come back to the questions introduced above.

5.1 The predictive framework: Predictive Processing and the Free Energy Approach

Let us begin by discussing briefly the key concepts and main features of predictive processing. I specifically focus on the idea that cognitive systems are taken to be prediction machines organized hierarchically, as well as on the principles of Bayesian inference, the computational model implemented by predictive processing.

5.1.1 Predictive Processing

Predictive processing is, standardly, a functional approach to neural dynamics.[56] The framework takes the brain to be driven by top-down processing, and thus opposes traditional approaches that take neural activity to be driven by stimuli or bottom-up processing. Predictive processing has become particularly appealing given its potential to provide a single strategy for understanding neural activity and the cognitive system.[57]

Predictive processing takes its name from predictive coding, a term that typically names a data compression strategy that is used in computational sciences and that has been used as a modelling strategy for neural networks. Predictive coding works under the assumption that processing is more efficient when it relies on information already possessed by the system. Consider, for instance,

[55] I refer to the predictive framework as an umbrella term to refer to both predictive processing and the FEA. Regardless, I take these as independent theses. While the FEA is a theory about the adaptive behaviour of biological systems, predictive processing is standardly taken as a functional approach to the brain with the potential of providing a unifying strategy to understand neural activity. Predictive processing is understood as a possible implementation of the FEA. It is worth noting that nothing in the FEA's view of the adaptive behaviour of biological systems mandates that surprise minimization should be implemented in hierarchical models such as those posited by predictive processing (Clark, 2017b). Nonetheless, others in the predictive processing literature indicate that the link between the FEA and predictive processing is tighter. Jakob Hohwy (2013), for instance, does not take these frameworks to be fundamentally different or independent. Buying into predictive processing already implies the acceptance of core principles of the FEA, which expresses principles that are essential to predictive processing.

[56] Although a lot of work has focused on the neural activity of the visual cortex, this model strategy has also been applied to other areas of the brain, e.g. the auditory system, the hippocampus, the ventral midbrain, and the frontal cortex (Huang & Rao, 2011, p. 11), and to perception, memory, affection, interoception, proprioception, and attention.

[57] Andrew Sims (2017) distinguishes between three different positions regarding the scope of predictive processing: *minimal predictive processing*, according to which only a few processes can be modelled, *mixed predictive processing*, which encompasses most processes, and *maximal predictive processing*, according to which predictive processing encompasses every neurocognitive function (e.g., the FEA and the proposal advanced by Bar (2009)).

Huang & Rao's (2011) application of this strategy to the processing of natural images. Natural signals, they claim, are "highly redundant" and exhibit temporal and spatial uniformity (Huang & Rao, 2011, p. 580). Instead of requiring the processing of incoming signals corresponding to each pixel, based on the assumption that neighbouring pixels tend to have an intensity of around the same value, their intensity is calculated by predicting it on the basis of surrounding pixels. Early sensory processing seeks to reduce redundancy found in the environment through the prediction of regularities, thus "transmitting only what is not predictable", namely residual differences between the prediction and the stimuli.

Although the terms predictive processing and predictive coding are at times used interchangeably in the literature, these are not the same. Predictive processing combines predictive coding's processing strategy with a view according to which predictions are advanced from a layered or hierarchically structured model (Clark, 2016). It has been mostly developed as a model of neural activity of the visual cortex and as providing, mainly, a mechanistic account of perception. As a theory of perception, the framework is best articulated as a constructivist approach. According to Nico Orlandi, for instance, for predictive processing, "perceptual processes consist in internal *constructions* of distal objects and scenes where *prior knowledge* about the likely causes of sensory stimulation is deployed to figure out what is present in the environment" (Orlandi, 2016, p. 329).

Importantly, as has been shown by Clark (2016, 2017a), a description of predictive processing as a neuro-centric approach is not mandatory. For him, there are some valuable lessons to be learned from developments in embodied cognition, e.g., that problem-solving can be distributed between brain, body and world (Clark, 2016, p. 246), which are already reflected in predictive processing. Neural dynamics, as they are modelled by predictive processing, should be considered to be in the service of "world-engaging action" (Clark, 2016, p. 250). In line with this, Clark (2017a) argues that neural processing can be understood as enabling efficient and reliable interactions with the world.

Cognitive processing is described as a cascade of predictions that is met by and adjusted on the basis of incoming stimuli. To illustrate this framework, consider visual tracking, the capacity to follow a visual cue, e.g. a bird flying in the sky.[58] According to predictive processing, this capacity

[58] Adapted from (Friston et al. 2012; Seth 2014).

can be explained by taking the brain to predict, based on prior information and its best model of the world, the position of the bird as it moves. Importantly, this model, namely the *generative model*, has a hierarchical structure,[59] and it is constantly updated on the basis of the system's interactions, with the aim of achieving better predictions over time or of minimizing prediction error over time. The idea, roughly, is that the advanced prediction is met, at a lower level, by incoming stimuli. The prediction is compared against this input, thus providing *error feedback* on the prediction. In addition to processing error feedback, the brain also processes its reliability and salience, namely its *precision*.[60] Error feedback is given more or less weight depending on its reliability (Clark, 2016, p. 57).

Next, I present some of these features in more detail. I begin by discussing the hierarchical nature of the predictive processing.

The hierarchical organization of the system

There are a number of advantages to predictive processing's proposal. It represents an efficient way of processing information. Information is coded in an economical manner, and it allows the system to deliver more adequate predictions over time (Friston, 2005). Moreover, if this is the right way of modelling neural activity associated with different cognitive processes, the cognitive system would turn out to possess an efficient strategy to "uncover" the objects that populate the world, their properties, and their relations (Clark, 2016). The problem is, of course, that many different things could be the hidden cause of the stimuli received by the system. Wiese & Metzinger (2017) call this the problem of perception, which arises because the system has access to sensory information that could have been caused by alternative objects, e.g., objects that share the same colour and shape.

This problem is solved by considering, for starters, predictive processing's basic strategy of advancing predictions of hidden causes of stimuli and refining them based on the feedback received. This is a kind of self-supervised learning in which the system arrives to more refined

[59] The term "generative model" might refer to two different things. On the one hand, it might refer to the coding of the system's prior knowledge along with probability of certain scenarios being the case. On the other hand, the hierarchical generative model refers to the entire system as a system that is characterised by its layered structure and the cascade of predictions that traverse said structure.
[60] For an overview of the notion of precision, see Clark (2016) and Kirchhoff & Kiverstein (2019a).

predictions over time. In addition to this, the generative models from which predictions are advanced are hierarchically organized. For instance, if we had a two-layer structure, the model of the lower level would be constituted by the predictions advanced by the higher level. This enables the system to capture different levels of complexity that enrich the kind of predictions that are advanced, combining predictions about the objects that populate the world, its properties (e.g. shape, size, colour), and the pattern or retinal stimulation that they might cause (Clark, 2016, p. 21). Moreover, the different levels of the layered structure of the generative model can also be updated and adjusted in a parallel fashion. This adds up to the system's capacity to compute features of external objects simultaneously, but at different levels of the generative model (Wiese & Metzinger, 2017, p. 7).

One way to articulate and formalize the thought that the brain is a predictive machine is by means of Bayesian inference. Let me turn to this next.

Bayesian inference

Bayesian inference is a computational method that allows us to model the thought that "what we experience is the most *probable* worldly state, given what we already know about the world (our priors) and the current sensory stimulation (the evidence, or 'likelihood')" (Clark 2013, n. 5).[61] Bayes' theorem is usually associated with theories of rationality. An agent is said to be rational when she updates her system of beliefs in conformity with this theorem. The reason for that is that this theorem models the way prior beliefs are updated in light of new evidence or new information, thus, resulting in an updated (and improved) set of beliefs (Orlandi, 2016, p. 330).

Let us translate this to the case of perceptual processing: sensory information presented to the system plays the role of available evidence, while a mock or proxy stimulus that is delivered by the system plays the role of the prediction. Huang & Rao formalize this with "a generative model $P(\boldsymbol{I}|\boldsymbol{r})$, which is the probability of an image $\boldsymbol{I}$ given a set of hidden internal model parameters $\boldsymbol{r}$ (represented by firing rates in a network of neurons). For a given input image $\boldsymbol{I}$, the neural system is assumed to select the parameters $\boldsymbol{r}$ that maximize the posterior probability $P(\boldsymbol{r}|\boldsymbol{I}) = P(\boldsymbol{I}|\boldsymbol{r})P(\boldsymbol{r})/P(\boldsymbol{I})$ obtained using Bayes theorem (…)" (Huang and Rao 2011, 2).

[61] Clark traces the roots of this view to Hermann von Helmholtz (2005), who held the view that perception is mediated by inferential reasoning.

It is important to note that this computational method describes prior information probabilistically. The information already possessed by the system is the generative model, a probability distribution that includes information about possible external causes, along with their likelihood. Consider, for instance, a typical case of perception, e.g. perceiving an apple. In this case, the proximal stimuli would play the role of the evidence that is available to the perceiver. How probable it is that the perceiver encounters an apple, rather than a pear, is the likelihood of the prediction. This is obtained given prior information and current evidence of that scenario.

Bayesian inference, however, has a downside: it is computationally very complex, and, in many cases, it is intractable. The computations the brain would have to process to implement Bayesian inference are too complex. To make this more tractable, most approaches model the behaviour of the brain using approximate Bayesian computations: rather than modelling or representing the probability of having a certain scenario in light of the current evidence available, they posit a function that approximates that probability. According to these approaches, the estimations in neural activity approximate the results one would obtain when applying Bayes' theorem. Approximate Bayesian computations posit a likelihood function that limits the space of probability from which the prediction is derived, making them computationally tractable.[62] This means that the system does not actually have to represent the probability distribution to infer the causes of stimuli (Wiese, 2017b).

Now, the computational strategy chosen to model neural activity must be justified. This is the role played by the Free-Energy principle. The FEA is the story behind the choice of free-energy as the function that limits or bounds the space of probability (Friston & Stephan, 2007, p. 425). It is with this approach that I show SMEn to be compatible. So, let me now turn to the FEA.

5.1.2 The Free-Energy Approach

While the FEA is a theory about the adaptive behaviour of biological systems, predictive processing is one way to implement this theory. In that sense, within the context of the FEA, predictive processing can be taken as a possible account of the workings of the central nervous

[62] One concern is that, by positing approximate Bayesian computations, predictive processing loses the rich models that would allow the framework to "scale to higher-order cognition" (Kwisthout & van Rooij, 2020).

system and its contribution to the adaptive behaviour of biological systems. To understand the FEA, let us begin with its notion of a biological system.

The biological system in the Free-Energy approach

According to the FEA, the biological system is an embodied agent whose boundaries are defined in virtue of its interactions with the environment. This notion fulfils a heuristic function, since it serves as a guide for the formulation of the model that would explain and predict neural processing. Biological systems are open systems that are not in equilibrium and seek to minimize dispersion and reach viability. Adaptive behaviour, in this approach, consists in the minimization of surprising dealings with the environment. Here, surprise does not refer to a personal-level phenomenon, but to a measure of the probability of states in which an agent might find itself given its model. Surprising states are those that have a low probability of being visited by an agent on average and over time, and which would jeopardize an agent's viability. Unsurprising states are, in contrast, states that have a high probability of occurring (Kiverstein, 2020).

The free-energy principle states that, in order to minimize disorder or dispersion, biological systems "make implicit inferences about their surroundings" (Friston & Stephan, 2007, p. 420). The process described by predictive processing becomes, here, the system's strategy for survival. Let us put it this way: to ensure their viability, biological systems minimize surprise by inferring which states they are expected to visit over time, i.e. which states have higher probability or are less surprising. The biological system, however, lacks access to the measure of probability of the states it is expected to visit over time because it would need to evaluate all the possible states it could visit. This is the information provided by the notion of free-energy (Kiverstein, 2020).

Now, the free-energy principle originated in statistical thermodynamics. In that context, free-energy measures the discrepancy between the energy of a system and its entropy, i.e. its tendency towards disorder. In our context, the notion at play is *informational* or *statistical free-energy*. Free-energy is a function of probability distributions (Friston & Stephan, 2007, p. 420), an upper-bound on surprise, given a generative model. It measures the probability of states given the system's model. For Friston, free-energy can be evaluated because, simply put, it is a function of sensory states and the current state of the brain. Given that free-energy is always greater than surprise, by minimizing free-energy, the agent minimizes surprise (Friston, 2009, p. 294).

Neural activity is, thus, understood in terms of surprise minimization, as an instance of the adaptive behaviour of biological agents. A system deploys two strategies to reduce the probability of error in its predictions over time (and, thus, minimize surprise): (a) changing its internal states or parameters (i.e. perceptual inference), and (b) acting upon the environment (i.e. active inference). Note here that the strategy mainly discussed in the previous section is associated with perceptual inference. These two strategies are deeply intertwined and must be understood as belonging to the same cycle of surprise minimization that defines the interactions of a system (Friston, 2010, p. 129).[63] For the FEA, interaction acquires the primary role over perception. Perceptual inference is subordinated in this sense to active inference. The system seeks its preservation in its engagements with the environment; however, for these engagements to be unsurprising it requires updated beliefs about the world. Perceptual inference as developed by traditional predictive processing is thus subordinated to the active engagements of the system. Just as in perceptual inference what is predicted is the most likely state of the world, in active inference what is predicted is the state of the world that would serve the preservation of the system through a motor command which will bring about this state. Think back to the visual tracking example mentioned above. In this case, the result of the inferential process is the movement of our head directed towards the predicted future position of the bird.

The generative model, the recognition density, and the generative process

At this point, the distinction between the *recognition density,* the *generative model* and the *generative process* becomes relevant (Ramstead et al., 2019). According to Ramstead et al. this distinction marks an important shift from hierarchical predictive processing to the FEA. The *recognition density* refers here to the structure that recapitulates the environment as it becomes relevant to the system. It is coded by the internal states of the system and it constitutes the guess work of the system to approximate the generative model, the model that tracks the hidden causes of sensory states. According to Ramstead et al., this might be encoded neurally or phenotypically. They add that biological systems *embody* the recognition density.

[63] In this context, the subordination of perceptual inference to active inference is such that some have proposed to take the two strategies of surprise minimization as constitutive of active inference. See, e.g., (Adams et al. 2012, p. 6).

The recognition density, in turn, entails, the *generative model*. Unlike the recognition density, the generative model is enacted. It is a "normative [model] of 'what ought to be the case, given the kind of creature I am'" (Ramstead et al., 2019, p. 233). The information it provides is not just the hidden causes of sensory states, but also the consequences of given actions. In that sense, the *generative model* is a "policy selector". It is not physically instantiated, but instead is enacted in the dynamic interactions of the agent. In other words, it is instantiated in the normative behaviour and actions of biological systems. The generative model is a statistical description of the states that a biological system ought to occupy to minimize surprise, and it is not something that can be abstracted away from the biological agent as a whole (Kirchhoff & Kiverstein, 2019a, p. 57).

In active inference, the agent generates its own constitution and maintains itself as a self-model. While the generative model "selects (…) ways of acting", the *generative process* brings forth or enacts those actions. The *generative process* corresponds to the dynamic coupling with the environment that "enables an agent to reduce uncertainty" over time (Kirchhoff & Kiverstein, 2019a, p. 57). It is through the generative process that the agent brings about the states that serve its preservation and that appear in the generative model.

These distinctions lead us to the representational profile of the FEA. According to the latter, by encoding the recognition density, the biological system as a whole is taken to model or represent the environment.[64] This can be seen in the following claim by Friston & Stephan: "adaptive systems (…) represent the state and causal architecture of the environment in which they are immersed. Conversely, this means that causal regularities in the environment are transcribed into the system's configuration" (Friston & Stephan, 2007, p. 426). The proponents of the FEA emphasize that a system represents the causal architecture of the environment and that this architecture is embedded in the structure of the system. In line with this, Clark claims that: "(…) the free-energy minimizing 'model' that does the mirroring is in fact the whole embodied, active organism" (Clark, 2017b, p. 6).

Friston, for instance, characterizes the recognition density as hierarchical, nonlinear and dynamic features that are inherited from the world and reflected in the model according to the timescale of

[64] The systematic distinction between generative model and recognition density was advanced by Ramstead et al. (2019). Arguments about representations in the context of the FEA have typically referred to the generative model. Ramstead et al. propose to read these arguments in relation to the recognition density, instead. I have adjusted the claims to reflect this distinction.

the phenomenon represented (Friston, 2012, p. S 174). The recognition density codes external changes based on their different timescales, distinguishing thus between: (i) rapid environmental changes caused by "structural instabilities or other organisms"; (ii) changes that happen over seconds, such as illumination or "slowly varying fields", i.e., contextual changes that are state dependent in that they are expected given previous states or cues in the environment; and, finally, (iii) long-term changes that might correspond to "an invariant aspect of [the environment's] causal architecture" which corresponds to "physical laws and structural regularities" (Friston & Stephan, 2007, p. 429).

It is important to note that the recognition density posited by the FEA is not agent independent. As Williams shows, it does not track the world in objective terms, but as it appears to the agent, in agent-dependent terms, so that it is a recollection of "the environment as it *matters* to the organism and its physical integrity" (Williams, 2018, p. 158). Given that the recognition density serves the agent's ultimate goal which is surprise minimisation, these are both agent- and interest-dependent. The agent is attuned to relational information that is useful to guide its own interactions. This point is also emphasised by Bruineberg et al. (2018), who note that the FEA's description of the biological agent is not that of a system that advances perfect predictions or predictions that correspond to an unbiased and objective perspective of the environment.[65] I later come back to a more detailed description of their perspective. For now, it is important to keep in mind that, for the FEA, the recognition density is an agent specific model of the environment. Furthermore, what is recapitulated in the recognition density is not an objective environment, but the environment as it is significant for the survival for the biological system, the system's *Umwelt* (Clark, 2013a, 2017a; Williams, 2018).

The Markov blanket and the boundaries of the biological system

The FEA holds that it is the whole biological system that constitutes the model or representation from which predictions are advanced and that recapitulates the environment. The Markov blanket, a probabilistic model introduced in the context of Bayesian networks (Kirchhoff & Kiverstein,

[65] Another proposal in the vicinity is Wiese's (2017a) view that, in active inference, the system's predictions are misrepresentations.

2019b, p. 2; Pearl, 1988), provides a way of modelling the biological system when it is understood in these terms.

Consider a complex system—a causal net—comprised by a set of interacting nodes. Within this system, we can identify different sets of nodes and their states based on the way they influence each other. Take a given set of nodes. The nodes they influence are their children. The nodes that influence them are their parents. Finally, we can also identify the nodes that are influenced by their children and any other node that influences their children. The Markov blanket names a boundary between "families of nodes", i.e. a node is shielded by a Markov blanket when its state can be "fully predicted" on the basis of its neighbouring nodes, independently of any other state of the system outside the Markov blanket. The Markov blanket, in other words, names the boundary of a set of nodes whose dynamics we can fully know independently of other states of the system (Clark, 2017b). A node within a complex system can, therefore, be enclosed within several nested Markov blankets.

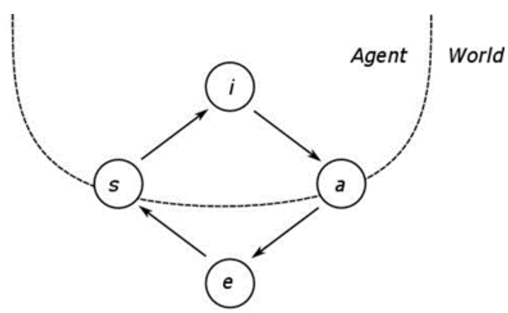

Figure 2. The Markov Blanket. Based on (Hohwy, 2017, p. 10).

The biological system can be modelled in this way, by distinguishing between internal, external, active, and sensory states. Sensory states are the children of external states, i.e. they are determined by external states. External states, on the other hand, are the children of action states, i.e. they are determined by action states. Action states are, in turn, determined by internal states. And internal states are, finally, determined by sensory states (see figure 2). This illustrates the cycle of surprise minimization that is a feature of self-organization. States that are beyond the Markov blanket are hidden from the internal states of the system. In active inference, the system 'infers' these hidden states and acts upon the world to bring them about. In connection to this, Friston claims that:

> [T]he partition of states implied by the Markov blanket endows internal states with the apparent capacity to represent hidden states probabilistically, so that they appear to infer the hidden causes of their sensory states (…). By the circular causality induced by the Markov blanket, sensory states depend on active states, rendering inference active or embodied (Friston, 2013, p. 5).

According to Hohwy (2017), sensory and action states form the Markov blanket that bounds the mind. The mind is secluded from states external to it, which are inferred. The epistemic seclusion implied by the Markov blanket leads us to a debate regarding the boundaries of the system in question. Clark reconstructs Hohwy's view in the following way: according to Hohwy, everything that is beyond the Markov blanket is "otiose" in relation to predicting the states internal to the mind. The mind only has access to the evidence that is provided by the sensory states (Clark, 2017b, p. 3). Moreover, the system reinforces its own boundaries. By advancing predictions that are improved over time, the system maximizes the evidence not only of what it predicts but of itself (Friston, 2010, p. 2, as quoted by Hohwy (2016, p. 264); see also (Hohwy, 2017)). When the agent successfully predicts and brings about unsurprising states, it gathers evidence that not only confirms the predictions, but also confirms itself as a model that "fits the domain it is modelling" (Kirchhoff & Kiverstein, 2019b). The idea is that the structure of the Markov blanket results in an isolated system. This speaks of the epistemic situation of the agent, as well as of the boundaries of the system.

However, this interpretation is rejected by some proponents of the free-energy approach (Clark, 2017b; Allen & Friston, 2016). Clark, for instance, emphasizes the flexibility of the Markov blanket. He argues that there are at least some cases in which an external object is incorporated into the strategy for surprise minimization. Clark illustrates this with the case of a cane. When thinking about the role it might play for someone who is visually impaired, the cane appears not as an object perceived, but as an extension of the arm for the perception of the environment. In this case, the prediction is advanced in virtue of the evidence provided by the use of the cane (Clark, 2017b, p. 11). If we take the system to be defined as a surprise minimization engine, so to speak, the cane can be considered as part of the system.[66] So, there are at least some cases in which the system involved in the cycle of surprise minimization is not just the brain, or the biological system. There are instances where the boundaries of the system, understood as a surprise minimization engine, are flexible.

[66] This might face the bloat problem, the concern that anything external that is once integrated into a cycle of surprise minimization becomes part of the system. To avoid this, Kirchhoff & Kiverstein complement this argument by claiming that only those elements that are integrated as part of an action policy on which the system has high confidence—i.e. a reliable course of action—should be considered to be part of it (Kirchhoff & Kiverstein, 2019b).

It follows that, from the perspective of the FEA, the cognitive system is not just the brain. Instead, it is constituted by the whole biological system as a surprise minimization machine. Moreover, the generative model seems to be a good candidate for the mechanistic story of SMCs. The embodied skills that constitute sensorimotor knowledge are enabled and explained by this cascade of predictions. Furthermore, predictive processing provides a way to make sense of SMEn's claim that the content of perception is partly fixed by expectations. In line with this idea, for example, Clark claims that the models proposed by the predictive framework "embody rich sets of expectations concerning how visual input should vary with (for example) movements of eyes and body, and even with covert shifts of attention" (Clark, 2013a, p. 494).

5.2 Representations in the predictive framework

The description of perception advanced by the predictive framework, according to which perception is an anticipatory phenomenon, sits nicely with SMEn's claim that expectations play a role in fixing the content of perception. While for SMEn, the character of perceptual experience is partly fixed by our expectations, predictive processing claims that the brain advances predictions about the object of perception. That said, there are some reasons that speak against the project of invoking predictive processing to account for sensorimotor knowledge. In this section, I begin by discussing these concerns. I focus, for the most part, on the representational profile of the predictive framework because this constitutes the greatest challenge to the viability of a predictive approach to SMEn.

5.2.1 Some concerns about the predictive framework

Before turning to concerns that are directly related to the compatibility between SMEn and the predictive framework, it is worth reviewing some general concerns surrounding the predictive framework.

Take, for instance, the worries raised by Flament-Fultot (2016), who argues that predictive processing has two important shortcomings. Firstly, he questions the plausibility of the model posited by predictive processing. The theory states that the architecture of the brain must be hierarchical, but for him it is not clear that hierarchical models can be implemented given the anatomy of the brain. Secondly, given the complexity of the models posited by predictive processing, he worries that these might be computationally intractable. These points deserve attention. Note, though, both are open empirical questions. In the next paragraphs I only offer a brief review of what has been said about these points.

Let me start with the second point. Kwisthout & van Rooij (2013) call this the tractability paradox. According to them, if we were to adopt the model of neural activity proposed by the predictive framework, a tension would arise between the overly demanding computations that would need to be implemented by the brain and "the ease and speed" with which the human brain performs its normal functions. The proponents of the predictive framework are not oblivious to this, which is why they rely on Approximate Bayesian computation. Regardless, as Kwisthout & van Rooij observe, whether the proposed approximation tackles the tractability paradox is something that needs to be reviewed on a case by case basis. For instance, they offer a method to test the tractability paradox for a specific Bayesian model. Furthermore, Kwisthout & van Rooij (2020) claim that, given that approximations rely on simplified models, it becomes less plausible that these are the rich models that predictive processing claims support higher order cognition. Of course, this depends ultimately on the proposed approximation. I briefly come back to this point in section 6.3.

In connection to Flament-Fultot's first concern, namely the plausibility of the neural implementation of hierarchical predictive processing, it is worth beginning by addressing whether this constitutes an anatomical or a functional description of the brain or the cognitive system. As mentioned, predictive processing has been mainly concerned with modelling neural activity. Furthermore, it is an approach to neural processing that intends to be biologically plausible. Nonetheless, it does not stand at the level of implementation, it is rather a functional approach (Wiese & Metzinger, 2017).

In order to reconcile the functional and the anatomical notions of hierarchical processing, Karsten Rauss & Gilles Pourtois (2013) propose the following. Firstly, they assume (a) that these are not opposing considerations; and (b) that both processes—top-down and bottom-up cognitive processing—involve ascending and descending neural connections. In other words, that the functional distinction between top-down and bottom-up functional processing does not map a distinction between the kind of anatomical connections that are involved. Furthermore, they believe that behind these two ideas lies the assumption "that bidirectional information exchange between the levels of hierarchical systems serves to reconcile incoming information with internally generated predictions" (Rauss & Pourtois, 2013, p. 4). So, according to the authors, predictive processing would be committed to the following: (a) there is an exchange of information between different functional levels, (b) the exchange of information is bidirectional, (c) the system advances predictions, and (d) these predictions are reconciled with incoming information external to the relevant level.

In a recent paper, Walsh et al. (2020) review the neurophysiological evidence in favour of the predictive processing, focusing on its key claims.[67] While predictive processing can accommodate important evidence related to sensory processing, more needs to be done to support predictive processing against other competing models. According to the authors, there is mixed evidence in what concerns the hierarchical organization of the processing. In this respect, predictive processing is concerned with the claim that the hierarchical organization of the processing enables the system to direct processing that occurs at different levels and at different time-scales towards a coherent and unified global representation (Walsh et al., 2020). Evidence that neural activity can be interpreted in terms of expectation- and error-units is insufficient to show that hierarchical predictive processing might be implemented in the brain. Instead, evidence that might suggest the latter relates to the suppression of neural activity at different levels. More specifically, while neuronal responses seem to be modulated by the suppression of neuronal activity following repeated stimuli in early phases, at later phases neuronal responses are instead modulated by the suppression of activity following what can be interpreted as global statistical inferences. This case too, however, is consistent with other alternative hypotheses. It is insufficient to show that there is influence in early areas from descending connections. What needs to be shown is that there is an interplay between different areas that are simultaneously implementing predictive processing. Again, these matters are open to further empirical research.

In addition to these concerns, there are other worries linked directly to the compatibility of SMEn and predictive processing that need to be addressed before bringing these frameworks together.

[67] They list the following key claims of predictive processing: "1. Error-signaling neural responses to sensory stimuli should scale inversely with expectation. 2. Top-down signals represent sensory prediction. 3. At each level of the cortical hierarchy there are two functionally distinct neural subpopulations representing predictions and prediction errors. 4. Prediction error minimization is achieved through reciprocal exchange of error and prediction signals across levels—a process known as "hierarchical inference"" (Walsh et al., 2020, p. 244).

The first relates to SMEn's idea that, as discussed in section 2.2, the interactions of the embodied agent should be prioritized over the specification of the neural mechanisms that enable perception. However, as can be seen from the discussion so far, even when predictive processing and the FEA approach seem to focus on neural dynamics these are understood within the wider dynamic of the biological system. So, it is not only the brain that should be considered the prediction machine, but the whole biological system. In line with this, for instance, Clark proposes that, in the context of the FEA, problem solving is distributed over brain, body and world.

The second concern about a predictive approach to SMEn is that the predictive framework is representational. Given that, as discussed in section 2.1, SMEn constrains the role played by representations, as well as the content these might have, more needs to be said to assess whether these approaches can be brought together. In the following section, I discuss the predictive framework's stance towards representations.

5.2.2 The representational profile of the predictive framework

Arguments in favour of the representational approach to the predictive framework

Representational language is pervasive in the predictive framework. If the FEA is rightly characterised as a representational framework, a predictive approach to SMEn becomes considerably more difficult to justify. In this section, I discuss the arguments in favour of a representational reading of the predictive framework and the FEA to show in the following sections that, even under such an interpretation, the FEA can be used to explain sensorimotor knowledge. With this in mind, I build the strongest case possible for the representational reading of the FEA to show that the two frameworks are compatible. My aim here is not to defend the representationalist interpretation of the FEA, as such. For that reason, the arguments presented here should not be taken to be knock-out arguments in favour of this reading.

Why does the FEA deserve the 'representational' label in the first place? After all, the relation between encoded environmental changes and the recognition density could be merely a causal correlation. If this was the case, some might find the representational label worrisome. Ramsey, for instance, considers that if the term refers only to the activation of an internal state of a system given an input that "plays a mediating role in the processing" (Ramsey, 2015, p. 10), then an extremely deflationary notion of representation is on the table and, in consequence, it becomes trivial. In order to identify cognitive structures that are truly representational, Ramsey (2007) advances what he calls the job description challenge. A theory is expected to provide an account of the essential features in virtue of which a cognitive structure can be considered a representation. It is necessary to provide "a job description that tells us what it is for something to function as a representation in a physical system" *qua* representation (Ramsey, 2007, p. 27). Meeting the job description challenge is sufficient for a structure to be described as a representation.

So, is the FEA's use of the term representation trivial? Gładziejewski (2016) offers a good case for the representational profile of the predictive framework. He claims that the predictive framework meets the job description challenge advanced by Ramsey. To make his case, Gładziejewski compares the functional features of the recognition density posited by the predictive framework with the functional features of cartographical maps, a paradigmatic case of representations.[68] He proceeds as follows: if the recognition density shares with cartographical maps the functional features that allow the latter to fulfil a paradigmatic representational role, then the representational profile of the predictive framework is justified. These features are: i. action-guidance, ii. (partial) detachability, and iii. error detection. Gładziejewski's claim is not that the recognition density displays some functional features of a paradigmatic case of representation, but that the recognition density displays *all three functional features* (Gładziejewski, 2016, p. 570). For this reason, he takes his position to be provisionally immune to anti-representational arguments that seek to trivialize the use of representations in the framework.

Let us review how these features are displayed by the recognition density. Just like cartographical maps, these models track certain structural features of the environment. Recognition densities can be characterised as structural representations (S-representations) because, as showed above, they

[68] This is what Gładziejewski calls the 'compare-to-prototype' strategy, the argumentative strategy recommended by Ramsey (2007). The idea is that generative models posited by the predictive framework and cartographical maps share some functional features. These structures, however, do not need to feature the exact same characteristics. For instance, cartographical maps are allocentric representations of space, and they do not guide an agent by providing possibilities of action, among other differences (Gładziejewski, 2016, p. 569). Importantly, Gładziejewski notes, the paradigmatic representation and the posited representation might exemplify these features in a different manner (Gładziejewski, 2016, p. 564). These differences should not break the functional analogy established in this argument. Featuring the distinctive functional characteristics presented by the cartographical map warrants the label of representation for the recognition density.

track the causal structure of what is represented in agent-dependent terms.[69] Furthermore, it is in virtue of sharing the structural features of an agent-dependent environment that recognition densities can fulfil the role they do. The relation of similarity between the generative model and the world is, for Gładziejewski, an exploitable relation (Gładziejewski, 2016, pp. 566-567). The functional features attributed to S-representations are due to the structural similarity between the representation and the represented.

Let us start with i. action-guidance, the first functional feature of S-representations. This kind of representation serves to guide the actions of an agent. Maps, for instance, allow an agent to move from one point to the other. An agent knows how to find her destination by following a map (Gładziejewski, 2016, p. 567). Recognition densities also exhibit this functional feature. Think back to FEA's proposal. On the basis of predictions advanced by the model, the agent minimizes surprising dealings with the environment and achieves viability. With the notion of active inference, the FEA emphasizes the active character of predictions: the agent advances, as a prediction, a command for acting in one way or the other. At this point, it is important to recall that perceptual inference is subordinated to active inference (Gładziejewski, 2016).

Secondly, S-representations are ii. detachable. They can fulfil their role, i.e. guiding actions, without the represented being present or available. Detachability can be either complete or partial. Think back again to the case of the cartographical map. The agent guides her actions by following the map. Although her destination is not present, the information about the location of her destination is made available by the map (Gładziejewski, 2016, pp. 567-568). The agent decides which path to take based on the map. It is in this sense that S-representations are detachable. In a similar way, the agent's interaction is mandated on the basis of the recognition densities. Predictions advanced as a result of the predictive process are anticipations: non-actualized states of affairs. However, detachability, in both cases, is not complete but partial. Although a route is decided based on the map, it is verified in the world: either one gets to the intended destination or not. Similarly, the prediction advanced on the basis of the recognition densities meets actual sensory feedback and it is advanced as an anticipation of what will be encountered in the world.

[69] See also Williams & Colling (2018) for a discussion of iconic representations, which they define as representations that fulfil their functional role "through their resemblance to their target" (Williams & Colling, 2018, p. 1951). As Williams (2018) notes, in the case of the predictive framework, the idea is roughly that the representation and its target can be described by the same model.

Given that predictions are verified by the interactions of the agent, recognition densities are attuned to the environment and never fully detached (Gładziejewski, 2016, p. 577).[70]

Finally, S-representations allow iii. error-detection, i.e. the detection of misrepresentations. It is possible for the agent to detect an error in the representation when she encounters a practical failure: e.g. when a map leads to unsuccessful actions (Gładziejewski, 2016, p. 568). Recognition densities exhibit this functional feature as well because predictions advanced by the model are tested constantly. When the model delivers an inaccurate prediction, the interactions of the agent will be unsuccessful. Furthermore, error prediction is an essential feature of the prediction process. The recognition density is updated on the basis of error feedback. Error detection plays a fundamental role: the model is updated to achieve more accurate predictions and reduce the possibility of error (and of surprising dealings) over time (Gładziejewski, 2016, p. 579).

Arguments against the representational reading of the predictive framework

The representational interpretation of the predictive framework has been met with resistance in the literature.[71] Moreover, it is possible to provide a non-representational alternative reading to the functional features presented by Gładziejewski. Bruineberg et al. (2018) for instance, offer an account of error detection in terms of dis-attunement. To make their case for a non-representational interpretation of the FEA, they start by emphasising that the FEA should not be conflated with predictive processing. While predictive processing is a strategy advanced to model neural activity, the Free-Energy principle is advanced to understand the behaviour of a biological system. The same features that are attributed to the biological system are taken to explain cognition, broadly construed. For Bruineberg et al., the FEA offers a formalization of self-maintenance, an essential feature of biological systems from the perspective of enactivism, according to which biological

[70] According to Gładziejewski, generative models are characterised by partial detachability, in part, because they have been mainly used to explain cases of online cognition. So, it is an open question whether they can be fully detached from the environment, that is, whether they can be used to explain cases of offline cognition. For Gładziejewski, one step in this direction is precisely the model advanced by Friston and Seth, according to which, generative models include coding of counterfactuals. These models involve "counterfactual sampling" of non-actualized actions and could be thought of, in that specific sense, as completely detached (Gładziejewski 2016, pp. 576-577).

[71] Hutto & Myin (2017) argue that although it comes naturally to interpret the framework in this way, the proponents of the FEA must face the gaps left by taking such stance. If the FEA posits contentful representations, it must explain where the contents of such representations come from. I do not address this discussion in more detail because Hutto & Myin's is an objection that is general to representational frameworks, and not specific to the predictive framework. For an account of the contents of representations within predictive processing, see e.g. Wiese (2017b) and Kiefer & Hohwy (2018).

systems bring forward their own boundaries in their interactions with the environment. These interactions are ordered to the system's viability (Bruineberg et al., 2018).[72]

The FEA mandates a view of the biological system as one that is attuned with its environment and that, in order to maintain viability, requires a recognition density that "embodies in its structure and organisation longer-term regularities between action, environment and the state of the organism" (Bruineberg et al., 2018, p. 2425). The environment is reflected in the recognition density as it is affectively significant for the biological system, i.e. in terms of what the biological agent cares about. Furthermore, if we stick to the FEA's perspective, it is the biological system (and not the brain in isolation) that should be considered as the system in question. This notion of the recognition density is advanced in opposition to a perspective that takes the system to be or encode an objective and agent-independent representation of the environment, which is, according to Bruineberg et al., the notion of the generative model that is in place in Bayesian models of the brain.

With the presentation of Gładziejewski's arguments I do not aim to show that a representational reading of the predictive framework is mandatory. However, if the non-representationalist's concern relates more specifically to the positing representations that are *action-neutral*, the position sketched has the tools to address the worry. As Williams (2018) has shown, the representational interpretation of the FEA can address this worry because the commitment to representationalism does not imply a commitment to the positing of objective representations that stand for an agent-independent environment. This is also recognised by Gładziejewski: recognition densities are not an objective mirror of the causal structure of the environment. They rather *recapitulate* the causal relations as they are relevant for the biological system (Gładziejewski, 2016, p. 571; Wiese, 2017b). These structures are ultimately at the service of the system's coping with the environment (Williams, 2018).

Nico Orlandi (2014, 2016) has contested the representational interpretation of the predictive framework.[73] In particular, she objects to an understanding of vision as an inferential process. Her

[72] For an overview on the FEA's notion of autopoiesis as self-maintenance and how it compares to enactivism's notion of self-production as developed mainly by Humberto Maturana and Francisco Varela, see Kirchhoff (2018a).
[73] It is worth noting that Orlandi (2014) advances a proposal of perception that opposes that of SMEn. For her, while early visual processing should not be characterised as representational, the end-result of visual perception should be accounted for in representational terms.

point is that visual processing does not consist in “transitions between representational states” (Orlandi, 2014, p. 18). To show this, she argues that the states that are involved in the inferential process that constitutes the Bayesian view of vision should not be understood as representations. For Orlandi, the only key concept within the predictive framework that can be described representationally is that of the perceptual hypotheses (i.e. the result of the process, and not the process itself). Hypotheses concern distal or absent conditions, they can misrepresent, and they can be described as performance-guiding states.[74]

Orlandi notes firstly that when it comes to early stages of visual processing and to intermediate levels of the cascade of predictions that lead to perceptual hypotheses, characterising these states as representations becomes difficult. After all, these states (i.e. error- and prediction-signals) do not concern distal or absent causes at all, but the immediate lower level or what is immediately present to the senses (Orlandi, 2016, p. 344). On the other hand, in what concerns the recognition density, i.e. priors and likelihoods, Orlandi claims that similar considerations apply. Priors and likelihoods are not about distal conditions such as the causal structure of the environment. Rather, they are about other states at lower levels of the predictive hierarchy.

There are a couple of things to consider in relation to Orlandi’s objections. Firstly, her objections are directed at the Bayesian considerations of the brain without considering the FEA as a story about the adaptive behaviour of biological systems. One of the reasons she gives to inscribe predictive processing within ecological approaches to visual perception is that Bayesian approaches do not provide a story of how the space of probability from which predictions are advanced is bounded. Importantly, this is precisely the story offered by the FEA and the role of the free-energy principle.

Secondly, the FEA’s claim about the representational nature of recognition densities is that the structure of the recognition density represents the causal structure of the environment. The recognition density is constituted by variables that represent not only likelihoods and prior probabilities, but also the dynamic relations between variables at different levels. The proposal of the FEA is that the relation between the different levels, i.e. the structure of the recognition density, can be described by the same model that describes the system’s environment (Wiese, 2017b). So,

[74] Downey (2018) draws on Orlandi’s arguments. However, he rejects the claim that perceptual hypotheses are representational.

Orlandi might be right in that the process is not inferential because it does not involve transitions between representational states. However, that argument does not address the point made by the proponents of the representational reading of the FEA that the recognition density, as a whole, recapitulates the structure of the system's environment.

The arguments just sketched show how the representational account of the FEA can respond to the worries that drive anti-representational interpretations of the framework. To that extent they support the plausibility of the representational reading of the FEA. Nonetheless, it is important to note that these arguments are not conclusive since they do not show that the representational interpretation is mandatory. More would need to be said to make a decisive case for either a representational or a non-representational reading of the predictive framework. This has not been the purpose of this section. Again, for the purposes of arguing for the viability of a predictive approach to SMEn, it is important to take seriously the plausibility of the representational profile of the FEA because this represents an important obstacle for the joint account. If we can claim that SMEn is compatible with such an account, arguing for the viability of a non-representational predictive approach to SMEn comes easily.

A non-representational alternative

As mentioned, Bruineberg et al. (2018) also take issue with representational readings of the FEA and, more particularly, with the link between the representational interpretation of the FEA and the Bayesian brain. This is because the latter recommends a view of a brain that behaves as "an exemplary scientist", inferring from representational states the hidden causes of sensations, and not as a system that is biased in its recollection of the environment and its relation to it. Bruineberg et al. thus offer an alternative non-representational reading of the framework. They do not reject the possibility that some internal structures of the system are best understood in terms of representations (Bruineberg et al., 2018, p. 2439). However, they emphasize that the function of the generative process that instantiates the generative model is to "steer" the interactions of the cognitive agent to maintain its viability and structure (Bruineberg et al., 2018, p. 2440).

According to the authors' reading of Friston's (2013) position, the notion of 'inference' that is at play in the FEA is not as robust as it is typically taken to be. The biological system is not inferring from a rich model of the world a possible hidden state that lies beyond it. Instead, the FEA relies on a much more minimal notion of inference according to which the biological system and its immediate environment possess "high mutual information". According to this interpretation, rather than inferring from a robust and rich representation, a biological system coupled to its environment is "said to "infer" the "hidden cause" of its "input" (the dynamics of [its environment]) when it reliably covaries with the dynamics of [its environment] and [the biological system] is robust to the noise inherent in the coupling" (Bruineberg et al., 2018, p. 2436). Another way to put this is to say that there is mutual predictability from the biological system to its immediate environment, and *vice versa*. Moreover, the biological system can maintain this coupling discriminating from noisy information.

Even when this process can be read in representational terms, there is a more minimal reading according to which the process of achieving high mutual information is just a process of synchronization between two bi-directionally coupled systems. Systems whose coupling is best understood in this way are engaged in a synergetic circular causality, in which it becomes difficult to distinguish between the dynamics that are internal to the systems and those that are external. Bruineberg et al. illustrate this with a very minimal case of two synchronized clocks. The same system, they claim, can be interpreted in terms of Bayesian inference: "[b]y synchronizing, clock 1 can be interpreted as maximising its Bayesian model evidence for the state of the environment" (Bruineberg et al., 2018, p. 2438). The authors do not mean to claim that this is necessarily the case when it comes to biological systems and its environment. Whether the dynamics of the biological system and of the environment to which it is coupled are such that they can be described as *synchronized* depends on the specific features of the dynamics of the systems. Regardless, this leaves the door open to the possibility that there is nothing representational at play. Moreover—as a corollary—the strategy of relying on synchrony is not strange to enactive and ecological approaches. If this is the right interpretation of the kind of coupling that the FEA describes, the FEA sits closer to these approaches.

Before turning to the next chapter and the predictive approach to SMEn, let me briefly summarize the views presented here. In the second section of this chapter, I focused on the FEA, according to which biological systems order their actions in ways that avoid jeopardizing their viability and that lead to their flourishing. To accomplish this, biological systems should avoid finding themselves in situations that challenge their survival. This is achieved, so the FEA claims, by anticipating and predicting which states they should avoid. The brain's predictive dynamics are taken to be an instance of this: neural activity is driven not by bottom-up stimuli, but by top-down predictions.

In what concerns perceptual processing, for instance, the brain is taken to advance a prediction of the stimuli that is more likely to be received, based on a model. This prediction is compared against received stimuli. Error feedback is, then, processed to update the model from which predictions are advanced, and to assess its precision. The consequence of this is that not all error feedback is deemed equally precise. I emphasize this point because it is a crucial aspect of Seth's proposal for a predictive approach to SMEn, which I introduce in the next chapter.

Finally, it is worth emphasizing a point concerning the nature of the recognition density. As discussed, there are compelling arguments to interpret the recognition density as a rich representation that recapitulates the aspects that are relevant from the environment to the biological system or agent, and that it does so in agent-dependent terms. However, other theoretical alternatives are available according to which the coupling between the environment and the biological system can be read in a more minimal way.

Chapter 6. Predictive Sensorimotor Enactivism

In the previous chapter, I presented the main features of what I called the predictive framework, an umbrella term that refers to predictive processing, the thesis that the brain implements hierarchical predictive processing, and to the Free-Energy approach (FEA), a theory about the adaptive behaviour of biological systems. I also mentioned that, according to proposals for a predictive approach to Sensorimotor Enactivism (SMEn), this framework provides the elements that can offer an account of sensorimotor knowledge. Furthermore, according to Seth, this might provide SMEn with an account of the neural mechanisms that enable perceptual experience. In this chapter, I focus on the predictive approach to SMEn. I begin with some remarks about SMEn's position regarding the role of the brain. In the second section, I review the proposals advanced by Downey (2017) and Seth (Seth, 2014) for a predictive approach to SMEn. I particularly focus on Seth's proposal, given that Downey's advances a non-representational account. As I mentioned, I want to make plausible the idea that SMEn can be made compatible with the predictive framework, even when the latter is interpreted in a representational key. In the final section, I review and address the concerns that arise in connection with the compatibility of SMEn and the predictive framework, specifically, in relation to the positing of representations and the internalist profile that standardly characterizes the predictive framework. I show that the predictive approach to SMEn results in a fruitful collaboration.

6.1 Neuro-talk in Sensorimotor Enactivism

As I mentioned at the beginning of Part III, it is worth recalling that SMEn is a naturalistic theory of perception that aims at being continuous with a causal account of the mechanisms that enable perceptual experience. The causal story favoured by SMEn is one that accounts for perception in terms of the interactions of an embodied agent with her environment. This might explain something that has been pointed out in the literature: SMEn says very little about the neural mechanisms involved in perception (Chemero, 2009; Clark, 2009; Downey, 2017; Seth, 2014, 2015). It is important to consider that this omission is not seen as a challenge within the framework. After all, SMEn starts off from the claim that looking inside the head does not take us very far in our understanding of perceptual experience, and it emphatically denies that an account of the neural mechanisms of perception is sufficient to explain perceptual experience. Given that perceptual experience is constituted by the interactions of the agent with her surroundings, we should turn to these interactions to account for it.

Nonetheless, the theory does not deny that the brain is involved in perception. See, for instance, the following claim by Noë:

> The brain has a job to do (…) a careful examination of the way experience and the brain's activity depend on each other makes plausible the idea that the brain's job is, in effect, to coordinate our dealings with the environment. It is thus only in the context of an animal's embodied existence, situated in an environment, dynamically interacting with objects and situations, that the function of the brain can be understood (Noë, 2009, pp. 64–65).

Neural mechanisms, although not sufficient, are nevertheless necessary for perceptual experience. Furthermore, these should be understood within the wider context of the interactions of the embodied agent. As can be seen, taking this approach in what concerns the role of the brain in perceptual experience is already somewhat reminiscent of the FEA. I say more about this below. For now, just recall that according to the FEA, neural activity is an instance of the adaptive behaviour of biological systems.

But, does SMEn owe us an account of the way the brain is involved in perception? One might think that it does not. Surely the brain plays a role in supporting perceptual experience, but a story at the neural level does not place a strong constraint on our account of perception because perception is explained at the level of the whole organism and its interactions with the environment. Nonetheless, by claiming that the causal and the agential level are continuous, SMEn establishes a tighter link between the two levels. The causal story goes hand in hand with the story at the level of the agent’s interactions. And part of that causal story is about the neural structures that underpin perceptual experience. The causal story is not complete until some account of the role of the brain is provided.

This seems to be the thought behind the proposals for a predictive approach to SMEn. Predictive processing offers an account of the neural mechanisms involved in perceptual experience that might allow SMEn to explain how the brain contributes to perceptual experience. For instance, Downey (2017) appeals to predictive processing in order to respond to the objection advanced against sensorimotor approaches to perception that claims that they are empirically vacuous, while Seth (Seth, 2014) appeals to predictive processing to provide, among other things, a causal story that accounts for sensorimotor knowledge. In the next section, I discuss these proposals.

6.2 The predictive approach to Sensorimotor Enactivism

To outline the project of a predictive approach to SMEn, let me begin by reconstructing Downey's (2017) proposal. For my purposes, I am specifically interested in his arguments concerning the explanatory role fulfilled by predictive processing in connection with SMEn's account of perception.

6.2.1 Non-representational Predictive Sensorimotor Enactivism

Downey proposes to outline a predictive version of SMEn that can tackle the challenge of providing a complete causal story of sensorimotor knowledge. Moreover, he sets out to address a challenge faced not only by SMEn, but by any radical approach in cognitive science, i.e. any approach that avoids relying on representations to explain cognitive phenomena. Downey's (2017) goal in constructing a non-representationalist predictive approach to SMEn, then, is to show that radical approaches can provide an empirically sound theory of perception that takes into account the role of the brain in enabling it.

Downey is specifically interested in SMEn's account of perceptual experience. He argues that predictive processing can provide an account of both sensorimotor knowledge and attention, what Downey takes to be the key concepts in SMEn's account of perceptual experience. Given that the representational profile of predictive processing constitutes one of the greatest obstacles for a predictive approach to SMEn, Downey makes the case for a non-representational reading of predictive processing, drawing on the arguments advanced by Orlandi.

To understand Downey's position, it is worth recalling SMEn's key claims. According to this view, perceptual experience consists in perceivers' possession and exercise of the mastery of SMCs. An important concern arises in relation to SMEn's proposal of sensorimotor knowledge. The view assigns two different and conflicting roles to sensorimotor knowledge, namely, mediating perceptual experience and constituting it. Take, for instance, the case of the tilted circular coin. On the one hand, perceiving the coin is itself an exercise of sensorimotor knowledge. On the other hand, that the coin is perceived as circular depends on knowledge of the way the appearance of the coin changes after an interaction. So, how could sensorimotor knowledge fulfil these two roles? One possibility is to interpret sensorimotor knowledge as know-how. The other is to take it to be knowledge of counterfactuals carried by representations. Downey opts for the former interpretation—sensorimotor knowledge as know-how—which can more easily do justice to the practical character of sensorimotor knowledge.

Downey distinguishes between a personal and a subpersonal interpretation of SMCs. He follows a behaviorist interpretation of sensorimotor knowledge, according to which mentality just is a relation between sensory inputs and motor outputs. From this perspective, sensorimotor knowledge is concerned with such relations and the brain "acts merely as a causal mediator" (Downey, 2017, p. 3). Downey proposes to understand sensorimotor knowledge, at the personal level, as behavioural dispositions manifested in perceivers' behaviour. [75]

Something to notice about Downey's proposal is that he does not address specifically how his view accounts for the two roles that are meant to be fulfilled by sensorimotor knowledge. Recall that this is a problem that attaches to the know-how reading of sensorimotor knowledge.[76] It is in order to address this that Silverman (2018) proposes to interpret this knowledge as a capacity to exercise the embodied skills that are constitutive of perceptual experience. Moreover, as discussed, the two roles can be more easily fulfilled by interpreting sensorimotor knowledge in counterfactual terms. So, if we stick to the reading of sensorimotor knowledge in terms of know-how, more needs to be said about the allegedly conflicting roles sensorimotor knowledge fulfils in SMEn's project. I say more about this below.

Despite this shortcoming, Downey rightly notes the role that predictive processing is meant to fulfil in an account of sensorimotor knowledge: predictive processing can provide SMEn with a story about the contribution of the brain in enabling sensorimotor knowledge, it "can therefore be used to explain the specific relations between certain neural inputs and certain other neural

[75] For reasons of space, I do not offer a thorough reconstruction of the notion of attention proposed by Downey (2017). Roughly, he offers a non-representational account of attention. According to him, attention can be best understood, subpersonally, as biased-competition between "transiently-assembled-neural-local-subsystems" (Anderson, 2014) that are formed "on the fly" in response to the competing affordances that are offered by the environment.
[76] See also (Flament-Fultot, 2016).

outputs" (Downey, 2017, p. 10). We should add, though, that some constraints need to be in place to avoid making this position trivial. As I mentioned before, SMEn does not deny that the brain plays a role in enabling perceptual experience. However, there is a way in which this can amount to nothing more than claiming that, surely, the brain plays a role in enabling perceptual experience, but the exact details are irrelevant. This, as discussed above, is not SMEn's claim. For this reason, more needs to be said in favour of the compatibility between the views. The question is: what is exactly the brain doing such that it would support perceptual experience, as described by SMEn? Seth's more detailed proposal offers an answer to this question.

6.1.1 Representational Predictive Sensorimotor Enactivism

Recall that SMCs are the systematic patterns of co-variation between sensory information and movements of the perceiver or the object(s) perceived. Perception depends upon sensorimotor knowledge, the knowledge of SMCs. Seth (Seth, 2014) proposes to account for SMCs, and the mastery thereof, by means of the generative model. It is important to bear in mind that, according to Seth, for predictive processing, neural mechanisms are necessary and sufficient to enable perceptual experience because the content of perceptual experience is fully determined by the neural generative models. His proposal is to integrate SMCs into the generative models, thus providing the causal story that SMEn is missing (Seth, 2014, p. 98). Incorporating the theory of SMCs comes naturally for the FEA because the generative model includes the specification of stimuli that follow a given movement, which is crucial to implement active inference. In what follows, I briefly discuss Friston's et al. (2012) arguments in favour of this. They propose to integrate the views advanced by O'Regan & Noë (2001) as well.

Visual search movements

Friston et al. (2012) defend the view that active sensory sampling is one strategy deployed by biological systems to acquire reliable beliefs about the world. Active sensory sampling, the authors claim, is a behaviour that emerges from the engagement of a system in cycles of surprise minimization. As explained in the previous chapter, active inference can be thought of as a process in which a system brings about the sensory input predicted. However, this means that what the system learns is the way a certain movement can lead to certain sensory information. The information coded by the brain should be such that it supports this. According to the authors "this means that prior beliefs have to be encoded physically (neuronally) leading to the notion of fictive or *counterfactual representations*; in other words, what we would infer about the world, if we sample it in a particular way" (Friston et al., 2012, p. 2). Moreover, the authors add, this idea should take us from *predictive* to *prospective* coding, i.e. from advancing the prediction with the highest probability given current evidence, to entertaining possibilities about the different sensory states that could be brought about given the current situation of the system.

Friston et al. argue that this proposal follows naturally from the principles established by the FEA. Recall that, according to the FEA, biological systems aim at minimizing dealings with the world that would challenge their viability, i.e. they aim at minimizing surprise. Recall too that biological systems do not have access to external states. They are shielded or isolated by a Markov blanket. However, given previous interactions, systems 'infer' how certain sensory information correlates with the objects that cause it. One way to ensure finding unsurprising sensory states is to actively seek these states. This depends, in turn, on perceptual inference, since the latter allows systems to finesse their beliefs about the world. This rough outline describes the action-perception cycles of surprise minimization introduced in the previous chapter. What I want to emphasize here is that, to implement these cycles—to actively seek certain sensory states—the system requires information about the way sensory states change after a movement, i.e. it requires information "of how we sample our environment" (Friston et al., 2012, p. 4).

According to Friston et al., sensory sampling movements are involved in this cycle of surprise minimization. Perception is understood, in this model, as an exercise of "hypothesis testing". The idea is that hypotheses or percepts are advanced in order to select the one that provides the best explanation for the available evidence (Friston et al., 2012, p. 13). Based on previous information, perceivers deploy saccadic eye movements to sample visual input. Information sampled in these movements is not merely accumulated. Posterior beliefs are assigned varying levels of confidence throughout the process. The hypothesis or percept that endures is the one that can ensure that salient features are sampled. In other words, the hypothesis that endures is that which can predict a reliable process for sampling sensory information. The authors claim that, in this process, "the hypothesis prescribes its own verification and can only survive if it is a correct representation of the world. If its salient features are not discovered, it will be discarded in favour of a better hypothesis" (Friston et al., 2012, p. 16).

The FEA incorporates, thus, an active view of perception. By means of cycles of perceptual and active inference, the system enriches its generative model. The model includes information not just about the world, but about the way the system brings about certain features of the world. SMEn's claim that perceivers possess knowledge of the way sensory information changes after an interaction becomes more plausible in light of this proposal. In the next section, I present Seth's view, according to which, SMCs are encoded in the generative models.

The case of synaesthesia

According to Seth, the predictive approach to SMEn provides an operationalized notion of sensorimotor knowledge. To better understand his proposal, let us review the case of synaesthesia, which, according to Seth, can be explained by the predictive approach of SMEn.

Synaesthesia can be a challenging phenomenon to explain from the perspective of SMEn.[77] Consider a case of chromesthesia, in which a sound triggers a chromatic experience: e.g. a high-pitch sound like a whistle that triggers an experience of the colour yellow. SMEn struggles to account for these cases because the pattern of SMCs that would explain the experience of the inducer (i.e. the whistle) have little or nothing to do with the pattern of SMCs that would explain the experience of the concurrent (i.e. yellow). Given that these experiences occur simultaneously, SMEn must explain how it is that they are both underpinned by deploying a consistent repertoire of SMCs. Furthermore, it is necessary to explain that, in a view that takes perception to be a skill, the experience of the concurrent persists. In other words, if perceptual experience is explained in terms of the mastery of SMCs, then, the concurrent should adapt away as the agent becomes more and more skilful (Ward, 2012).

Seth offers a solution to this problem. According to the view he recommends, both the inducer (in our case, the whistle) and the concurrent (the colour yellow) are coded in the generative model, they are both underpinned by the same structure of SMCs. Moreover, this is meant to explain synaesthetic experience, an experience that is characterised by a difference in the veridicality of the inducer in comparison to the concurrent. Consider that, while the experience of the inducer involves objective veridicality, the experience of the concurrent does not: the synaesthetic subject

[77] For competing accounts of synaesthesia, within the context of SMEn, see (Hurley & Noë, 2003; Ward, 2012).

identifies the concurrent as an experience that accompanies the experience of the inducer, a veridical experience.

To explain this difference, Seth claims that although they are both encoded in the generative model, they are coded (i) at different levels of the model, (ii) with different levels of precision, and (iii) with different degrees of learning. Inducers (in our example, the whistle) are encoded in higher levels of the generative model and, thus, have a rich structure that results from learning. Concurrents (in our example, the experience of yellow) are encoded in intermediate levels because they are taken to be associated with secondary processing. The intermediate level from which the concurrent is predicted should be reshaped through learning, thus leading to the elimination of the experience of this non-existent stimulus. However, for Seth, synaesthetic experience arises in virtue of an unusually high level of precision of this intermediate level which resists reshaping and, thus, the experience of the concurrent persists (Seth, 2014, pp. 107–109).[78] It is the unusually high level of precision that does not allow the concurrent to be eliminated.

For Seth, the veridicality of the experience of the inducer is due to it being coded at a higher level of the generative model. This is due, in turn, to the fact that there is a greater history of interactions associated with the inducer, as opposed to the concurrent. The concurrent is coded in an intermediate level and lacks a rich history of interactions. The richer structure of the inducer in the generative model explains the veridical experience of the synaesthetic subject regarding the inducer.

Sensorimotor Contingencies and the generative model

Seth's proposal is that SMCs are encoded in the generative model, which, he claims, encodes counterfactual "perception-action couplings" that result in active inference. However, in light of the distinctions introduced by Ramstead et al. (2019) between the generative model, the generative process and the recognition density, it is necessary to make some adjustments to Seth's position. This is relevant to reject his claim that SMCs are exhausted by what, up to this point in this chapter, I have called the neurally implemented generative model.

[78] Seth's view is compatible with associative accounts of cross-modal and intra-modal perceptual experience. Inter-modal and cross-modal perceptual experiences result from associative processes that take place in learning and development that might be strengthen given genetic predispositions and structural features. What is characteristic of synaesthetic experiences is that these associations are resistant to reshaping (Seth, 2014).

According to Seth's approach, then, the perceiver deploys a set of SMCs as a result of top-down inferential processing. Recall the visual tracking example introduced before. In this case, the movement of the eyes and head as the trajectory of the bird is followed results from the prior belief about the position of the bird and generates a posterior belief from which a new inference—hence, a new interaction—results.[79] If we adjust this idea to the FEA, the thought is that perceivers deploy a generative process.

The *mastery* of SMCs, for Seth, relates to the "conditional aspect" of perception, i.e. to the counterfactual knowledge of, for instance, how the appearance of a tomato would change if I were to turn it around. This mastery is attributed to the hierarchical structure of the brain. Seth claims that this can be reflected in the layered structure of recognition densities: predictions are generated from a cascade of top-down processes that are associated with more abstract predictions (Seth, 2014). These predictions are observations that, in turn, serve the cycle of surprise minimization and are the basis on which perceivers adjust their interactions (Ramstead et al., 2019).

At this point, it is worth going back to sensorimotor knowledge and the two roles it is meant to fulfil. The problem was that the exercise of sensorimotor knowledge is meant to support both the *actual* deployment of skilful sensorimotor interactions, as well as the sense of presence that is framed in terms of counterfactual information. The predictive framework provides a way to support this. Moreover, this is consistent with SMEn's view according to which representations might fulfil an enabling role in relation to perceptual experience (see section 2.1). Sensorimotor knowledge was defined as the capacity to respond with sensitivity to the way sensorimotor information changes after an interaction. From the perspective of the predictive approach to SMEn, this is explained in the following way. Perceivers manifest this capacity when deploying a certain sensorimotor interaction that results from a process of surprise minimization. Here, the interaction is, itself, the result of the process. This explains that the process of surprise minimization supports the actual deployment of sensorimotor knowledge. In what concerns the sense of presence, this is explained by the cascading top-down process. According to Seth, the richness of perceptual experience comes from the process from which the prediction results. However, if we take the perspective of the FEA, the richness of perceptual experience comes from the full cycle of surprise

[79] Adapted from (Friston et al. 2012, p. 15; Seth 2014, p. 104).

minimization that incorporates the richness of neurally and bodily implemented recognition densities.

Let me take stock here. Seth's predictive proposal regarding SMEn provides a story about the neural mechanisms involved in perception. In this way, this account fills in the gap left by the question of the brain's role in perceptual experience. As discussed in the previous chapter, the predictive approach to SMEn faces several challenges. For starters, the predictive framework is typically read in a representational key. Even when there are non-representational alternative readings, the arguments advanced in favour of non-representational readings do not provide sufficient reasons to support it over the standard representational reading of the predictive framework. On the other hand, the standard view of the cognitive system advanced by the predictive framework as a system that is secluded from its environment is incompatible with SMEn's view that perceptual experience does not supervene on the activation of certain brain regions.

In addition to these challenges, the predictive framework faces a number of challenges itself. These are open empirical questions that I do not address here. However, it is worth returning to one important concern. As can be seen, the predictive approach to SMEn relies on the richness of the generative models to account for the sense of perceptual presence that accompanies perceptual experience. Here Kwisthout & van Rooij's (2013, 2020) concerns about the tractability of predictive processing become relevant, since the brain might not be able to support these rich generative models. In the next section, I come back to this point, since it is connected to a worry raised by Di Paolo (2014) in light of Seth's proposal.

6.3 Addressing the concerns about the compatibility of Sensorimotor Enactivism and the predictive framework

In this section, I address the concerns discussed previously. Let us begin with the concern that the representational profile of the predictive framework hinders any attempt at bringing together these approaches.

6.2.1 Representations

Before assessing whether the position on representations defended by the predictive framework is compatible with SMEn, let me briefly recall why I focus on this position. The predictive framework is more often than not read in a representation-invoking way. Establishing a marriage between SMEn and the predictive framework in its more common representational key is much less straightforward than establishing such a marriage when the predictive framework is read in its less common non-representational key. This explains why Downey (2017) opts for establishing the collaboration between a non-representational predictive processing and SMEn. Here, I aim at making the case for the compatibility between SMEn and what justifiably qualifies as the default understanding of the framework in the wider cognitive science literature.

It is worth recalling the position defended in section 2.1 of sensorimotor knowledge as a representational-hungry domain. Several arguments have been offered against the idea that these domains require representations. Nonetheless, positing representations remains one strategy available to treat these cases. Sensorimotor knowledge is, consequently, a domain that admits the positing of representations without requiring them. This is what I called, following Silverman (2018), a *prima facie* representational-hungry domain. Moreover, this is compatible with the view that representations might play an enabling role in what concerns perceptual experience. Recall too that SMEn constrains the role that representations might play in perceptual experience. Firstly, SMEn rejects the view that a representation is the end-result of perception. Secondly, it defends a position according to which, if SMEn were to invoke representations, these should not be neutral with respect to the actions and constitution of the perceiver.

Consider Orlandi's (2014) position. According to her, the end-result of perceptual processing—of perceptual inference—are representations. Her thesis is that while visual processing is not representational, i.e. it is not a transaction between states that are best described as representations, it produces representations. To challenge the plausibility of this view, I want to emphasize the circular character of the process of surprise minimization.

It is important to begin by noting that, according to the view defended in this thesis, it is the recognition density that is often conceived representationally. Furthermore, although the cycle of surprise minimization does imply feeding and adjusting this model, the creation and subsequent adjustment of said model is not the end-result of perception. Firstly, this is because perceptual inference (the adjustment of the model) belongs to a cycle of surprise minimization that has as its ultimate goal delivering successful interactions between the biological system and the environment. The adapted biological system is a model of the world that allows unsurprising engagements with its environment. In this model, perception contributes to the adequate engagement of the agent with the environment which allows the agent to maintain its viability. Moreover, perception instantiates the process of surprise minimization defended by the predictive framework. Recall here, too, that the generative model—instantiated in the generative process—refers to the dynamic coupling with the environment, i.e. the engagement by means of which the biological system ensures its viability. From this perspective, it becomes difficult to claim that the end-result of perception is the updating of the recognition density.

In what concerns the content of the recognition density, or put otherwise, the kind of information that is coded in the recognition density, consider firstly that this information includes sensory information that is expected to follow an interaction. The recognition density is thus indicative of further interactions as it presents the agent with possibilities of action. The recognition density is agent-specific in that it is in part the embodied agent herself that constitutes said model, and in that the neural implementation of the recognition density recollects the environment as it is significant for the system. Furthermore, the recognition density is not neutral to the agent's history of interactions; on the contrary, it is enriched by learning.

One final thing to note relates to the idea that sensorimotor knowledge is a *prima facie* representation-hungry domain. The predictive approach to SMEn proposes that sensorimotor knowledge is manifested in the deployment of sensorimotor interactions. Sensorimotor knowledge relies on the recognition density which is implemented both in the embodied structure of the agent and in neural activity. In consequence, the distinction between representational and non-representational structures does not correspond to a distinction between the brain and the body. In relying on the recognition density, perceptual experience is supported not only by neural activity. Rather, it is supported by the non-neural body too. Moreover, from the perspective of the FEA, these two explanatory strategies (i.e. relying on the embodied constitution of the agent and her interactions, and positing representations) are not incompatible.

As can be seen, the representational reading of the predictive framework does not represent a challenge to the viability of a predictive approach to SMEn. Let us turn, now, to the second concern.

6.2.2 Internalism

A second concern about the predictive approach to SMEn arises in light of the standard 'internalist' profile of the predictive framework.[80] Some versions of predictive processing recommend a view of the cognitive system according to which the cognitive system is secluded from its environment. In addition to this, some versions hold that perceptual experience supervenes solely on the brain. An example of the latter is Seth's position. However, this is incompatible with SMEn's position. In this section, I address this incompatibility. I begin by addressing a question about the content of perception that arises from the marriage of SMEn with the FEA.

The content of perception

In section 3.1, I argued that SMEn defends a version of direct realism according to which, in perceptual experience, perceivers enter into unmediated contact with the world. Perception is a relational state. The objects of perception are constituents of perceptual experience. I also claimed that I take this view to be compatible with the claim that this is enabled by representations, since some of the elements posited in our causal story might mandate such an interpretation. In what concerns the predictive framework, there are compelling arguments to support the view that the generative model is a representation. Furthermore, representational accounts are standardly interpreted as introducing a mediation between the perceiver and the world. So, a question arises, is perception—as described by the predictive approach to SMEn—a process by means of which we come into unmediated contact with the world?

Despite the compatibility between the representational-friendly profile of SMEn and the representational stance of the predictive framework, it seems that we run into problems here.[81] To see why let us go back to what the content of perception is, according to the predictive

[80] In this section, I mostly focus on the FEA, because it is the thesis that can more easily be made compatible with SMEn.
[81] See Anderson & Chemero (2019) for a version of this concern.

framework.[82] Recall that the process of surprise minimization instantiated by the cognitive system is such that it makes sense to describe it in terms of a Markov blanket. The cognitive system is shielded by a Markov blanket and only has access to sensory states. Under this description, it becomes challenging to say that the system perceives the world. Surely, the states that give rise to perceptual experience are about the world. But, it is not straightforward that the system perceives "the structured and meaningful world itself" but, instead, the "structured and meaningful world generated by the brain's own top-down predictions" (Kirchhoff, 2018b, p. 757). The latter would be Hohwy's position. For instance, he claims that "normal perception is (…) at one remove from the real world it is representing" (Hohwy, 2013, p. 138). Even when this response might seem like the most plausible given the principles of predictive processing, we can arrive at a more nuanced response. This is Clark's (2013b) view. According to him, even when the main source of the content of perception are perceptual predictions, these are constantly being revised against the world. In line with this, Clark claims that "such contents are constantly being checked, nuanced, and selected by the prediction error signals consequent upon the driving sensory input" (Clark, 2013b, p. 199).

According to Kirchhoff, when we consider the position advanced by the FEA, we can take this idea further. The relevance of the process of surprise minimization instantiated by the brain does not strictly imply that it is only the brain that is the source of the contents of perception. The view would be that whatever instantiates the process of surprise minimization comprises the vehicle of perceptual experience. As noted above, the recognition density is comprised by the whole embodied agent. So, it might be that, given the hierarchical structure of neural activity, the brain contributes substantially. Nonetheless, this does not imply that the brain is the sole vehicle of perceptual experience (Kirchhoff, 2018a, p. 760). The recognition density "need not be neurally instantiated but rather reflected in our bodily biomechanics, bodily shape, and so on" (Kirchhoff, 2018a, p. 762). Accordingly, the process of surprise minimization—as it is instantiated in

[82] It is worth introducing, here, a distinction advanced by Wiese between the *mathematical* content of the generative model and its *cognitive* content. To establish this distinction, he draws on Egan (2014). Mathematical contents are those ascribed to the states of a system as described by a computational model. On the other hand, cognitive contents serve an explanatory purpose (Wiese, 2017b, p. 721). According to Wiese, these are not independent. Although the mathematical content does not determine the cognitive content, it does constrain it (Wiese, 2017b, p. 723). Here, I am concerned with the cognitive contents assigned by the predictive framework.

perception—constitutes the process by means of which we come into unmediated contact with the world.

Explanatory strategies

Understanding the recognition density as comprised by the whole embodied agent goes against the idea that perception is at one remove from the world. But it does not go against the idea that sensorimotor knowledge is supported solely by the brain. This idea might seem implausible given the previous discussion. However, consider the following. According to Seth, sensorimotor knowledge depends upon SMCs coded in the neurally implemented recognition density. Another way to put this is to say that, for Seth, lawful patterns of co-variation to which the perceiver is sensitive are those that are encoded neurally. What is relevant is not the recognition density, but the hierarchical predictive processing that characterizes the neural process of surprise minimization. This explains that the perceiver exhibits the mastery and sense of presence that is characteristic of perceptual experience. Furthermore, we can claim that the only "adequate vehicle" that can implement hierarchical predictive processing is the brain. Note that this is contingent upon facts about human nature (Clark, 2009, p. 983).

Di Paolo (2014) offers a similar argument to Kirchhoff's, as reconstructed above. Di Paolo starts off from the claim that the relevant SMCs are not just neurally implemented. Here, SMCs and the recognition density come apart. According to Di Paolo, SMCs are better described by the generative model, as a virtual structure of interactions. Firstly, the recognition density includes information that is not about the systematic co-variations that constitute SMCs. Furthermore, given that the recognition density constitutes the priors from which the prediction is advanced, the recognition density would include SMCs that are already incorporated into the perceiver's repertoire, SMCs that are embodied. However, in addition to SMCs included in the recognition density, there are cases in which interactions that are neither the current interactions of the agent, nor appear in the agent's repertoire of possible interactions that belong to the recognition density, constitute SMCs as well. These are described by the generative model.

To understand Di Paolo's objection, we must note firstly that he distinguishes between a virtual structure of interactions and actual interactions. So, for instance, going back to the example about our experience of seeing the coin, we must distinguish the interaction currently taking place from other possible interactions that are implied in the current interaction: from the way the coin looks when we hold it in one hand, certain interactions are implied, e.g., that being shaped as a circle it can project any kind of ellipse onto the retina.[83] Di Paolo refers to the former as the actual structure of interactions and to the latter as the *virtual* structure of interactions.[84]

For Di Paolo, given that the recognition density requires regular updating to maintain a certain level of accuracy, it requires regular feedback from the world. The brain is insufficient because the world is necessarily involved and already constitutes, in part, the virtual structure of interactions. The virtual structure of interactions is not exhausted by the neurally implemented recognition density. In the interactions of the perceiver with the world, there is already a structure of possible states that can be visited by them. Some possibilities of action are available to the agent not in virtue of the counterfactual structure of the neural recognition density, rather they are implied in the agent's engagement itself.

To illustrate this, Di Paolo takes a case of dynamical operationalizations of SMCs: "if I walk on a slope there is a strong downward tendency for my movement, even when I walk uphill. This is real and does not depend on my having enough sensitivity to detect it. Most "nearby" trajectory options are implied in the enacted movement" (Di Paolo, 2014, p. 2). In this case, Di Paolo claims, the virtual structure is entailed by the movement itself. The movement of walking uphill on a slope entails a virtual structure that constrains the possibilities of actions to which we are sensitive. The neural recognition density does not suffice to explain why *these* possibilities of action are available and not others.[85]

Di Paolo objects to Seth's claim that the neural recognition density is sufficient to account for perceptual experience because it is insufficient to account for the virtual structure of interactions to which the agent is sensitive. Di Paolo's claim is that the available sensorimotor interactions

[83] I took this fact about the retinal image of ellipses from (Morales et al., 2020).

[84] From the perspective of enactivism, the environment is constituted by the virtual interactions that are available to the agent and that are actualized as the agent interacts with her surroundings. Kiverstein & Rietveld (2018), for instance, take the agent's environment to be virtual precisely because it is enacted by the interactions of the agent. The virtual structure of interactions are the states a system could occupy given its tendencies and current interactions. For them, the virtual conditions that constitute a system's environment are "the set of states of the agent-environment system that 'surround a currently actualised trajectory of states' (Di Paolo et al. 2017)" (Kiverstein & Rietveld, 2018, p. 152).

[85] Flament-Fultot calls this change in the sensorimotor patterns available a change in the "landscape of *possibilia*" (Flament-Fultot, 2016, p. 170 italics in original).

exceed those that appear in the neural recognition density. On the face of it, this argument is not too different from that advanced by Kirchhoff. Particularly, in one respect, this is still compatible with the idea that the neurally implemented recognition density supports the *mastery* of SMCs. There is a difference, though. While Kirchhoff claimed that the recognition density is not only neurally implemented, Di Paolo adds that SMCs cannot be exhaustively accounted for by referring only to the recognition density, even to the bodily recognition density. The possibilities of action offered by the generative model are already constrained by the agent's current interaction and the virtual structure of possibilities that emerges from it. Put differently, the recognition density does not suffice to explain *how* the world contributes to perception: the interactions of the agent already constrain the structure of interactions that constitute SMCs. This is the generative model.

This objection should not lead to the rejection of Seth's model. And, importantly, nothing in Seth's model prevents him from accepting Di Paolo's proposal of constraining the available sensorimotor interactions by way of the agent's embodiment and interactions. It might be objected that Di Paolo and Seth rely on incompatible views of sensorimotor knowledge. While Di Paolo is committed to a practical view of sensorimotor knowledge, Seth is committed to a counterfactual reading. However, as I have argued previously, sensorimotor knowledge should be understood as grounding perception, an embodied skill (see section 1.2). Sensorimotor knowledge is the agent's capacity to respond to possible sensorimotor interactions, and this capacity can be enabled in different ways: by the agent's embodiment or by internal representations.

This shows that, in part, the conflict that arises between the frameworks is due to their conflicting explanatory strategies.[86] While Seth seems to think that referring to the inner structure of the agent suffices to explain perception, Di Paolo claims that it does not because the virtual structure of sensorimotor interactions to which the agent is sensitive cannot be thus explained. It can only be explained by the generative model which includes, already, the interactions of the perceiver.

[86] In the vicinity of this concern is the worry that the predictive framework and SMEn commit to different models of explanation. This concern can easily be addressed by claiming that both views are committed either to mechanicism or to dynamicism. What is at play when we decide to hold one or the other is the nature of the cognitive system. Those who subscribe to dynamicism argue that there are mental properties attributed to cognitive systems that emerge from the behaviour of the system as a whole and over time. These are features that cannot be described when decomposing a system and studying just the causal relations between the elements of the system. I do not discuss this matter. However, one of the key questions within this discussion relates to the conception of embodiment that is in place in each of the frameworks. I discuss these two conceptions in chapter 7.

Di Paolo's claim can be summarized as follows: in order to have a full profile of the SMCs that constrain perceivers' sensorimotor interactions, it is necessary to refer to the interactions of the embodied agent. By accepting this claim, we are left with a view that is compatible with SMEn's explanatory externalism introduced in section 2.2. According to the position defended by Hurley (2010) and Noë (2007), the interactions of the agent in part constitute the explanatory substrates of perceptual experience and constitute, therefore, the realizers of perceptual experience. Furthermore, the model of perceptual processing that I have described in these chapters can be assimilated to the online simulation cases introduced by Hurley to build her case.

Where does this leave the predictive approach to SMEn? According to the arguments discussed in this section, neural dynamics are subordinated to and should be understood within the context of the agent's interactions. Deploying this strategy is consistent not only with SMEn, but also with the predictive framework described here. Although the latter is an approach that aims at accounting for neural activity involved in perception, this does not necessarily mean that it takes the brain to be explanatorily sufficient for perception. The predictive framework is compatible with this reading of the explanatory priority of the interactions of the embodied agent since it offers an account of the biological system that underlies its theory of perceptual processing. The claim that referring to the interactions of the embodied agent is necessary can be further qualified: referring to the interactions of the embodied agent is necessary to explain perception in the sense that, even when referring to the inner structure of the agent, the interactions of the embodied agent should throw some light on our understanding of the agent's inner (neural) structure.

The world is its own best model: contributions of embodiment

Let me address briefly one last point. In the last section, I argued that the explanatory strategies of the predictive framework and SMEn can be made compatible under the idea that some SMCs are not included in the neurally implemented recognition density, and yet still constrain the sensorimotor interactions of the perceiver. This position is still compatible with the claim that the *mastery* of SMCs is supported only by the hierarchical predictive processing implemented by the brain. The brain is such that it is the only adequate implementational basis of the kind of processing that relies on rich models. But is the brain the adequate implementational basis for hierarchical predictive processing? It is in this respect that predictive processing can be enriched by working alongside embodied approaches to cognition, such as SMEn.

As mentioned in the last chapter, hierarchical predictive processing faces an important challenge. According to Kwisthout & van Rooij (2013, 2020), the brain might not be able to support the rich models that would be necessary to explain higher order cognition within the predictive processing model. According to these authors, it is implausible for the brain to handle rich generative models: the probability distribution that comprises the generative model is too complex to be computed by the brain. This explains that the probability distribution has to be approximated. This is the role taken by the recognition density. Added to this is the worry that approximations might lead to results that diverge from the results one would obtain using Bayesian inference, as Kwisthout & van Rooij show in their analysis (2020, p. 185). This is, of course, a concern only if it is shown that the Free-Energy principle fails in that way. Furthermore, this is not something that can be decided aprioristically. Nonetheless, a concern remains about the capacity of the brain to process these extremely complex computations.

This is particularly challenging for the predictive approach to SMEn because these rich models do play a role in explaining the sense of presence that characterizes perceptual experience. One alternative, of course, is to scrap predictive processing altogether. Given that this is an independent thesis from the FEA, we could even retain the conception of the biological agent. However, this seems too radical an idea. After all, predictive processing is an exciting project that offers a good account of the anticipatory dynamics of the brain. Moreover, there is relevant evidence that supports it as a framework to explain neural dynamics (Walsh et al., 2020).

On the face of this challenge, we might consider the possibility that surprise minimization does not require the rich, reconstructive models that are allegedly neurally implemented. Hohwy (2019) contemplates this possibility. When considering Clark's (2016) proposal that the generative model is rich, but context-sensitive and poised for action,[87] Hohwy argues that doing without rich representations goes against the spirit of the predictive framework. His concern is that the predictive framework requires rich, reconstructive models given its conception of a world filled with uncertainty. Accordingly, the world, as it is encountered by the biological system, is a changing uncertain world. Building rich, reconstructive models is the best way to achieve long-term surprise minimization. If the system encountered an already structured world, this would not

[87] At the core of the discussion between Hohwy and Clark is the discussion about the epistemic position of the cognitive system.

be needed. In line with this, he claims that “[t]he only way to avoid a rich inner model is basically to represent the agent as cycling regularly and utterly predictably between situations with one- to-one mappings where quick- and- dirty solutions work. But such a world is one in which [predictive processing], and Bayesian inference, are not needed in the first place” (Hohwy, 2019).

Hohwy entertains another argument to consider the possibility that rich and reconstructive models are not necessary. According to him, the biological systems that the predictive framework is meant to consider are, often, simple organisms. It is unlikely that these organisms rely on rich, reconstructive models. Despite that, they implement surprise minimization (Hohwy, 2019, p. 201). The question would then be: what kind of models are at play and how are they embodied? The avenue to explore is the avenue typically explored by embodied approaches to cognition. The question turns out to be about the possibility of relying on the structure of the embodied agent and its interactions. In the case of perceptual experience, this is the structure modelled by, e.g., Buhrmann et al. (2013). This goes in line with Di Paolo’s (2014) proposal that the virtual structure of interactions—the generative model—incorporates SMCs and constrains the space of possibilities of perceivers’ sensorimotor interactions.

To tackle the concerns raised by Hohwy, Clark (2019) emphasizes that, in his model, precision weighting plays a key role. According to him, the different weights assigned to precision estimation (namely, the precision of the error feedback) balance the system’s reliance on information coming from the world (i.e. bottom-up processing) and the advanced predictions (i.e. top-down processing). The resulting model is one that is context- and action-sensitive. Moreover, according to this proposal, interactions are, at times, governed by the cues found in the world. Another way to put this is to say that the biological system does not always face a world filled with uncertainty. This is, at times, the case. And in these cases “prediction error penetrates deeper and deeper, requiring more and more neural (and perhaps extra-neural) resources to damp it down to the tolerances of anticipated noise” (Clark, 2019, p. 289). At other times, the resources required are kept to a minimum. Furthermore, this should help us tackle the concern of tractability, since “in the ordinary unfolding, [there is] no reason to think that the richness of the full world-model need actively be in play” (Clark, 2019, p. 290).

The recognition density, nonetheless, is still a rich model, but it does not reconstruct the world as such. Instead it recapitulates it as it becomes relevant for the biological system and its interactions.

Furthermore, the process of surprise minimization relies heavily on the system's embodied structure which comprises, too, the recognition density. There is a sense in which the body, from this perspective, is already a rich recapitulation of the world in agent-centred terms, i.e. from the perspective of the biological system. This rich model, however, is not always at play. More often than not, the system relies on its interactions and the possibilities that are entailed by them.

6.4 An overview of the predictive approach to Sensorimotor Enactivism

Let me now, briefly, summarize the view to which we have arrived. Perception consists in the exercise and possession of practical knowledge of the systematic patterns of co-variation between sensory information and possible interactions between the perceiver and the object(s) of perception. This is what has been called sensorimotor knowledge and can be characterized as perceivers' sensitivity to SMCs. On this version of SMEn, sensorimotor knowledge is supported by the repertoire of SMCs coded or recapitulated in the neural and bodily recognition density, as well as by the virtual repertoire of SMCs that is described in the generative model. In this way the view incorporates a virtual structure of SMCs, namely SMCs that are not (yet) part of the repertoire of interactions possessed by perceivers as elements that can also constrain the process of surprise minimization that characterizes biological systems. The navigation and incorporation of these virtual SMCs manifests, too, sensorimotor knowledge.

The predictive approach to SMEn that I have articulated draws on the FEA. From this perspective, perception is supported by the coding of SMCs both at neural and bodily levels. In that sense, perception is an instance of an embodied and enactive process, the process of surprise minimization. It is *embodied* in that the recognition density recapitulates the environment, and it is *enactive* in that it enacts or brings forth its own rules and parameters of constitution—the generative model. The recognition density coded in neural and bodily structures recapitulates the history of interactions of the system. Given that it is sensitive to different timescales, it includes the history of interactions at an individual and at an adaptive level.

This version of the predictive approach to SMEn departs from Seth's (2014) proposal according to which the coding of SMCs in the neurally implemented recognition density is sufficient to explain the mastery of SMCs that characterizes the SMEn's model of perception. If perception is constituted by the possession and exercise of implicit practical knowledge of SMCs, limiting the account to the way this is supported by the neurally implemented recognition density is insufficient. One needs to look at the way SMCs are embodied in the system and at the way SMCs are enacted. In other words, one needs to look at the generative model and the generative process.

This predictive version of SMEn departs, as well, from the proposal to operationalize SMCs advanced by Buhrmann et al. (2013) discussed in chapter 3.2. The departure from this view is more subtle, since it seems that the models proposed by the authors can be incorporated within the FEA's story. To understand why, let us take a look briefly at the general structure of these dynamical models. As discussed in section 3.2, Buhrmann et al. (2013) identify four kinds of SMCs that are described in the following models: the *sensorimotor environment*, the *sensorimotor habitat*, the *sensorimotor coordination*, and the *sensorimotor strategies*. As explained, SMCs fix, at least in part, the character of perceptual experience, and play a role in individuating sense modalities. By describing possible sensorimotor interactions of a system, these models describe sensorimotor regularities.

However, these operationalizations can be captured by the predictive framework. Although the models move from being less agent-centred to being more agent-centred, and capture slight differences in timeframes (Buhrmann et al., 2013), the variables from which they all start are the sensor states of the system, the position of the system, and the system's internal and active states. These aspects are reflected in the proposals advanced by the predictive framework. Think back, for instance, to Friston et al. (2012). The authors simulate sampling eye movements and model these taking into consideration the system's internal states (modelled in terms of predictive coding), sensory states and active states. These are modelled in terms of cycles of surprise minimization, with the aim of capturing the anticipatory dynamics of biological systems in accordance with the FEA.

In this chapter, I have argued in favour of a predictive approach to SMEn. To get a sense of the proposed view, I began by reconstructing the positions advanced by Downey (2017) and Seth (2014). I particularly focused on Downey's arguments about the support that the predictive framework can lend to SMEn. The predictive framework provides a view on neural dynamics that can be accommodated within SMEn's causal story. As a prelude to Seth's position, I reconstructed the arguments advanced by Friston et al. (2012). The authors provide a view that incorporates the active character of perception and interprets sensorimotor interactions in terms of a cycle of surprise minimization. Finally, I focused on Seth's position and his account of synaesthesia. According to him, SMCs are encoded in the neurally implemented generative model.

The fit between the two views is due, in part, to the anticipatory character SMEn and the predictive framework attributed to perceptual processing. However, a lot more work is needed to justify this marriage, particularly because these approaches seem to fall on different sides of the spectrum of cognitive theories in what concerns their position about the boundaries of the mind and about the need to posit representations. In the second part of this chapter, I discussed the compatibility of these views. Firstly, I argued that SMEn can be made compatible even with a representational predictive framework. Secondly, some interpretations of the predictive framework already take into account the flexibility of the boundaries of the cognitive system. However, I argued that there is an important aspect that is left out in Seth's position: the 'coding' of SMCs is not exhausted in the neurally implemented generative models. On the one hand, the generative model encompasses non-neural elements. On the other hand, there are systematic co-variations that are implied in the interactions of the system. The generative model—implemented neurally and bodily—recapitulates SMCs that are relevant to the organism, which are incorporated over time and through learning. Some SMCs might not be incorporated in the repertoire possessed by the system, and yet they can constrain the process of surprise minimization that is instantiated in perceptual processing. The mastery that is constitutive of perceptual experience relates to perceivers' sensitivity to SMCs. This mastery is manifested in the way perceivers manage their interactions on the basis of the SMCs as they are encoded in or recapitulated by the recognition density. This mastery is not only neural, but bodily.

Now, there is one aspect that needs to be addressed by the predictive approach to SMEn that relates to a feature of Buhrmann et al.'s (2013) model. This model captures the concreteness of the agent, for instance, it captures their skills and their most visited sensorimotor trajectories. So, one open question is whether the predictive approach to SMEn could capture something similar. This relates to the conception of embodiment that is at play in the predictive approach to SMEn. This is the focus of the following chapter.

IV. Predictive Sensorimotor Enactivism and Situated Embodiment

Chapter 7. Situated embodiment

In Part IV, I argue that there is a conception of embodiment that has not yet received much attention within the predictive framework and, in consequence, in the predictive approach to Sensorimotor Enactivism (SMEn). This is the notion of *situated* embodiment. Several analyses already note that the predictive framework can incorporate the fact that our social interactions contribute, in important ways, to cognition.[88] My aim here is to articulate a notion of situated embodiment and discuss the way it can be captured in the predictive approach to SMEn. In this chapter, I draw on Wheeler (2010b) to distinguish between two conceptions of embodiment: implementational and vital embodiment. I begin by discussing the conception of implementational embodiment. I, then, discuss an objection to hylomorphic views of entities (Di Paolo, 2020; Di Paolo et al., 2018; DiFrisco, 2014) to lend support to the claim that the views that subscribe to a notion of implementational embodiment risk being unsatisfactory. I present this objection to motivate a move to vital embodiment, which I then discuss. The first section aims to identify the conception of embodiment that is at play in SMEn, as it was originally conceived, and the notion of embodiment we would be left with in the Free-Energy approach (FEA).

In the second section, I draw on arguments advanced by feminist phenomenologists to articulate a conception of embodiment that captures the idea that the body constitutes someone's situation. According to these approaches, the body carries with it (or *embodies*) someone's historical, social, political, cultural and economic circumstances, in opposition to the view that someone is just embedded in these circumstances. The body is not just situated, it is the *situation*. I argue that, to fully meet the challenge advanced by Di Paolo et al. (2018)—namely, doing justice to the materiality of the body—it is necessary to incorporate this aspect into the conception of embodiment at play in the view of perception advanced. My claim is that, unless we incorporate the situated aspect of embodiment into the notion of vital embodiment that is at play in the FEA and, in consequence, into the predictive approach to SMEn, we risk offering an account that is unsatisfactory. I argue that this notion of embodiment is rightly captured by the concept of biosocial embodiment advanced by Miriam Kyselo (2014) and refined by Joe Higgins (2018). At the end of this section I come back to the conception of embodiment in the FEA.

[88] See (Friston & Frith, 2015; Gallagher & Allen, 2018; Kirchhoff & Kiverstein, 2019a).

7.1 From implementational embodiment to vital embodiment

When discussing Lambros Malafouris's (2004) work, Wheeler (2010b) introduces a distinction between implementational and vital materiality. Wheeler notes that, for Malafouris, in order to take material culture seriously one needs to be "systematically concerned with figuring out the causal efficacy of materiality in the enactment and constitution of a cognitive system or operation" (Malafouris, 2004, p. 55, as quoted by Wheeler (2010)). In this case, what is at stake is the materiality of artefacts and objects that, at times, are incorporated into our problem-solving routines. These artefacts occasionally take part in these routines and the question is whether, when, and in virtue of what they become constitutive elements of "cognitive systems and operations". Wheeler distinguishes between two different kinds of causal contributions that can be attributed to these objects: *vital materiality* and *implementational materiality*.

The distinction between these two senses of materiality can be understood in terms of the aspect of the material basis that carries explanatory weight. When a view subscribes to an *implementational* notion of materiality, the explanatory weight falls on the abstract features of the material basis in virtue of which it can fulfil certain causal roles. What matters is whether the material basis can support a certain causal role or implement a certain function. On the other hand, if a view subscribes to a notion of *vital materiality*, the explanatory weight falls not on abstract features that permit the implementation of a certain function, but on the nature of the material basis as such (Wheeler, 2010b). Unlike implementational materiality, the principles that govern the material basis are relevant in that they become determinants of the kind of cognitive processes a system sustains, rather than just constraints on them.

To mark the distinction between the conception of embodiment that can be at play in different versions of SMEn, I draw on this distinction. Given that I focus exclusively on embodiment, I refer to implementational materiality as *implementational embodiment* and to vital materiality as *vital embodiment.* Let me turn now to these two conceptions of embodiment.

7.1.1 Implementational embodiment

According to an *implementational* conception of embodiment, then, the body can be said to contribute to cognition in virtue of *abstract* features that are attributed to it. These abstract features allow the body to fulfil a certain causal role. Accordingly, the body matters for cognition in virtue of the causal role it fulfils.

Two views that exemplify this conception of materiality are the functionalist theory of the mind and the mechanistic theory of the mind. Consider, first, functionalism. According to Wheeler's characterization, functionalism holds that "a mental state is constituted by the causal relations that it bears to sensory inputs, behavioural outputs, and other mental states" (Wheeler, 2010b, p. 32). The explanatory weight, here, falls on the description of the causal relations between inputs, outputs, and other mental states. The body matters insofar as it is implementing a certain function.

Consider, now, the mechanistic theory of the mind. Mechanicism is, roughly, the view that a phenomenon is explained by providing an outline of the spatiotemporal structure that supports it or is "responsible" for its occurrence (Milkowski et al., 2018, p. 4). This outline specifies the components of the relevant spatiotemporal structure and its processes or operations. The focus of the view is on the composition and organisation of a mechanism (Bechtel, 2008, p. 4). The explanatory weight falls, then, on the description of this causal structure.

Importantly, this description can be as specific as needed for the relevant explanatory purposes. Milkowski et al. (2018) illustrate this with the mechanistic description of a pair of scissors. One might be interested in explaining e.g. the efficacy of the scissors in cutting a piece of rope. In this case, it might be relevant to have an explanation that includes details about the material from which the blades are made (Milkowski et al., 2018, p. 6).[89] Details of the body matter insofar as it is an instance of a certain mechanism. And these details might be mighty important for the mechanistic description of a structure. Consider, for instance, the concept of engineering tolerance. As Miłkowski (2016, p. 31) points out, engineers indicate the limits up to which a device can vary in order for it to be able to process a certain computation. A description that can be implemented only by a few systems has a very low degree of tolerance and the other way around.

[89] Milkowski et al. (2018) argue that 4EA approaches to cognition—what they call wide approaches to cognition—provide heuristic stories for the specification of mechanisms. In line with this overarching perspective of the mechanistic theory of the mind, the authors also argue that functionalist accounts should be considered to be outlines of mechanistic explanations.

Implementational materiality in embodied approaches to cognition

Now, there is a sense in which the implementational conception of embodiment continues to be at play in some embodied approaches to cognition. Furthermore, it seems that this is the conception of embodiment that is at play in SMEn as it was originally conceived. To articulate this further, let me turn to a concern developed by Mog Stapleton (2013) when making the case in favour of a more comprehensive notion of embodiment.

Stapleton focuses on the embodied turn in cognitive science and robotics. In these areas, she notes, researchers found a way of “decentralizing cognition” and minimizing the cognitive workload of the brain by relying on the interactions of a system with the environment and on its morphological features (Stapleton, 2013, p. 2). The body becomes central, from this perspective, for problem-solving tasks. It shapes cognition in that cognitive processes rely less on a central controller, and more on the way different body parts communicate among each other, coordinating in this way to tackle certain tasks. Moreover, cognitive processes rely upon and exploit the *morphology* of the system to efficiently tackle cognitive tasks. Rather than depending on computational processing, a cognitive task is “offloaded” to the body and the interactions of the system.

An example of this can be found in some cases of embodied robotics. The principles under which embodied robotics operate can be illustrated by Barbara Webb’s (1996) robotic crickets. Webb intended to replicate the behavior of a female cricket locating and following the song of the male cricket. She did so by reproducing certain anatomical features of the sensory mechanism that allows crickets to identify the location of a sound, since this is achieved in virtue of the cricket’s specific sensory mechanism, and not just by means of neural processing.

The cricket’s ears are located on its forelegs. If a sound were to come from one side or the other, it would arrive at one ear before it arrives at the other. The ears are connected to each other through the tracheal tube, and to the rest of the body through structures called spiracles. The sound arrives at each ear, then, not only externally but also through the tracheal tube that connects them. As it travels through the tracheal tube, the soundwave changes its phase. When the sounds reach the eardrum, depending on each of its phases, the waves reinforce or cancel each other. The amplitude of the vibration at each of the eardrums corresponds to the location of the sound. The larger the amplitude, the closer the sound source is to that ear (Webb, 1996, p. 64).

In addition to this, each eardrum is connected to one dedicated interneuron in the cricket's nervous system. Importantly, the two interneurons connected to the eardrums only fire in response to the right rhythm, the male crickets' sounds (Webb, 1996, p. 97). These neurons only fire when "their activation level reaches a certain threshold" (Wheeler, 2005, p. 202). The neuron that will fire first will be the one corresponding to the eardrum that has received the strongest input (i.e. the one with the largest amplitude). When a neuron is activated, the female cricket starts moving in the direction of the corresponding ear. These neurons only activate at the beginning of the sound. After firing, the neuron returns to its rest state. However, in the period before it returns to its rest state, the level of the activation threshold will be lower and the neuron will, therefore, fire again more quickly. Male crickets' song is attuned to this mechanism. This song must have the right pattern and the right pace to activate the neurons of the female cricket (Wheeler, 2005).

In virtue of the very specific mechanism just described, the female cricket is able to tackle the task of locating the male cricket. What I am interested in here is the contribution of the body to solving a certain cognitive task. Certain aspects of female crickets' morphology are central to the strategy deployed by the cricket in locating the male cricket. However, these aspects are relevant in virtue of the causal role they are fulfilling. To see why consider again mechanicism. According to this view, depending on our explanatory purposes, we might need a more specific or a more abstract description of the spatiotemporal structure that is responsible for a certain phenomenon. In this case, to explain the specific strategy deployed by the cricket to tackle a cognitive task, we require a very specific description, one that includes e.g. the location of the eardrums and the connection between them. What becomes relevant is the role the eardrums play in this mechanism.

It can be argued that a similar kind of embodiment is at play in SMEn, as it was originally conceived. Recall here that, according to this view, perception consists in practical knowledge of SMCs. As discussed in chapter 3, SMCs fulfil the important explanatory role of accounting for the character of perceptual experience. SMCs depend on certain features of perceivers' bodies, e.g. the way colour cells are distributed across the retina. However, what matters are the systematic *patterns* of co-variation that comprise SMCs. The aspect of embodiment that is relevant is structural or organizational properties of the system. The morphological features of the body of the perceiver matter because they make certain information available, and shape and sustain the skills that are constitutive of perceptual experience.[90] In this case, similar considerations apply as to Webb's case, since the explanatory weight falls on the pattern of sensorimotor interactions. Consider again, for instance, the case of TVSS devices. Whether these devices support visual experience depends on whether they enable sensorimotor interactions that can be described as homomorphic to the sensorimotor interactions enabled by the organic visual system.

Now, in both cases—embodied robotics and SMEn—the body is granted more and more relevance. Moreover, both seem to be steps in the right direction in that they both "decentralize cognition" (Stapleton, 2013, p. 2). So, what is it that remains problematic? To motivate the shift towards vital embodiment, it is worth understanding the hylomorphism objection, a criticism that is advanced to implementational materiality. Let me turn to this objection.

7.1.2 The hylomorphism objection

The hylomorphism objection involves a concern that a given theory conforms to a set of metaphysical assumptions that are inadequate to characterize the phenomenon of interest. According to this objection, some theories of, e.g., cognition or life might fail to be satisfactory in virtue of their subscription to a version of hylomorphism. This, according to James DiFrisco, arises from "the combined refusal to reduce life to its material substratum and the recognition that life cannot be any type of thing or substance distinct from matter" (DiFrisco, 2014, p. 501). Accordingly, the material base or substratum matters insofar as it instantiates a certain structure, form, or organization. The explanatory weight, thus, falls on the structure of the material base. Drawing on the work of Gilbert Simondon (1995, 2005), DiFrisco characterizes hylomorphism as the view according to which a substance is composed by two elements, its form and its matter (DiFrisco, 2014). This is, of course, the Aristotelian view of substances. DiFrisco does not intend to reconstruct this position, but instead notes, following Simondon, that it is a widely established position, and he aims at outlining certain problematic features.

[90] One might argue that Buhrmann's et al. (2013) definitions of SMCs move away from this conception of embodiment since these are models that are specific to a system. These models consider, for instance, which sensorimotor trajectories are prioritized by a system. However, the different models of SMCs provide a more or less constrained model or description of the pattern of sensorimotor trajectories of a system. The explanatory weight falls, in this case too, on the pattern of sensorimotor trajectories.

Simondon is particularly interested in the way hylomorphism is reflected in the discussions concerned with the individuation of a system. According to him, there have been two views typically advanced to answer that question: hylomorphism and substantialism. What both positions have in common is that they assume that the question about the individuation of a system must be resolved by referring to some feature of the system already individuated, e.g. its form or its matter, rather than to the *process* of individuation itself. The mistake committed by both views is that they start from a "“static” (…) conception of individuation" (DiFrisco, 2014, p. 503). In opposition to this, Simondon holds that individuation is dynamic. Moreover, in what concerns hylomorphism, he claims that by focusing on the abstract configuration of an individual, what is missed is the transaction between "organized, and organizing matter" (Di Paolo, 2020, p. 10). Simondon is committed, then, to the claim that it is the *process* of generation than brings about the individual.

Drawing on similar considerations, Di Paolo et al. (2018) reject views that have hylomorphic assumptions. They find an example of this in mechanicism as a model of biological and mental phenomena. According to Di Paolo et al., underlying mechanicism lies an old-fashioned conception of "active material systems" (Di Paolo et al., 2018, p. 35). Mechanicism relies on state determinism—the view that current states of a system are determined by past states. However, state determinism is only adequate to describe systems "with well-behaved interactions" and "situations of near thermodynamic equilibrium", i.e. systems whose future states can easily be predicted and that will not change spontaneously (Di Paolo et al., 2018, p. 35). To model a system, mechanicism relies on approximations. These approximations often fail when applied to systems that do not have these features and that, instead, are in situations far from equilibrium and "behave chaotically" (Di Paolo et al., 2018, p. 35). Simply put, the claim is that the tools on which mechanicism relies fail to capture the materiality of the body.[91] Ultimately, this leads to bad explanations. The models advanced will be inaccurate descriptions of the relevant systems.

[91] Here, my aim is not to make a knockout case against mechanicism nor to defend the claim that hylomorphism is, in fact, one of mechanicism's assumptions. Far more would need to be said for that purpose. Instead, my aim is to set the stage for the conception of embodiment that is at play in the FEA. This conception falls close to the one defended by Di Paolo et al. (2018) and Di Paolo (2020). Moreover, this sets the stage for the conception of situated embodiment that I discuss in the second section of this chapter. My main goal is to show that, unless a view does justice to the situated aspect of embodiment, this view risks being unsatisfactory.

Di Paolo et al. claim that they reject hylomorphism in favour of a view of open and active materiality.[92] In line with this, they claim that "[b]odies are active and organized material processes (…) Their becoming is ongoing and they are concretely embedded in situations, in space and time, and in relations to other bodies, not incidentally but fundamentally" (Di Paolo et al., 2018, p. 99). Systems characterized by open materiality, according to Di Paolo et al., not only depend on previous states in often unpredictable ways, but also keep track of or trace their past. These traces are manifested in the system's behaviour (Di Paolo, 2020, p. 23).

The conception of active materiality expressed by Di Paolo et al. (2018) and Di Paolo (2020) is a version of what I call, drawing on Wheeler (2010b), *vital embodiment*. This notion opposes *implementational embodiment.* So, let me turn now to vital embodiment.

7.1.3 Vital embodiment

When it comes to *vital embodiment*, the relevance of the body for cognition is not due to abstract features of the body, as is the case with the implementational conception of embodiment. Instead, the explanatory weight of the body falls on the *nature* of the material basis as such (Wheeler, 2010b). According to Stapleton (2013), the implementational conception of embodiment falls short of what can be said from the perspective of embodied cognitive science. To support her claim, she draws on the contributions and lessons that can be learnt from enactivism. According to her, the shift that characterizes these approaches is linked not only to the relevance given to the implementational base and the interactions between the system and its environment, but to the relevance of "the constitution of cognitive systems and the relation between their constitution and their interaction with the environment" (Stapleton, 2013, p. 2). To account for the constitution of the cognitive system, then, one needs to look at the process by means of which the system is constituted, and by means of which the system comes to be differentiated from the environment. It is this shift, in part, that allows enactivism to deliver a more comprehensive conception of embodiment.

[92] In the version of autopoietic enactivism defended by Di Paolo et al. (2018), the authors advance a series of formal concepts that are meant to capture this active materiality and the features that characterize living systems. According to them, living systems are precarious complex temporally rich systems engaged in open-ended processes and interactions.

One first step in the direction of vital embodiment is the subscription of enactivism to a strong version of the life-mind continuity thesis, according to which "life and mind share a set of basic organizational properties, and the organizational properties distinctive of mind are an enriched version of those fundamental to life" (Thompson, 2007, p. 128).[93] These are the organizational principles that allow us to explain the constitution of the cognitive system. However, this commitment only takes us so far. Recall that we are concerned with the relevance of the body in our explanations and in claiming that this relevance is due either to the *nature* or to the *abstract features* of the material basis of the cognitive system. If our claim is just that cognitive and living systems share a set of *organizational* principles, we are still relying on features that would seem to fit with the conception of implementational materiality introduced above.

Thompson (2007), however, develops a specific version of the strong continuity thesis, what he calls the *deep continuity thesis*. This thesis allows us to take a second step to show that it is the *nature* of the material basis that should concern us. According to the deep continuity thesis, life and mind share not only a set of organizational principles, but a set of phenomenological principles. Based on Hans Jonas' (1966) thought, he claims that "the continuity includes the subjective and experiential aspects of mental life as well as the cognitive aspects. For Jonas, certain basic concepts needed to understand human experience turn out to be applicable to life itself" (Thompson, 2007, p. 129). Here, I draw on Wheeler's (2011) reconstruction of Thompson's position. As Wheeler notes, to characterize the *nature* of living and cognitive systems, Thompson draws specifically on two of Jonas' concepts, namely needful freedom and self-transcendence. Needful freedom refers to the idea that living systems depend upon material exchanges—e.g. living systems need to eat—while also exhibiting freedom from their own material constitution. Living systems are characterized by their constant "material turnover" that is central to metabolic processes: e.g. turning food into energy. Following Jonas, Thompson claims that the metabolism of an organism is the first manifestation of the system's "immanent purposiveness". Organisms, by means of their metabolic processes, establish their own norms of subsistence and constitution, and "assesses" whether external conditions are good or bad for them. Accordingly, there is a sense in which metabolic processes are already self-affirming processes (Thompson, 2007). Consequently, organisms are characterized by going beyond their current organization and states.

[93] See also (Godfrey-Smith, 1994; Wheeler, 1997).

Self-transcendence refers to this idea that living systems constantly go beyond themselves. In that sense, organisms are already *projective* (Wheeler, 2011).

Unlike the strong continuity thesis, the deep continuity thesis is committed to a movement of mutual enlightenment between life and mind. Accordingly, it is not only that by looking at the organizational principles that govern the constitution and maintenance of living systems we come to learn about the cognitive systems. Instead, it is also that by looking at the features of our mental lives we can come to learn something about living systems and their material basis. In this case, for instance, we learn that the material basis of living systems is projective, i.e. it exhibits a feature typically attributed to our mental lives.

These concepts allow Thompson to characterize living systems as those that bring about their own *material* boundaries and constitute themselves as a *material* unity. According to these concepts, although living systems are always "realized by a finite collection of material elements", they can never be fully identified with these elements, since they are always in the process of turning over material (Wheeler, 2011, p. 158). This ultimately leads to a different conception of embodiment. The explanatory weight concerning the contribution of the body to cognition falls, here, on the nature of the material basis: i.e. on the way this material basis constitutes, bounds, and individuates the cognitive system.

7.1.4 Embodiment in the Free-Energy Approach

It is fairly easy to see that the FEA is committed to the strong continuity thesis. The FEA is an overarching theory of life and cognition that focuses on "the interdependency of brain, body, and local environment" (Kirchhoff, 2018a, p. 2524). To understand why it is worth recalling some details of the FEA.

Life and mind are governed by the same principle, the free-energy principle. The FEA is a theory about the adaptive behaviour of biological systems. It departs from the claim that biological systems are open systems that minimize disorder (or dispersion) by making "inferences about environmental states" (Kirchhoff, 2018a, p. 2523). Moreover, these inferences are made on the basis of an *embodied* recapitulation of the world, namely the recognition density. The inferences made by the organism are constrained by a model that is implemented both neurally and in the body of the living system. In previous chapters, I referred to this as the cycle of surprise minimization: a cycle in which by means of active and perceptual inferences, living systems maintain themselves within boundaries that do not challenge their viability. The system seeks unsurprising states, and, in this engagement, it continuously updates its model. To use a phrase coined by Kirchhoff to capture the strong continuity thesis, the FEA is committed to the view that life "foreshadows" mind. Key features of cognition already make an appearance in the process of self-preservation of living systems. In this context, cognition is "a relational feature" of living systems that enables them to "improve conditions of self-maintenance" (Kirchhoff, 2018a, p. 2530).

According to the view, the materiality of biological systems takes on a new relevance. Consider, for instance, a point made by Gallagher & Allen. For them, the FEA approach offers a view of a brain that "is uniquely equipped to exploit the finely tuned properties of an organism's dynamic morphological body and associated *Umwelt*" (Gallagher & Allen, 2018, p. 2636). The material basis of the living system matters, so to speak, because it *embodies* the model from which the system advances predictions of its sensory and active states, namely, the recognition density. Furthermore, the material basis of the system matters because, via active inference, it *enacts* the normative model of the states that ought to be occupied by the living system, namely the generative model. In its interactions with the environment, living systems enrich and update their models. Importantly, this means that living systems change themselves. The models from which predictions and actions are generated, then, are not enriched in the sense that systems gather more and more information. Instead, models are enriched in that the system becomes more attuned to the environment. Active inference involves a rearrangement of both the living system and the environment. To emphasize that the living system and the environment are co-defined, Gallagher & Allen propose to draw on the notion of situation and claim that the movement of a living system is a movement of the situation. Accordingly, "a rearrangement of objects in the situation is a rearrangement of [the system] as well" (Gallagher & Allen, 2018, p. 2638). In virtue of this, they refer to the cycle of surprise minimization as a process of *predictive engagement*.

This also points to the deeply social character of embodiment. The recognition density keeps track of regularities in the environment as they matter for the system, as well as routines of interactions to engage with that environment. Our brains keep track, then, of the routines of interactions as we engage with others and as we engage in social and cultural practices. This is noted by Kirchhoff

& Kiverstein, who claim that "[t]he stable and persisting regularities that our brains are able to track are made through our activities, in particular through the activities people engage in when they take part in social and cultural practices" (Kirchhoff & Kiverstein, 2019a, p. 94).

Let us take stock. I started this section by discussing two conceptions of embodiment—implementational and vital embodiment—that take different approaches to the question of why the body or the material basis of a system matters for cognition. To motivate the move from implementational embodiment to vital embodiment, I drew on the hylomorphism objection advanced by DiFrisco (2014) and Di Paolo et al. (2018). According to this objection, a view might risk being unsatisfactory when it commits to a conception of the cognitive system that fails to do justice to certain key features of its material basis. I, then, moved on to discuss vital embodiment and argued that this is the conception of embodiment that is at play in the FEA. The body matters for the FEA not only in its capacity as the implementation basis of cognitive processes, but because it embodies and enacts the models that enable the system's self-maintenance. From this perspective, active inference just is embodied inference. The last point in this story, though, turned to the social aspect of this process. The system keeps track of its interactions and routines as they are developed in a social world. To make sense of the social aspect of embodiment, we need first to articulate a conception of embodiment that captures an interplay between the organic and the social. So, let me follow suit and, in line with Thompson's approach, look in the next section at what phenomenology can tell us about this aspect.

7.2 Situated embodiment

A concern about the social aspect of embodiment is often present in analyses from the feminist and phenomenological traditions. In opposition to the views that see the body both as an obstacle for knowledge and as the realm of the feminine, feminist phenomenologists focus their reflection on challenging the idea that it is a disembodied mind that acquires knowledge. Furthermore, these traditions highlight the fact that the body that matters is that of a *situated* agent.[94] By this I mean

[94] Edmund Husserl (1989, para. 36) introduced the distinction between the physical body (*Körper*) and the lived body or the body as it appears in experience (*Leib*) (see, for discussion (Wehrle, 2020)). Most well-known figures within the phenomenological tradition have drawn on and finessed this distinction. Although, Martin Heidegger is mostly

that these analyses aim at showing that our mental lives are shaped by a body that is the concrete locus of our social, political, economic and historical circumstances. Thus, the idea is not just that the body happens to be embedded within a specific set of social, political, economic and historical circumstances. Instead, the claim is that the body is governed by social principles and norms, as much as it is governed by organic principles. In this context, the body is understood as a condition of possibility of experience rather than just a feature of experience. Importantly, one key aspect of the analysis advanced by feminist phenomenologists, as will be discussed, is the idea that the body is a *situation*. Their claim, then, is not only that we are embedded in a social world, but that we come to embody this world. To make sense of this idea, let me turn to Dorothea Olkowski's analysis of embodiment.[95]

7.2.1 Embodiment as an intersection of the vital and the symbolic

To make sense of the notion of embodiment, Olkowski (2017) starts off from an analysis of Maurice Merleau-Ponty's (1963) distinction between three milieus or orders of phenomena, namely the *vital*, the *physical,* and the *symbolic*.[96] According to her, the phenomenological notion of embodiment lies at the intersection of two of these domains: the symbolic and the vital. While conditions that belong to the vital milieu limit the organism, what is relevant for its survival is not fully explained by accounting for the functions of each of the physical parts of the body. Instead, it is necessary to look at the aspects and elements of an environment that are *significant* for an organism. This corresponds to the *vital* milieu, which Olkowski characterises "the plane of the virtual" (Olkowski, 2017, p. 9). By this, she means that, in this domain, the possibilities for engagement that are available to the organism become significant and, in this way, the organism brings forth or enacts a meaningful world in accordance with these possibilities.

silent about the embodied character of experience, in the *Zollikon Seminars*, for instance, he claims that the boundaries of the lived body surpass those of the physical body (Heidegger, 2001, p. 112). However, it is Merleau-Ponty who gives embodiment the most relevance by claiming that the body is the *locus* of experience. It is on this notion of the body that De Beauvoir draws and on the basis of which she develops the idea that the body is situated. Given that our experience is deeply temporal, our circumstances are marked in the body. I come back to this point in the next chapter.
[95] Olkowski's aim is to provide an analysis of intuition as a phenomenological method of analysis, from the perspective of feminist phenomenology. She offers an analysis of women's embodiment as a step towards this.
[96] Merleau-Ponty (1963) aims at distinguishing the forms or regularities that can be observed in each of these milieus or domains without collapsing them, and without making them wholly independent from each other.

The milieu of the vital is, nevertheless, overwhelmed by structures that go beyond it, structures that belong, instead, to the symbolic milieu, the realm of language and thought. This is what Olkowski calls the domain of "nascent intelligibility" which structures the vital. This nascent intelligibility is noticeable in our perceptual experience of structured wholes. Moreover, this occurs without the interference of thought, or of "logic or language". What is given to us in perceptual experience is given already structured.

This explains why Olkowski locates embodiment at the intersection of the symbolic and the vital. The world appears in perception as something that makes sense and that matters for the organism, and as something that indicates possibilities of action for the individual (Olkowski, 2017, p. 11). What is vitally significant for an organism emerges from its interactions. The body is, from this perspective, "a structure of sensations and movements" (Olkowski, 2017, p. 13). The phenomenological notion of embodiment stands, then, as the bridge that links the vital and symbolic realms, since it reveals what is in the world, i.e. the objects and the environment with which an organism interacts, and this is revealed as already structured and as meaningful for the organism.

7.2.2 Temporality in embodiment

Olkowski adds another element to the story: temporality. Considering this aspect is, for her, crucial to understanding how embodiment links the vital and the symbolic, and to account for the structure of perceptual experience. For her, the structure that is given in perceptual experience should be explained by alluding to the role played by past interactions in conditioning our current interactions. Stimuli received become relevant on the basis of the history of interactions of an individual, which is summoned as the individual prepares for action.

What is relevant, though, is not only that the body summons this history of interactions, but that the actual interaction is a moment of *creativity*. In that sense, it is not only the past that is relevant, but the future as well. While the object of perception appears determinate and structured, the multiplicity of possibilities of action that can be set in motion appear as "an embodied indeterminacy" as a result of the affective character of experience (Olkowski, 2017, p. 14). Drawing on Henri Bergson (1988), Olkowski claims that in experience the body either becomes visible or is given from within in the form of affections, what she calls the memory of the body

(Olkowski, 2017, p. 15). Affections have an element of indeterminacy, since they are located between the multiple possibilities that are received from the outside and the multiplicity of possibilities of action that can be set in motion.[97] By emphasizing the affective character, we can also emphasize the idea that experience manifests the perspective of the organism.

From this phenomenological analysis we can gather that meaning arises from an interplay between three aspects: (1) that the world appears to us in terms of possibilities of engagement (i.e. what Olkowski names the virtual), (2) that, in this encounter, our own body appears to us affectively, and (3) that this all appears symbolically or linguistically structured.[98] Nevertheless, what is crucial for our purposes is the *interplay itself*. *Situated embodiment* names the structure of this temporal encounter between what is offered in the present by what Olkowski calls the sensorimotor mechanisms—i.e. the structure of sensory stimuli—, and the possibilities that are offered by the recollections of the body—manifested, in turn, in the structure of affections that accompany perceptual experience. The interactions that take place are, then, open as a project from the encounter of current and past possibilities.

As mentioned, the interplay between affectivity, virtuality, and language characterizes the analysis of the temporality of experience that is central to the phenomenological tradition. Consider Merleau Ponty's claim that the subject constitutes temporality (Merleau-Ponty, 2002, pt. III), or Husserl's (1991) and Heidegger's (1996) idea that consciousness is temporal. This temporal character of our mental lives not only makes experience possible, but it is meant to explain the way things appear in experience. The mark of mentality is, from this perspective, that we are directed towards a world. For Heidegger, for instance, when engaging with the world, we encounter certain possibilities of interaction that become available given our socio-historical, economic, political and personal circumstances. That some of these possibilities are actualized becomes evident as they are thematized in our discourse, as we speak of them or incorporate them in what Merleau-Ponty and Olkowski call the symbolic realm.

[97] Affectivity is crucial to understanding embodiment. Indeed, some emotions are the clearest cases of embodied mental phenomena. However, I will not say more about affectivity as such because my main focus is on perception and the relevance of situated embodiment for it.

[98] The idea that existence is characterized by these three structural moments has been articulated by Heidegger. The three moments he identifies are attunement or affectivity, understanding or projection, and discourse or language (Heidegger, 1996, para. 29,31,34,38).

7.2.3 The body as the situation

But, where is the body in that structure? The idea of the body as the situation—and not as merely situated—is that our circumstances are marked in our body. De Beauvoir, for instance, claims that our body "is the instrument of our grasp upon the world, [and] a limiting factor for our projects" (De Beauvoir, 1997). Consider Linda Alcoff's (2006) articulation of what she calls visible identities, e.g., race and gender. Alcoff understands visible identity as an interpretive horizon, which encompasses both the public—i.e. how one is visible to others—and the subjective character of the self—i.e. how one sees oneself. Identity is an *interpretive* horizon because its two aspects constitute an open and dynamic perspectival location from which the subject engages with the world (Alcoff, 2006, p. 95). The body shapes experience in a non-deterministic and non-transparent way in that:

(a) It provides us with motor and spatial metaphors that shape our thinking (Lakoff & Johnson, 1999).

(b) However, the way each individual understands these bodily metaphors depends on one's body image which results from "tacit bodily knowledge [practiced movements and mannerisms], (…) perceptual orientation and a conceptual mapping that determines value, relevance, and imaginable possibilities" (Alcoff, 2006, p. 107).

(c) All these factors influence the position of the body when dealing with different tasks that require integrated and unconscious bodily movements. Our body is meaningful not only because it constrains our social interactions and creates social expectations imposed on us, but because it already manifests a perspective. The body constitutes the meaningful frame of our subjective experience.

The claim that the body is the *situation* is certainly different from claiming that the body is *situated*. The latter is a claim about the body being embedded in a set of circumstances. As we will see in the following section, that idea risks granting the body a mere causal role in cognition. Claiming that the body is itself the situation, instead, denies the claim that the body is merely an interface, i.e. the boundary where information is exchanged. The body is shaped by the convergence of certain circumstances. The individual carries these circumstances with her in that the history of interactions of the individual, her social, historic, economic, political and personal circumstances are *sedimented*—in different ways, in different measures, and at different levels—in the body.

7.3 The biosocial body

Can the notion of vital embodiment present in the FEA do justice to the idea that the body is a situation? And does this notion tell us something about the FEA's conception of embodiment? I think the answer to both questions is positive. However, this is not straightforward. The difficulty becomes evident when we look at the tension between the social and the embodied turns of cognitive science. This is articulated by Kyselo (2014) who argues that the tension is due to the relevance each approach grants to the body and to the social.

7.3.1 The tension between vital and situated embodiment

One way to see how the problem arises is to think of the social and the biological as two different domains, each providing a set of principles and norms, which are taken to govern the constitution and maintenance of the system. The problem arises when bringing these two sets of governing principles together. This is what Kyselo (2014) has termed the body-social problem: the problem of reconciling the sets of norms and principles that arise in virtue of the body's belonging to what appear to be different spheres or dimensions of meaning.

Kyselo is specifically concerned with the notion of the human self when posing this question. She asks about the way "bodily and social aspects figure in the individuation of the human individual self as a whole" (Kyselo, 2014, p. 2). The challenge encountered when addressing the question of the constitution and preservation of the self is to reconcile the different factors that seem to play a role in this. The self is, ultimately, at the intersection of sensorimotor, affective, physiological and cultural processes. The problem here lies in the implicit assumptions about what counts as the individual and about the principles that are at play in the individuation of the self. We can put this more generally and ask instead about the role each domain plays in the constitution and maintenance of a system.

One possible reply to this is to point to the myriad of factors and layers that play a role in the individuation of the self. For Kyselo, simply stating these is not helpful in addressing the matter precisely because what is in question is the coherent unity that constitutes the self, as well as a question about the interaction between the different factors.[99] Kyselo notes that the tension has not been relieved in the attempts to address the individuation of the self that can be found in the embodied and social turns of cognitive science. The different views associated with these two approaches make assumptions about what counts as an individual system. These views have typically emphasized either the social or the embodied. They have opted for one of the following claims:

(a) The embodied perspective opts for *the minimal self claim,* according to which there is a core self, individuated on the basis of biological principles that rule the organization, maintenance, and individuation of the system. The social is *contextual* in the sense introduced in chapter 1. Social aspects are contextual factors because they might produce variations in the self (De Jaegher et al., 2010, p. 443). However, there can be a self without the social (Kyselo, 2014).[100] If we were to put this in terms of the distinction advanced in the previous section, this view holds that the body is *situated* rather than the *situation*.

(b) The social turn opts, instead, for the *social as constitutive claim*, according to which the social is “the primary source of individuation” since it is constitutive to the process that produces the self. Moreover, the self relies for its constitution on social processes and could not be such without them (Kyselo, 2014, p. 4).[101] This view falls closer to claiming that the body is a situation, but misses the idea that—according to the phenomenological analysis, at least—the body is still an encounter of the social and vital or biological realms.

Kyselo advances a biosocial account that aims at reconciling these claims. The goal is to avoid, on the one hand, dropping the distinction between the individual and the environment—which is the claim advanced by embodied approaches; and, on the other hand, to avoid claiming that the individual is an “isolated being parachuted into the social world” (Kyselo, 2014, p. 8). Let me turn to the biosocial account.

[99] See also, (Higgins, 2018).
[100] See, e.g. (Zahavi, 2008, 2010).
[101] Kyselo notes that there are very few examples of this position. She attributes this position to (De Haan, 2010).

7.3.2 Biosocial embodiment

Embodiment as an enabling condition of the self

To advance a notion of the body that reconciles both aspects of experience—the vital and the social—Kyselo picks up on Jonas' concept of needful freedom. She distinguishes the notion of needful freedom as it is manifested in organismic individuation, from the notion of needful freedom that is manifested in the individuation of the self. She calls the latter social needful freedom, and the former organismic or embodied needful freedom.

From the notion of organismic needful freedom, Kyselo develops what she calls the principle of individuation *through and from the world* (i.e. social needful freedom). She characterizes this principle in the following way: "an individual identity reflects, in its structure and existential needs and concerns, the world from which it continuously emerges; but, in order to exist as an individual, it thereby also emancipates itself from the world through those very same processes" (Kyselo, 2014, p. 8). Self-identity is, for Kyselo, a social phenomenon all the way down. The self is never fully given or determined. Instead, it is in constant and continuous enactment. For its constitution, the self gets all its 'material' from the social world and, from there, the self is demarcated. The self, then, arises "from a sea of social relational, not merely bodily processes" (Kyselo, 2014, p. 9).

Something crucial to the understanding of the individuation of the self is the relation between social and embodied freedom. Kyselo notes that we should not leave out the fact that self-individuation presumes the organic or bodily individuation of the system from the environment, i.e. it presumes that there is a system that is body-bounded or individuated as a body. However, this feature of embodiment is not what delivers and explains the emancipation that would be characteristic of self-individuation. Previous enactivist proposals of self-individuation, according to Kyselo, have made a mistake, though, in what concerns their characterization of the movement of freedom from the body. For Kyselo, there are always some moments of complete emancipation or freedom. These are moments in which self-identity is bracketed from the body. The process of self-individuation "escapes the body". The body plays a role as the *means* and *mediator* for this process of self-individuation, since it serves to "monitor social engagements and relations with the goal of social homeostasis" (Kyselo, 2014, p. 12).

From this perspective, embodiment is not constitutive of selfhood. If we look back at the distinction between constitutive, enabling and contextual elements of a phenomenon advanced by De Jaegher et al. (2010), we can argue that embodiment plays, instead, an enabling role in the individuation of the self. Recall here that an enabling condition is one whose absence prevents the relevant phenomenon from occurring. The problem with this approach, however, is that it ends up being unsatisfactory if one is looking—like I am—for a notion of situated embodiment that captures the thought that the bodily and the social are deeply intertwined. Kyselo is right in characterizing the individuation of the self from the social in terms of needful freedom. However, it would not capture just yet, for instance, Alcoff's view that the body constitutes the locus where this movement between emancipation and participation takes place.

The biosocial self

Higgins (2018) notes that there is still one step missing in Kyselo's analysis. His assessment differs slightly from mine, in that he claims that Kyselo's account grants a contextual role to embodiment. I think Kyselo's claim is stronger than that. Kyselo recognizes instead the enabling role of embodiment. Accordingly, embodiment is necessary for the individuation of the self, but it does not constitute the self. In fact, for Kyselo, the individual self escapes its embodiment. Furthermore, as Higgins rightly notices, this brings back the body-social problem.

In contrast to Kyselo's position, Higgins advances a notion of the biosocial self, which he characterizes as "a mode of being in which the constitutive biological-bodily and social processes (i.e. 'biosocial processes') of selfhood are non-decoupleable" (Higgins, 2018, p. 446). One way to put this is to say that the process of organismic needful freedom is already social. To support this claim, Higgins looks into well-known evidence from developmental psychology that looks, for instance, at newborn babies' capacity to imitate others. In addition to this, we can also lend support to his position by looking at a recent proposal about the relation between mother and fetus during pregnancy. Martínez Quintero & De Jaegher (2020) develop an analysis of pregnancy from the perspective of autopoietic enactivism. They argue that, at a certain developmental stage, there is already a kind of minimal intersubjectivity between mother and fetus. They claim that: "[f]irst experiences in the fetus are already confronted with the mother's alterity in minimal ways, both locally, as a moving and rhythmic world; and globally (and later in development), when mother and fetus engage as two agents, from their own perspectives" (Martínez Quintero & De Jaegher, 2020, p. 14).

Sociality, as one might put it, is already in place and present in the moment of constitution of the bodily self. Higgins argues that "[a]s far as humans are concerned, our biological bodies are socially saturated from the first moments of life through to our very last moments, meaning that 'bodily' and 'social' processes should both be considered 'bodily-social', or biosocial processes" (Higgins, 2018, p. 448). Moreover, living systems are in what Di Paolo (2020) describes as *perpetual becoming*. Consequently, in this perpetual movement of self-constitution, the social and the biological are co-constitutive. Another way to put this is to claim that embodiment is social all the way down.

With the notion of biosocial embodiment in place we can now go back to the FEA. The question about the reconciliation of the two conceptions of embodiment—the vital and the situated—in one concrete body remained. To answer this, I drew on Higgins' (2018) position. Vital embodiment is deeply social because in the perpetual becoming of the individual, the body is already "saturated" by the social. How is this reflected in the FEA? The process of surprise minimization manages our social and cultural practices. And it does so on the basis of a model that is shaped by the specific circumstances we inhabit. To incorporate the *biosocial conception of embodiment* to the FEA, then, we must take the process of surprise minimization to be "saturated" by the social.

Let me briefly summarize where we are. In this chapter, I have articulated three conceptions of embodiment: the implementational, the vital and the biosocial conceptions of embodiment. I have suggested to make sense of the conception of embodiment at play in the FEA in terms of *biosocial embodiment*. In relation to the FEA, I began by reviewing the conception of *vital embodiment* that is present in the FEA. According to this conception, the material basis of the living system matters because it embodies the system's model of the world from which it generates the predictions and interactions that ensure its preservation. This is a model that reflects the perspective of the system. On the basis of this embodied model, i.e., the recognition density, the system enacts its norms of preservation. As mentioned, the FEA already imbues embodiment with a social aspect. This is noticeable in Gallagher & Allen's (2018) and Kirchhoff & Kiverstein's (2019a) analyses. The world that is modelled is a social world, and our interactions are, more often than not, social and cultural practices.

In contrast, according to the conception of *situated* embodiment reviewed, it is not only that the individual is subject to certain organizational principles that bring about its constitution and material boundaries with the environment; the idea is also that this material basis—the body—is marked by past social circumstances of the individual. Although this social aspect is already present in the FEA, I invoked the analysis from feminist phenomenology to emphasize some features of situated embodiment: (a) that the situated body stands at the intersection of the vital and the symbolic, (b) that the body constitutes the situation, and (c) that, since the body is the locus of experience, the situated character of experience is irreducible. Drawing on Alcoff (2006), I argued that this is sedimented in the materiality of the body. Situated embodiment names, ultimately, the concrete consolidation of our social, political, economic and historical circumstances, as they have appeared to and have been reconfigured by the agent herself. Finally, I drew on Higgins' analysis to claim that the conception of embodiment present in the FEA already reflects the idea that the social saturates the body of the agent.

Of course, at this point we might wonder why we are, suddenly, discussing what constitutes the self and how to reconcile the different dimensions that are relevant in its individuation if we are concerned with a theory of perception, namely SMEn. The thing is that, when it comes to perception, a problem reminiscent of Kyselo's body-social problem remains, since we might wonder whether and how social factors shape perception. I address this in the next chapter.

Chapter 8. Situating perceptual experience in the predictive approach to Sensorimotor Enactivism

In the previous chapter, I discussed that a problem that arises by advancing the *situated* and the *vital* conceptions of embodiment is that, in what concerns its individuation, the system—agent, organism or person—seems to be pulled in different directions and governed by different principles. This is what Kyselo calls the body-social problem. To address this problem, I drew on Higgins' conception of the biosocial problem, according to which the *vital* individuation of the system is already saturated by the social. My proposal is to understand this "saturation" in terms of situatedness. In what concerns the Free-Energy approach (FEA), what is "saturated" by the social in this way is the process of surprise minimization. This is reflected in the idea that the embodied model from which predictions are advanced is deeply social or situated, since it reflects the perspective of the system, and it summons in a creative way its (adaptive and individual) history of interactions. Moreover, to accommodate this conception of embodiment, the process of surprise minimization must reflect the social, political, economic, historical and personal circumstances of the individual.

In this chapter, I focus exclusively on gender identity[102] as one aspect of the situated character of perceptual experience. As such, according to Higgins, gender identity serves to illustrate key features of the biosocial body, in that it reflects the way in which certain social norms constrain bodily subjectivity while, at the same time, bodily activity serves to maintain certain social norms (Higgins, 2018). My aim in this chapter is to outline one plausible way to accommodate this conception of embodiment in the predictive approach to Sensorimotor Enactivism (SMEn), and to make plausible the claim that perceptual experience is social all the way down, i.e. that perceptual experience is a situated phenomenon.

---

[102] I refer to gender identity to name, not only the social construct, but someone's attitude towards it. In what follows, I intend the terms "woman" and "feminine" to include both cis and trans women. Further, the analysis I offer does not imply a commitment to a binary conception of gender.

8.1 Gender and the situated body

In the last chapter, I reconstructed Olkowski's analysis of the situated body. I particularly focused on three of its features: (a) that the situated body stands at the intersection of the vital and the symbolic, (b) that, as the locus of experience, it has an irreducible character, and (c) that it constitutes the situation. Accordingly, the social, historical, political, economic, and personal circumstances inhabited by someone are marked in the body. We are never just embedded in a given context. Instead, this context shapes the material basis of our experience. One consequence of this view is that someone's social circumstances shape her mental life in a very particular way. This consequence is already noted by De Beauvoir. She claims that:

> [A]ll perception in general, presupposes an indefinite past lying behind oneself, a "communication with the world that is older than thought," which is made concrete by the fact of my birth. My history is incarnated in a body that possesses a certain generality, a relationship with the world prior to myself, and that is why this body is opaque to reflection (De Beauvoir, 2004, p. 163).[103]

If this analysis is on the right track, our situation should be, in consequence, reflected in the way we carry ourselves, in the way we relate to others, in the way we think, and, of course, in the way we perceive the world. Iris Marion Young offers an analysis of the way the situated character of the body is reflected in comportment. To make sense of how it might be reflected in perceptual experience, I begin by reconstructing her position.

8.1.1 Iris Marion Young's analysis of feminine comportment

To address how the situated character of the body is reflected in perceptual experience, let us begin by reviewing Young's (1980) analysis of feminine comportment. Her view is useful for our purposes because she discusses how gender identity shapes the way, e.g., women carry themselves in the world and execute certain tasks and activities. I am not so much interested in Young's analysis of femininity *per se*. Moreover, Young's analysis could be extended to other aspects of

[103] Here, De Beauvoir quotes Merleau-Ponty (2002).

someone's situation that are embodied in a similar way to gender.[104] My interest here is to understand the way these aspects of our situation shape our mental lives, particularly perceptual experience.

A difference in comportment

Young draws on arguments from both De Beauvoir (1997) and Merleau-Ponty (2002). Her aim is to show that women's situation is embodied and enacted in their *comportments* towards and interactions with the environment, i.e., in the way someone carries herself in the world, or, put differently, in the way we use our bodies to interact with the world and with others. According to Young, these comportments are situated; more precisely, they are gendered.

Young's point of departure is her analysis of Erwin Straus' observations of the difference in the way women and men throw a ball. Straus noticed that there is a clear difference in the comportment of someone, depending on their gender, both towards the world and towards space. For him, there is a clear 'feminine' comportment towards the action, a distinction he aimed to make intelligible, and he argued that this difference boils down to something like an essential difference. While Young dismisses this idea of an essential difference, she remains interested in the difference in comportment of the men and women.

Unlike Straus, Young aims to show how this difference of, what she calls, the "modality of the lived body" of men and women is due not to a biological difference, but to a difference in situation. Drawing on De Beauvoir, she argues that there is a unity that grounds this typical feminine comportment. However, this is not the unity of an essence but a unity of situation, a situation that is, of course, common only to a group of people living in a specific time and context. She defines an individual's situation in terms of the circumstances that shape existence. For her "[e]very human existence is defined by its situation; the particular existence of the female person is no less defined by the historical, cultural, social, and economic limits of her situation" (Young, 1980, p. 138).

According to Young, it is possible to identify as a modality of existence, i.e., a way of being in the world, a set of structures that constrains the way a group of individuals orient their body in the

[104] One open question is what other social identities are similar to gender in that they are manifested in bodily comportments. Alcoff (2006), for instance, considers that race is similarly situated. In addition to race and gender, we could perhaps think too of certain disabilities (see, e.g., Oliver, 2009).

world. As mentioned, she is interested in the modality that is typical for women. Young examines movements that are typically observed in women and finds certain features common to them. For instance, the way women restrict certain movements, or how they position their feet for certain tasks, etc. For her, this is due not to physical differences between individuals of different genders. Rather, it is due to the way they each use their bodies for certain tasks. Unlike women, men "[s]ummon the full possibilities of [their] muscular coordination, position, poise, and bearing" (Young, 1980, p. 142). Women, instead, do not summon these possibilities. These are, surely, available to them in virtue of the physical structure of their bodies: their strength, their shape, etc. And yet these do not appear as possibilities, they are not part of their counterfactual repertoire. For instance, driving a punch with the shoulder is not part of the repertoire of women when they throw a punch. According to Young: "a space surrounds them in imagination which we are not free to move beyond; the space available to our movement is a constricted space" (Young, 1980, p. 143).

The difference between feminine and masculine situated bodies, as it is indicated by Young, can be articulated in terms of their respective virtual structure of possibilities. Let me make sense of this by using some of the conceptual tools proposed by SMEn. This virtual structure refers to the different possibilities of interactions between a system, an organism or a person, and its (her) environment. As discussed in section 3.2 in connection with the different kinds of sensorimotor contingencies (SMCs), we can describe the different trajectories (in that case, sensorimotor interactions) a system might traverse. Moreover, depending on the variables we consider, the space of possibilities changes, and we can describe a more constrained virtual structure. Furthermore, we can consider variables such as habits and skills to constrain this space. Let us go back to Young. For her, this space of possibilities is shaped in part by gender identity. Even when, given morphological features of their bodies, women could run as men do, this interaction does not appear as part of the repertoire of possibilities that they would typically visit.[105]

[105] We might wonder, here, whether this is not true of other possibilities. For instance, consider the kick of a professional karateka. It is definitely different from the way I would kick. I have no training and no knowledge of the way I should move my leg with as much strength and speed as possible, without hurting myself. My repertoire of possible interactions is clearly different from that of the karateka. But how is this different from the distinction between the possibilities encountered by people with different gender identities? Young expects to find a difference in the virtual space of two professional karateka of different genders. This prediction might be right. But why? The question here concerns the way different aspects of our situation interact.

To analyse the difference in comportment between men and women, Young develops a set of existential categories that describe women's lived body.

The existential categories of women's situatedness

Young develops three existential categories: (i) *ambiguous transcendence*, (ii) *inhibited intentionality*, and (iii) *discontinued unity* (Young, 1980). These provide a description of certain constant features of women's experience. Her starting point is Merleau-Ponty's articulation of the body as the locus of subjectivity and intentionality, and which is analysed in terms of the different capacities it exhibits. The key idea is that, in providing possibilities of action—e.g., grasping this apple, turning to see the window, walking towards that door—the body opens the world to a subject. Put differently, the world appears to a subject in the orientation of her body and in the intentional actions that are enabled. It is in this way that Merleau-Ponty locates subjectivity in the body. In line with this, Young says that "[i]t is the body in its orientation toward and action upon and within its surroundings which constitutes the initial meaning giving act" (Young, 1980, p. 145).

Let me briefly discuss the existential categories identified by Young.

(i) *Ambiguous transcendence.* The feminine body is for Young characterized by an *ambiguous transcendence* in that it is "laden with immanence" (Young, 1980, p. 145). For Young, something characteristic of women's existence in a patriarchal society is that they live a contradiction in that they live in between transcendence and immanence. In other words, they live in between a structure that characterizes human existence—transcendence—and a structure that characterizes inanimate objects—immanence.[106] Very roughly, according to the phenomenological analysis, unlike inanimate objects, human beings are characterised by transcendence in that, in different ways, we always go beyond our current state and circumstances. Now, there is a sense in which this ambiguous transcendence belongs to every lived body: lived bodies are not mere transcendence in that they can also be approached and acted on like inanimate objects. According to Young, for Merleau-Ponty the body moves in "an open and unbroken directedness" (Young, 1980, p. 145).

[106] This distinction also tracks the Leib-Körper distinction mentioned in a note in the previous chapter. While the body of experience (*Leib*) is characterised by the movement to transcendence, the physical body (*Körper*) is characterised by immanence.

The transcendent character of the body is manifested in that the body acts upon the world, invoking and exercising its capacities.

What is characteristic of the feminine body, though, is that it does not manifest this movement from immanence into transcendence. Instead, it "remains in immanence" (Young, 1980, p. 145). For Young this is noticeable in that a woman's lived body appears at times as a burden. The feminine body is experienced as an object to be protected and guarded, and not an enabler of possibilities.

(ii) *Inhibited intentionality.* According to this category, the body is the subject of intentionality and this is manifested in motility. The lived body unites in and directs bodily motion towards a certain aim. Moving towards the accomplishment of this aim gives structure and order to the body's activities. Following Merleau-Ponty, Young claims that, in its movements, the body exhibits intentionality. Moreover, the possibilities that appear to the subject are those that are related to the subject's intentions. Whether and how these possibilities appear also relate to the subject's frustrations and limits. The lived body fluctuates between an "I can" and an "I cannot".

The feminine lived body is no different in this respect. It frequently reflects both the "I can" and the "I cannot" but in a particular way given her situation. When the feminine lived body manifests inhibited intentionality, the intention appears not as her intention, but as someone's intention. Young claims that, in these cases, women project an "I cannot". The emphasis is in the impossibility of owning that intention. For instance, given certain social norms related to gender, a certain action might not appear as a possible object of a woman's intention.

(iii) *Discontinued unity.* This category describes the body's "unifying and synthesizing function" (Young, 1980, p. 147). In moving with an aim, the body unites to its surroundings. These are organized with respect to the body and its intentions. Moreover, the body organizes itself in directing towards an aim: it is the body as a whole which is involved in reaching a certain goal or performing a certain action. Young claims that women's lived body stands in discontinued unity with its surroundings. This is manifested in the lack of participation of women's whole body in an action. While, e.g., an arm might be engaged in a certain action, e.g. throwing a ball, the rest of the body remains immobile or unengaged. Women act as if fragmented: actions do not involve the body as a whole. It is in this sense that the body is discontinuous with itself.

What is characteristic of the lived body, according to Merleau-Ponty, is that it exists directed towards the world of possibilities, unlike objects that exist in themselves, undirected. However, what becomes characteristic of the lived feminine body, according to the descriptions of Young, is that it often appears as a thing, as a body characterized by immanence. It often fails to open the world because, in many cases, these possibilities do not appear as possibilities available to her. Instead, when engaged in certain activities, what is revealed is a body that is perceived as fragile or incapable. This is reflected in the self-reference manifested in feminine movements. In line with this, Young claims that: "feminine bodily existence is self-referred in that the woman takes herself as the object of the motion rather than its originator. Feminine bodily existence is also self-referred to the extent that a woman is uncertain of her body's capacities and does not feel that its motions are entirely under her control" (Young, 1980, p. 148).

These existential features of the feminine lived body have consequences for the way space is experienced. The body opens a world of possibilities: the phenomenal space becomes apparent in relation to the body's possibilities of movement and action. But, if this is so, the phenomenal space experienced by women is also different. According to Young, there is a double spatiality in women's experience of space. There is the 'here' and the 'yonder', that is, a distinction between the space of her possibilities and a space that stands beyond the reach of her capacity, although not beyond someone else's reach.

The way women relate to space, which is, ultimately, reflected in their performance of certain tasks, originates from this relation to their bodies. Moreover, inasmuch as motility is crucial for the development of other cognitive and perceptual capacities, Young predicts that these existential categories will impact women's performance in related cognitive and perceptual tasks (Young, 1980, p. 149).

Beyond feminine comportment

One problem with the Merleau-Pontian analysis of the body concerns the neutrality of the description.[107] It is a description of a body that moves unconstrained out of its immanence, into fluid action, and into continuous application of its capacities. Young's analysis shows that when

[107] Although Merleau-Ponty's view of embodiment has been appropriated and further developed in feminist phenomenological analysis (Oksala, 2006, p. 231), it has also received important criticisms, see e.g. (Butler, 1989).

this lived body is feminine, this is not the case. The feminine body stays rooted in immanence. The same could be said of other identities, though. Is the masculine body rooted in immanence in some other way? And, how about bodies that are perceived, for instance, under a whole other set of damaging prejudices, or bodies that are perceived to exceed certain potentialities?

Just as it is expected that the feminine lived body is manifested in the way a woman performs certain tasks or engages in certain actions, other embodied social identities—and other aspects of identity—might be reflected in someone's comportment too. The situated aspects of the biosocial body are reflected in individuals in different ways. For instance, someone's racial identity might shape her experience in a way that reflects a similar fluctuation between the 'I can' and the 'I cannot', or reflect an experience between the 'here' and 'yonder' that is relevant.[108] Something similar is noted by Nick Brancazio (2019), who is mainly interested in the role of gender in the sense of agency. She argues that:

> [T]he idealization of an "I can" body as representative of the experience of men is problematic. The internalization of patriarchal norms limits the bodily comportment of men in different ways. For example, Western men are expected not to show weakness or fatigue. They are discouraged from using their bodies to express certain emotions. Thus, men's bodies are subject to limitations due to gender norms in a way that provides lived constraints on their bodily comportments as well (Brancazio, 2019, p. 436).

Katharine Jenkins' (2018) norm-relevancy account of gender identity comes in handy here. For her, gender identity should be understood in terms of an individual's "sense of their own 'locatedness' with regards" to social gender norms (Jenkins, 2018, p. 730). What is central is whether someone experiences a norm as relevant to them, and this includes too the possibility of contravening this norm. This is in line with one important aspect of situated and biosocial embodiment mentioned in the previous chapter, namely what Olkowski characterizes as *creativity* and Kyselo calls *emancipation*. According to this idea, although social norms, including gender norms, are part of our social context and partly constitute the material from which an identity is formed, there can be movement of emancipation and creativity in which, for example, someone moves away from these norms.

[108] I take racial identity to be a social construct by which we assign meaning to mere bodily differences (Taylor, 2004).

8.1.2 Pain and the situated body

To further support to the view that perception is situated in the full-blown sense we have discussed here, I want to briefly revise Gillian Einstein's (2012) analysis of female genital cutting (FGC). In her analysis, we find an example of the way the situated aspect of the biosocial body is reflected in a specific mental state. Einstein is specifically interested in the central nervous system and how it might change as a result of changes in the non-neural body (e.g., the removal of one part of the body), in this case the change caused by FGC. Her view and approach is consistent with the conception of biosocial embodiment. In line with the previous discussion, for Einstein, the central nervous system embodies—and any body part, for that matter—the situation of the individual. For instance, she claims that "the nervous system is an integrator of and integrated with the entire body and the world (…)." And adds a few lines after that "a practice that affects one part of the body will be owned by the entire body or, embodied through the interconnections of all body systems and the environment" (Einstein, 2012, p. 158).[109]

In the study, Einstein was interested in enquiring whether FGC had resulted in something akin to phantom pain. The subjects were a group of women of either Somali birth or Somali descent in Canada. Taking into account the specific circumstances, and the cultural and social context of the women who took part was crucial. For instance, Einstein avoided working with a pre-conceived notion of the way pain would appear in the experiences of the women involved in the study. In that way, Einstein rejects the universality of embodied experiences. For her, the body is marked by the context, history of interactions and experiences, of an individual. She claims that: "[a]ctivity in the world affects the body and these effects may or may not make their way to the brain but they still instantiate themselves in response, memory, language, and thought" (Einstein, 2012, p. 160).

Einstein found that FGC has an effect on women's central nervous system and body. Among the physiological findings, she enlists the following effects: (1) changes in the somatosensory cortex that lead to referred sites of pleasure; (2) regular experiences of exhaustion around women's period; and (3) various sites of pain in their bodies with a medium to high rate of pain (Einstein, 2012, p. 155). She also reports that, in quantitative sensory tests carried out with only a subset of participants, they found evidence of vulvar regions with low pain thresholds, although participants

[109] Einstein develops this view in opposition to a mechanistic view of the body.

did not report having vulvar pain (Einstein, 2012, p. 156). Furthermore, they did not show signs of post-traumatic stress or other similar distress. In addition to these findings, Einstein notes that they described, when pressed, having experienced moments of significant pain or discomfort. However, their perspective of this pain is of a typical ache. The participants considered themselves healthy, continued to lead active lives, and expressed pride about their experience and culture. The women who participated in this study did not have, in Einstein's words, "an illness identity" (Einstein, 2012, p. 156). Moreover, she claims that her study shows that the participants "embody their culture" (Einstein, 2012, p. 156). Note here that Einstein is very careful in pointing out that this is not a generalized nor generalizable experience. These results reflect the experiences of women that belong to a very specific set of circumstances and might not be reflected in the experiences of women whose circumstances are different. She claims that:

> [M]ost of the participants in our study were from wealthy, educated families in Somalia. They came to Canada as part of the first immigration in the late 1980s/early 1990s (…) This is a very different body than those of lesser socioeconomic status, less schooling, and less privilege who had to continue to live in war-torn Somalia and who then left only to end up in refugee camps in Tanzania with eventual immigration to North America. And the meanings of their circumcision and the body on which it is played out are likely very different (Einstein, 2012, p. 163).

Einstein's analysis shows the complex interplay between different aspects of someone's circumstances and the way these are embodied. Moreover, her analysis lends support to the idea that perceptual experience manifests the way individuals embody their situation.

8.2 The question about the social character of perception

Both Young and Einstein hold the view that someone's situation shapes her mental life and interactions with the world. Moreover, they both seem to adhere to something akin to the biosocial conception of embodiment described in the previous chapter. Although a view about the situatedness of perceptual experience needs to be articulated, I dare to say that they would both accept the plausibility of such thesis. Of course, these considerations only take us so far in what concerns this claim. After all, from the idea that pain and comportment are situated, it does not follow that perceptual experience is situated too. From the fact that our situation has an impact on how we move, it does not follow immediately that it has an impact on how we move to investigate the world sensorily. Similarly, from the impact it has on the experience of pain, it does not follow that it has a comparable impact on perceptual experience. We can articulate this concern in the following way: from the fact that the property of being situated is attributed to the whole person, it does not follow that it is a property that can be attributed to each of the capacities of the agent.

8.2.1 The puzzle of perceptual experience

From the perspective of SMEn, we might not feel too compelled to raise this worry. After all, SMEn defines perception as something the agent does. So, if someone's situation has an impact on or shapes her actions, it becomes plausible to think that the situated aspect of embodiment will be reflected in perception. So, the concern arises in light of the way we conceive of perception. When we think, for instance, of the position enjoyed by perception as a source of knowledge, we might worry that the social influence on perception might hinder this role. This has been addressed by Anne Jaap Jacobson (2012), who notes that, while perceptual experience is typically taken to be a reliable guide for the "discovery of truths", it is also the case that knowledge is affected in many ways by our socio-historical circumstances.

This topic has been greatly explored by feminist philosophy, where some have argued that society influences our knowledge. For instance, the structure and resources of an institution can influence the way certain research is pursued, as well as the way it is publicized. Society also influences knowledge production by excluding individuals from certain backgrounds from the institutions of knowledge production, as well as by discounting them as epistemic agents. In addition, some individuals might have better epistemic access to understanding certain circumstances. For instance, given their particular background, someone might be in a better epistemic position to identify whether a case would count as an instance of micro-aggression.

In contrast with this, Jacobson asks whether society influences knowledge "all the way down" (Jacobson, 2012). In a way, she distinguishes between two different questions, whether society influences who we are as epistemic agents, and whether it shapes who we are as cognitive agents. According Jacobson, perceptual experience is, in fact, influenced by our social circumstances, but only to a degree. Let me briefly reconstruct her position, which focuses exclusively on visual experience.

Jacobson starts off from a similar place to our discussion of SMEn by noting that there are two aspects of perceptual experience that need to be reconciled. As I discussed in section 1.1, while our visual system is quite limited, our visual experiences are rich, and do not reflect these limitations as such. Jacobson is interested in understanding the impact of society on perceptual reports, as well as on the way we fill the gaps of sensory information that are left given the limitations of our visual system. But for Jacobson this is not the only puzzling feature of perceptual experience. It is also impressive that, obtaining such limited information, vision achieves impressive things. Vision, Jacobson argues, is a highly selective system. Although it is gappy, it allows us to achieve impressive things. Consider, for instance, the contribution of the visual system in mirroring social behavior, as well as its contribution in guiding action, particularly in cases in which certain information is required for fast reactions. So, how is this achieved?

8.2.2 Perception as incidentally social

Jacobson distinguishes two stages of consolidation in the process that goes from saccades to objects. The first consolidation stage refers to the binding of qualities. Qualities that are processed separately are bound together by the workings of attention. At this stage, we go from processing colour, shape, and size separately, to processing these properties as belonging to the same object. The second stage of consolidation refers, instead, to the moment when the "jerky input" we get from saccades is transformed into objects (Jacobson, 2012, p. 222). Saccades give us very short takes of the scene in comparison to the way it appears to us. Attention plays an important role at this point, too, since it is involved in the coordination of saccadic movements. Furthermore, it is this latter stage that is said to "involve amodal completion" (Jacobson, 2012, p. 222), the phenomenon in which elements of a scene from which we do not get sensory stimuli appear, nevertheless, in our experience.

At this point in her argument, Jacobson takes issue with SMEn. Recall that, according to SMEn, amodal completion is explained by perceivers' possession and mastery of sensorimotor knowledge, knowledge of the systematic co-variations between sensory information and interactions (i.e. sensorimotor contingencies (SMCs)). In opposition to SMEn's view, Jacobson argues that more recent research supersedes this proposal.[110] In conflict with one of the core claims

[110] Here Jacobson refers to (Martinez-Conde et al., 2008).

of SMEn, research has shown that these first stages of consolidation do not involve nor rely on knowledge of any kind.[111]

According to Jacobson, as a result of these two stages of consolidation we get a scheme of three-dimensional, coloured objects. For Jacobson, this seems to be the first scheme of the scene perceived. Nonetheless, it still stands a long way from our rich perceptual experience. Furthermore, at this point, she claims, there is no clear evidence of the relevance of social influence. She characterizes these two stages of early visual processing as the "realm of the pure sensory bottom-up processing" (Jacobson, 2012, p. 223). After these stages, we finally come to the classification of objects. Jacobson looks at evidence from developmental psychology. She notes that, in what concerns what is sufficient for children's perception of the persistence of objects, 12-month old children go from "spatio-temporal continuity" to continuity of kind. To illustrate this she says that, before 12 months of age, "an object that is occluded may change into any other sort of object, at least as far as the [child] is concerned" (Jacobson, 2012, p. 223). From this point in their development, children seem to have an experience of a world populated by "stable, lasting objects" (Jacobson, 2012, p. 223). At this point, according to Jacobson, perceptual processing stops being solely bottom-up processing and involves top-down processing, since it relies on conceptual learning.

It is at this point that perceptual experience can be said to receive social influence. This is so because, in order to fill the gaps left by our sensory input, we draw on our socially learned concepts:

> The adding-in that takes us from the schema to the completed scene in fact draws on a host of factors, including surely our grasp of the relation between our environment and our movement, results from further saccading and conceptualization of its products, past experience with the location and kinds of things to be seen, cultural conventions about artifactual objects, the input of other senses, and so on and so forth. The result is that we no longer have a somewhat schematic and partial scene; rather, as our sensory reports reveal, we have a much fuller picture of the environment (Jacobson, 2012, p. 224).

[111] Here, I do not respond to her objection. Mainly, because she does not say much more about the way she understands sensorimotor knowledge. In the next section, though, I briefly discuss some of the literature on which she relies, since some of the papers she discusses already point towards the anticipatory character of perceptual processing.

Jacobson's thesis is that perception is not social in itself, but in virtue of the interactions between perception and higher-order cognition. For Jacobson, societal influence does not enter into the picture at an early stage, at the level of early visual processing, but at a later stage when top-down processing becomes relevant for filling-in details that are left under-determined in the first two stages of consolidation. If perceptual experience is social, according to Jacobson, this is incidental on visual processing relying on higher-order cognitive processes. And, importantly, perception is not social all the way down.

8.2.3 Some remarks about Jacobson's description of early visual processing

Perhaps one of Jacobson's more problematic claims is that early visual processing is the domain of bottom-up processing. For instance, Jacobson refers here to a paper by Shuler & Bear (2006) where they show that the brain of adult rats is more plastic than previously argued. For their study, rats were equipped with goggles that illuminate either one eye or the other. When the rat approached a water source, one of its eyes was illuminated. The neural response to these stimuli was monitored. After a few trials, the authors noted that neural activity seems to anticipate ("portend") the reward (Shuler & Bear, 2006, p. 160). So, what is interesting here is that this paper shows that the activity of some neurons in cortical area V1 is anticipatory.

There is now vast evidence against the view of early visual processing as an exclusively bottom-up domain (Huang & Rao, 2011, p. 12). Let me just briefly mention a couple of examples. Consider, for instance, Srinivasan et al.'s (1982) proposal to interpret the activity of interneurons in the retina as predicting the signal of neighbouring cells. Or consider Gilbert & Li's (2013) paper, where they emphasize the influence of top-down processing in early visual processing. The authors argue that influence from higher-order cognition interacts with retinal information. Furthermore, they add that this influence is pervasive in the visual pathway.

Jacobson is likely not unaware of at least some of this evidence (I mean, of course, proposals available at the time of publication of her paper). Moreover, the weight received by these proposals has shifted incredibly in the past years. Gilbert & Li (2013), for instance, note that top-down influences in early visual processing was deemed negligible. Finally, evidence of the anticipatory character of neural dynamics in the shape of, e.g., predictive coding, does not amount just yet to what Jacobson is looking for to support the claim that higher-order processing cascades all the way to these areas.

Let me try to make sense of Jacobson's position. For Jacobson, social influence enters the picture at a conceptual level. Society fixes our concepts, and it is in virtue of this that it shapes perceptual experience. We fill in the details of our gappy sensory stimuli based on concepts we have learned. Importantly, in her model, we encounter an already structured social world that shapes our experience. Unless processing associated with conceptual learning influences perception, one cannot say—according to Jacobson—that perception is influenced by society. Now, if this is the right way to capture her position, it seems that Jacobson's view would be compatible with a version of hierarchical predictive processing, since the proponent of hierarchical predictive processing could say that conceptual influence goes all the way down.[112] However, that is not the story I want to tell. Because by restricting the influence of society to the conceptual, Jacobson is missing part of the story. This is the story told by Olkowski, Higgins, Young, De Beauvoir, and Einstein, according to whom we need a richer conception of embodiment. This story makes the claim that perceptual experience is social all the way down viable. Moreover, it also indicates that the influence of society is exhibited not only through our concepts.

8.3 Situating perceptual experience

8.3.1 More than conceptual influence

To lend plausibility to the thought that social influence is not reduced to conceptual influence, let me turn briefly to some examples. I focus, first, on examples that show the relevance of social interaction to the development of perceptual capacities. And then, to differences in perceptual experience across cultural groups.

[112] Jacobson's position can also be interpreted as committing to a version of the cognitive penetration thesis, according to which, if it is possible for two individuals who are in identical relevant conditions (e.g. perceiving conditions) to have different perceptual experiences due to a difference in the content of higher-order mental states (i.e. beliefs, judgements and desires), then perceptual experience can be cognitively penetrated (Macpherson, 2017, p. 2).

Social interactions and the development of perceptual capacities

I draw here on Bermúdez's analysis of joint visual attention and social referencing. Joint visual attention is the phenomenon in which two individuals attend to the same object in virtue of where they perceive the other's gaze to be directed (Bermúdez, 1998). Scaife & Bruner (1975) and Murphy & Messner (1977) show that at around 8 to 9 months old, children are able to engage in joint visual attention. For instance, Scaife & Bruner show that at 8 months old, children can follow the gaze of another individual who has established visual contact. Murphy & Messner, on the other hand, show that 9-month-old children are able to follow the direction to which their carer points. This shows that children of this age understand that pointing is a cue for joint visual attention. More importantly for our purposes, Bermúdez notes that there is also evidence that children look back at the person who gave them the cue, looking for feedback. Finally, 6-month-old children exhibit the ability to look back and forward, alternating between someone's gaze and an object (1975).

These studies on joint visual attention show the relevance of social interactions in the development of perceptual capacities. At an early age, children learn that perceptual experience is something that can be shared with others. Furthermore, they learn that someone else's gaze and gestures are indicative of something that is worth attending to and investigating with their senses. In line with this, Bermúdez cites an interesting passage from Daniel Stern (1985) connected to children's movement as an exploration of different perspectives. He claims that "[i]n moving about, the infant continually alters the perspective held on some unknown stationary sight. Perhaps this initial acceptance of serially different perspectives is a necessary precursor to the more generic "realization" that others can be using a different coordinate system from the infant's own" (Stern, 1985, p. 129–130 as quoted by Bermúdez (1998, p. 257)).

Social referencing, on the other hand, is the regulation of one's behaviour on the basis of others' "emotional reactions to a particular situation" (Bermúdez, 1998, p. 262). In children's development, this is manifested when children face a puzzling, unfamiliar, or intimidating situation. Children look back at their carer, looking for guidance. According to Bermúdez, children's behaviour will be influenced by the emotional response of the carer to the situation. Bermúdez refers to an experiment by Klinnert et al. (1983). Just like the kittens in Held & Hein's (1963) experiment, children were tested on the avoidance of a modified visual cliff. Recall here that the test uses a modified platform with a clear glass plate with two sides, a shallow and a deep side. Below the shallow side, a chequered pattern is placed immediately below the surface. On the deep side, the same pattern is placed below, but at a certain distance below the surface to make it appear as a cliff. Klinnert et al. observed that 12-month-old children look for feedback from their carer after looking at the deep side of the platform. Bermúdez reports the following results: "Of the 19 infants whose mothers smiled, 14 crossed the deep side, while of the 17 infants whose mothers showed fear, none crossed the deep side to the mother" (Bermúdez, 1998, p. 263). What this shows is that we learn how to interpret sensory information based on feedback from others: the cliff becomes something children should avoid in light of their carers' reaction. Depth perception is developed, in part, socially. Moreover, the fact that this evidence is related to early cognizers puts pressure on Jacobson's idea that social influence can only by conceptual.

Differences in perception across cultural groups

Non-conceptual social influence is noticeable not only in relation to perceptual development. There is also evidence of variations in perception across individuals from different cultural backgrounds. Consider, for instance, differences across individuals from different cultural groups in their susceptibility to the Müller-Lyer illusion (see e.g. McCauley & Henrich, 2006). This optical illusion involves two or three double-headed arrows of the same length, whose ends differ in the direction to which they point, inwards towards the centre of the arrow or outwards. It was previously thought that everyone was equally susceptible to perceiving one of the arrows as longer, depending on the direction in which the ends point.

However, in an extensive study by Segall et al. (1966), the authors showed that there are substantial differences in the susceptibility to the illusion depending on the cultural environment in which an individual spent her first twenty years of life. The authors analysed 17 small scale societies, among which were groups of agriculturalists societies, foragers, horticulturalists, and goldmine-laborers, as well as a couple of groups of "Westerners" (McCauley & Henrich, 2006, p. 92). According to McCauley & Henrich, the difference has been attributed to the adaptation of the visual system to the local environment. Depending on the environment, the visual system "[builds] up biases that tend to produce useful inferences in that environment" (McCauley & Henrich, 2006, p. 95). Segall et al. examined on the hypothesis that the cause of these variances is individuals' exposure to carpentered environments, namely environments that are populated by buildings, rooms and furniture with sharp right angles. These angles are interpreted by the visual system, in consequence, as an indication of depth, leading, thus, to the illusion (McCauley & Henrich, 2006, p. 95).

Other studies of cross-cultural variation have shown differences between individuals from different cultural backgrounds. Chua et al. (2005), for instance, focused on the difference between American and Chinese individuals when processing a scene. It has been shown extensively that American individuals attend to focal objects, Chinese individuals focus on contextual information. Chua et al.'s research found, in addition to this, that the difference is due to differences in saccadic eye movement and fixation region. For instance, American individuals tended to focus for longer period of time in certain regions, while Chinese individuals exhibited a more balanced pattern of fixations that went from foreground object to background. In their discussion, the authors attributed these differences to differences in "experience, expertise, or socialization" (Chua et al., 2005, p. 12663). More specifically, for the authors, this might be due to childrearing practices and to the relevance of social roles, since a society with complex social networks and roles might require further attention to context. What I want to emphasize here is that these social differences lead to a difference in the way we exercise our basic sensorimotor skills.

8.3.2 The predictive approach to Sensorimotor Enactivism at work

According to the examples just reviewed, then, perception is influenced by society at least in three other ways, in addition to the conceptual influence identified by Jacobson. Firstly, social interactions are relevant to the development of perceptual capacities. Secondly, social and cultural circumstances have an effect in our susceptibility to perceptual illusions. Depending on our social context, we might be better adapted to certain visual information. And, finally, there are relevant cultural variations in the exercise of our more basic sensorimotor skills. What these examples suggest is that perceptual experience is social all the way down. Further evidence would be needed to show that is it situated in the full-blown sense discussed previously. For instance, we might wonder whether gender identity makes a difference and how; or how someone's profession and level of skills might be reflected in perceptual experience. Still, the evidence we have already takes us one step closer. However, we also need a framework that helps us make sense of these differences. Jacobson's proposal, for instance, fails to predict these results. In the next section, I show that the predictive approach to SMEn can provide this framework.[113]

Situated prediction machines

Recall that, in Part III, I defended a predictive approach to SMEn. I argued that the predictive framework can offer an account of SMCs, the systematic co-variations between sensory information and interactions. According to this view, the mastery that is constitutive of perceptual experience relates to perceivers' sensitivity to SMCs. This mastery is manifested in the way perceivers manage their interactions on the basis of the SMCs as they are encoded in or recapitulated by the recognition density, which is implemented both neurally and bodily. I think that, with this proposal in place, we can argue that perceptual experience is social all the way down.

Kirchhoff & Kiverstein (2019a) already make a case for the cultural aspect of the FEA. For them, the generative model and the recognition density track our engagement in social and cultural practices. This has an important consequence for our conception of the cycle of surprise minimization. For Kirchhoff & Kiverstein, what exerts an influence on perceptual processing are not just the priors that cascade down from higher levels of a neurally implemented generative model. Instead, they claim, "it is our ability to maintain attunement with the regularities in a cultural practice that can (depending on the context) exert a powerful influence on the perceptual processing of a sensory signal" (Kirchhoff & Kiverstein, 2019a, p. 95).

This is consistent with the view espoused here, according to which the generative model is enacted in the interactions of the agent. Perceptual processing is anticipatory, and it is constrained by priors. However, these priors constrain processing not only from within the system, but from the context of the agent. In line with this, Kirchhoff & Kiverstein claim that "[c]ultural practice is one of the driving forces behind uncertainty reduction" (Kirchhoff & Kiverstein, 2019a, p. 97). Furthermore, the authors identify that these constraints are reflected in precision weighting. Precision weighting

[113] In a recent paper, Arango (2019b) already notes that SMEn's framework might be able to accommodate some of the social aspects of perception. However, he takes issue with both the notion of sensorimotor knowledge and the problems we have discussed in connection to it, as well as with the relevance of morphological embodiment. Arango turns to a Wittgensteinean notion of perceptual practices, which he defines as that which people do, i.e. the patterns of our practices and actions. I think this is already present in SMEn as it is enriched by Buhrmann's et al. (2013) model. However, I agree with Arango in that the commitment to morphological embodiment (what I called implementational embodiment) is limiting. See Arango (2019a) and McGann (2019) for discussion.

reflects “how uncertainty is kept to a minimum” in a given context (Kirchhoff & Kiverstein, 2019a, p. 97). Put differently, precision weighting must align with specific cultural practices. These include our sensory explorations of the world and the sensorimotor skills that enable them. For instance, think back to the case of social referencing. This view would predict that sensory information consistent with the presence of a cliff is assigned a high level of precision. Moreover, this is so due to the interaction of the agent with others.

Kirchhoff & Kiverstein add that our attunement is not only to a context, but to “expectations between people” (Kirchhoff & Kiverstein, 2019a, p. 99). Drawing on Roepstorff et al. (2010), they argue that this might explain how prediction errors are propagated. We share a common world with those with whom we engage in cultural practices. Consequently, we likely share a model of the world. According to Kirchhoff & Kiverstein, isolated predictions and precision weighting might explain, for instance, some features of autism, everyday anxiety, and schizophrenia.

One aspect that is crucial and that is left out of Kirchhoff & Kiverstein’s analysis is that precision weighting will depend not only on the context of the individual and the task or practice with which they are engaged. It will also depend, in an important way, on who that individual is and on certain features of her situatedness. This is because *the context to which we have to adapt is not the same for everybody, nor are the expectations from others*. Our social identity makes a difference in the way we participate in these social practices.

That is, precisely, the point of Young’s analysis. Even when engaged in the same practice, within the same context, the gender identity of an individual plays a role in determining the space of possibilities and expectations. For instance, when throwing a ball, the proprioceptive prediction of a woman in a given context will be different from that of a man. The routines of surprise minimization exhibited by individuals will differ, in some way, depending on their social identities. Put differently, our patterns of sensorimotor interactions will vary due to features of our situation. Young, for instance, indicates that this might explain the difference in performance in tasks related to spatial perception. Accordingly, given our gender identity and other aspects of our socializations, we might sample sensory information differently to minimize surprise.

This is, of course, speculative. Much more is needed to argue that differences in social identities make a difference to perceptual processing and in the way we engage in what Gallagher & Allen

(2018) call embodied inferences. However, I have shown that it is plausible, within this framework, to think that they do.

Conclusion

My main goal, in this thesis, has been to tackle certain issues that arise around Sensorimotor Enactivism's (SMEn) conception of embodiment. SMEn is a theory that claims that perception is something perceivers do, that requires the possession and execution of practical knowledge about the way sensory information changes after an interaction. In short, SMEn claims that perception is active and knowledgeable. One of the key commitments of this view is to the idea that this knowledge—sensorimotor knowledge—depends on details of the body of the perceiver. The role played by the body, and the challenges that arise from the conception of embodiment in SMEn have served as the guiding thread of this thesis.

While Part I mainly focused on setting the stage for the different topics covered here, Parts II, III, and IV examined three main problems that relate to different aspects of embodiment. Part II discussed the dependence of sensorimotor knowledge on the body, as well as the concern that this dependence leads to a view that is chauvinistic in what concerns the content of perception and the individuation of sense modalities. In Part III, I focused on the concern that, while SMEn emphasizes the relevance of the body, it leaves out one crucial element of the material basis of perceptual experience, namely, the brain. To address this, I argued that joining forces with the predictive framework—specifically, with the Free-Energy approach—allows SMEn to provide a complete causal story about the mechanisms that underlie perceptual experience. Once SMEn is equipped with this causal story, in Part IV, I tackled one last concern connected to the conception of the body that is at play in SMEn. This worry is related to the phenomenologically inspired idea that the body is marked by the agent's social and historical circumstances. This is the idea that perception is situated. In what follows I summarize the main ideas of each chapter before offering final remarks in relation to a problem that I raised in the introduction in connection with the place of SMEn's in the wider 4EA context.

Part I set out to examine SMEn's main features. In Chapter 1, I began by discussing two assumptions abandoned by the framework, namely, the sandwich view of the mind and the snapshot conception of perception. The abandonment of these views highlights SMEn's main claims: its defence of the close relation between action, perception and cognition—in contrast to the sandwich view of the mind, and the claim that representations are not the end-result of perception—in contrast to the snapshot conception of perception. As an alternative to representational conceptions of perception, SMEn aims to account for the problem of perceptual presence, which relates to the fact that perceptual experience is rich and detailed, despite limits of our perceptual systems. Two key claims of SMEn account for this puzzling feature: that perception is active, and that it is knowledgeable. Although these claims have been greatly discussed, some aspects of these claims remained contested.

In relation to the first claim, I defended the view that perception is constitutively active in that it is something *done* by the whole organism or person. When this claim is considered in relation to movement (i.e., as perception as being kinetic), I argued that SMEn is committed to the view that movement is an *enabling* condition of perception. By this, I meant that SMEn would predict that a lack of movement hinders the development of someone's perceptual capacity or performance in a perceptual task.

In relation to the second claim, I drew on Silverman's (2018) proposal that sensorimotor knowledge is practical knowledge that grounds the execution of the embodied skills that constitute perception. This position satisfactorily addresses the concern that sensorimotor knowledge must fulfil two roles within the theory, namely, explaining that perception consists in the execution of skills and explaining the sense of presence characteristic of perceptual experience.

Chapter 2 moved on to examine SMEn's position regarding representations and explanatory externalism. While SMEn denies that referring to the brain is sufficient to account for perceptual experience, it does not deny that representations might be involved in perceptual experience. I defended, firstly, the claim that SMEn constrains the *role* played by representations, since the view rejects the idea that these are the end-result of perception. Secondly, I defended the claim that SMEn constrains the *content* of the representations that might be involved in perception, since the view accepts representations that are agent-specific and indicative of further possible interactions. In addition to this, drawing on Silverman (2018), I defended the view that SMEn's description of sensorimotor knowledge is of a domain that is *prima facie* representationally hungry. Accordingly, sensorimotor knowledge can be causally enabled in different ways, including by representational states.

With regards to the specific brand of externalism defended by SMEn, drawing on arguments by Noë (2007) and Hurley (2010), I argued in favour of the thesis of explanatory externalism. This thesis involves two steps. Firstly, given the close feedback loops between brain, body and world that characterize perceptual processing, explaining perceptual experience by referring solely to neural activations turns out to be insufficient. Secondly, the vehicles of perceptual experience must be identified on the basis of the elements that are necessary for our explanations. Against the claim that body and world are only causally relevant for perceptual experience, I argued that the causal-constitutive distinction becomes otiose.

Given my defence of explanatory externalism, Part II focused on examining the dependence of sensorimotor knowledge on the body. Chapter 3 examined the theory of Sensorimotor Contingencies (SMCs) as a bridge between a theory of perception and a theory of perceptual processing. This theory fulfils a crucial role when we consider that SMEn is a naturalistic theory of perception that is meant to be continuous with a scientific account of the mechanisms of perception. I began by examining SMEn's view of the content of perception. Drawing on Kiverstein (2010), I defended a weak version of SMEn and argued that SMCs contribute to fixing the character of perceptual experience. What appears in perceptual experience as pertaining to both the background and the foreground of experience is fixed in part by perceivers' sensorimotor expectations in virtue of their possession and deployment of sensorimotor knowledge.

Based on this discussion, I defined SMCs as the systematic patterns of co-variations between sensory information and possible interactions between the perceiver and the object(s) of perception. In discussing and examining the theory of SMCs as it was first introduced by O'Regan & Noë (2001), I set the stage for the claim that SMCs already involve a certain degree of sensorimotor summarizing, i.e. they already involve information about objects' invariant features deployed in planning, action, and reasoning. Besides accounting for the character of perceptual experience, SMCs also play a role in the individuation of sense modalities. Chapter 3 discussed two elements of the theory of SMCs that become crucial to the arguments advanced in subsequent sections. Firstly, I discussed Buhrmann et al.'s (2013) operationalization of SMCs. And, secondly, I discussed the embodiment claim, i.e., the claim that the pattern of SMCs is shaped the body.

The embodiment claim gives rise to a concern that, given the dependence of SMCs on the body of the perceiver, SMEn cannot extend certain attributions to all the cases of perceptual experience it should. Consequently, the theory of SMCs cannot fulfil its explanatory purposes of accounting for the character of perceptual experience and of providing parameters for the individuation of sense modalities. In this regard, Chapter 4 distinguished between two interrelated worries: the content chauvinism concern and the sense modality chauvinism concern. These two kinds of chauvinism have not been distinguished in the literature. Instead, chauvinism has been tackled by claiming that SMEn delivers something akin to multiple realizability. Although this is not wrong, more needs to be said, since there are proposals that deliver multiple realizability without avoiding chauvinism.

The content chauvinism concern arises from an interpretation of the embodiment claim according to which every low-level detail is relevant for fixing the content of perceptual experience. Against this interpretation, I argued that the embodiment claim involves instead a commitment to an irreducibly perspectival aspect of perceptual experience. Nevertheless, even if SMEn were to commit to this weaker claim, the agent-centred aspect of SMCs stands in opposition to the two visual system theory advanced by Milner & Goodale (2006). To tackle these concerns, I argued that the weak version of SMEn defended in Chapter 3 already involves sensorimotor summarizing. For this reason, SMEn cannot be attributed the chauvinistic claim. On the other hand, based on recent evidence, I argued—against Milner & Goodale's model—that the content of perceptual experience has a perspectival aspect that is consistent with SMEn's dual content theory and the theory of SMCs.

To tackle the sense modality chauvinism concern, I drew on Buhrmann et al.'s (2013) dynamic models of SMCs introduced in Chapter 3. I argued that the notion of sensorimotor coordination can provide a systematic way to distinguish between ambiguous cases of occurrences of sense modalities such as the case of the TVSS device and the organic visual system. On the basis of this model, we can establish a structural profile that allows us to determine if the profile of two systems is comparable. Moreover, by comparing the description of two systems' sensorimotor coordination—the model that considers the system's performance in a task—we can provide a good reason to classify two systems as supporting perceptual experiences of the same kind.

As mentioned, the theory of SMCs is meant to serve as a bridge between SMEn as a theory of perception and a causal story about the mechanisms that underlie perceptual experience. In that line of argument, the four kinds of SMCs developed by Buhrmann et al. are meant to provide an operationalization of the concept of SMCs. However, one aspect that is left unattended concerns the role of the brain in that causal story. Part III addressed this concern by arguing in favour of a predictive approach to SMEn.

In Chapter 5, I examined the predictive framework, which I used as an umbrella term to encompass predictive coding, predictive processing, and the Free-Energy approach (FEA). These theses emphasize the anticipatory nature of neural dynamics and claim that the brain is driven by top-down predictions rather than by bottom-up processing. I mostly focused on the FEA as the view that can be made compatible with SMEn. I explained that the FEA is a theory about the adaptive behaviour of biological systems according to which they aim at surprise minimization, i.e., at minimizing engagements with the environment that would jeopardize systems' viability. I defended an externalist version of the FEA, according to which the process of surprise minimization spans over brain, body, and world. In line with this, I drew on Ramstead's et al.'s (2019) distinction between the generative process, the generative model and the recognition density. They argue that the process of surprise minimization is both embodied and enacted. It is embodied in that the recognition density—the system's model—is encoded in the brain and body of the system; and it is enacted in that the generative model is entailed by the recognition density, rather than being physically instantiated.

Chapter 5 also addressed some concerns that surround the predictive framework. I particularly focused on its representational profile, since this presents the greater challenge in what concerns its compatibility with SMEn. For that reason, building on Gładziejewski (2016) and Williams (2018), I built the strongest possible case in favour of the representational reading of the FEA. I argued that the recognition density is characterised as a recapitulation of the environment in agent-dependent terms. With this in place, Chapter 6 argued in favour of the compatibility between SMEn and the FEA. I reviewed two proposals for a predictive approach to SMEn, Downey's non-representational alternative and Seth's representational alternative. I concentrated particularly on Seth's proposal, according to which the generative model encodes SMCs. I argued that SMEn can be made compatible with a representational version of the FEA, since the representations advanced by the FEA comply with the constraints discussed in Chapter 2. In what concerns the internalist profile of some views within the predictive framework—including Seth's own proposal—I argued that given the distinctions introduced in the previous chapter between the recognition density, the generative model and the generative process, SMCs are not instantiated only by the brain, but by a whole brain-body-world system.

Part IV aimed at articulating a conception of situated embodiment that could be incorporated into the predictive approach to SMEn defended in Part III. Chapter 7 began by examining the conception of embodiment that is at play in the FEA. Drawing on Wheeler (2010b), I distinguished between vital and implementational embodiment. I suggested that the latter risks not doing justice to the materiality of the body *as* it contributes to perceptual experience, and that, on the contrary, it is vital embodiment that is at play in the FEA. The FEA is committed to the strong continuity thesis between mind and life. On the one hand, the FEA is committed to the idea that life and mind are governed by the same principle, the free-energy principle. On the other hand, within this framework, the materiality of the body takes on a new relevance given the role it fulfils in the process of surprise minimization, namely *embodying* the model from which the system advances predictions about its sensory and active states and *enacting* its norms of constitution in its interactions with the environment (Ramstead et al., 2019).

In the second section of this chapter, I drew on the phenomenological conception of situated embodiment that is advanced by feminist phenomenologists. I argued that, according to the conception of situated embodiment articulated on the basis of the phenomenological analysis, the body stands at the intersection of the vital and the symbolic, that it constitutes the situation of the person and that, given that it constitutes the locus of experience, experience is irreducibly situated. Drawing on Higgins' (2018) conception of the biosocial body, I argued that the conception of embodiment that is at play in the FEA can be interpreted in these terms, i.e., as a body that is saturated by the social. For this reason, the FEA already reflects the situated character of embodiment.

Finally, in Chapter 8, I advanced a proposal to accommodate the situated character of perceptual experience into the predictive approach of SMEn. I focused exclusively on one aspect of someone's situation, namely gender identity. I argued that Young's (1980) and Einstein's (2012) analyses of feminine comportment and pain, respectively, lend support to the idea that someone's social circumstances shape her mental life. Moreover, I argued that these analyses suggest that perceptual experience is situated all the way down. Against Jacobson's view that perception is only occasionally situated due to the interaction between perception and higher-order cognitive processes, I argued that we have reasons to think that the influence of society is not limited to a kind of conceptual influence. I reviewed the relevance of social interactions in the development of perceptual capacities, which I illustrated with the case of joint visual attention and social referencing. I also reviewed the relevance of one's belonging to certain cultural groups relation to the content of perception, which I illustrated with cases of cross-cultural variation in susceptibility to the Müller-Lyer illusion, and with evidence of differences across individuals from different countries in the way they process a visual scene. In light of these cases, I argued that perception is influenced by society in other ways.

Finally, I argued that the predictive version of SMEn provides a framework to make sense of this. I drew on Kirchhoff & Kiverstein's (2019a) position, according to which our cultural practices and the expectations of others influence the predictions that are advanced as part of perceptual processing. To this I added that certain aspects of our situation contribute to the way we participate in these social practices and the expectations others have from us. Thus, our routines of surprise minimization depend in an important way on aspects of our situation.

Pairing SMEn with the FEA provides a story that can fill in the gaps left by the original conception of SMEn. It opens the door to the possibility of providing a more detailed account of the workings of the perceptual system, as well as to its relation to an account of perceptual experience at the agential level that is phenomenally apt. Moreover, it provides an account of a cognitive system's embodiment that does justice to the biological and social aspects that characterize at least some of these systems. In consequence, this account can also do justice to the thought that we are concrete cognitive agents. That is, it can start to explain how differences in embodiment result in differences to our mental lives.

Although, I have focused exclusively on perceptual experience and perceptual processing, this joint approach is also promising in relation to other mental states and processes. More specifically, in what concerns imagination. For instance, some have concluded that, if perception is enabled by processing structured in the way described by the predictive framework, the line between perception and imagination becomes blurry. Accordingly, what is perceived is enabled by imagining-like states (Kirchhoff, 2018b). This brings back in the worry that perception is at one-remove from the world, since the content of perceptual experience is not the world, but an internal construct, an imagining. However, this consequence only follows from an internalist interpretation of the predictive framework.

When we take on a different perspective, a perspective according to which the process of surprise minimization is a process of attunement with the world, the relation between imagination and perception can be read in a different way. Rather than proposing that perception is at one remove from the world, we can take a different perspective on imagination and explore the ways in which this is an embodied process that piggybacks on our sensorimotor interactions. Moreover, we can think of the way imagination contributes to the process of surprise minimization. Importantly, the close relation between perception and imagination is, in a sense, already present in SMEn. If we take imagination to consist in the representation of non-actual states of affairs, by claiming that perception is supported by the coding of counterfactuals, the sensorimotor theory might have to accept that there is a way in which perception is mediated by imagination. Of course, more needs to be said about this.

In addition to this, having equipped SMEn with a causal story, we have now a fully developed theoretical alternative to autopoietic enactivism. The FEA provides the story that is missing in SMEn about the mechanisms that support perceptual experience. In the introduction, I mentioned that there is an open question looming over SMEn. This question concerns its place in the wider context of the 4EA approaches to cognition. In connection to this, it has been suggested that SMEn is an enrichment of the orthodoxy rather than a departure from it (Vernazzani, 2019). So, where does the predictive approach to SMEn leave us in what concerns the place of SMEn in this wider context?

Perhaps a good place to start thinking about this is by reframing this question. Wondering whether SMEn is enriching the orthodoxy or not stops making sense when we consider, for instance, the position of the predictive framework in the debate. The interpretation of the FEA and of the predictive framework as in close proximity to the 4EA approaches of cognition is not the prevalent treatment they receive in the literature. More often than not, they are read in a representational and internalist key. In a similar way to SMEn, the predictive framework seems to be a contested territory. Moreover, as Kirchhoff (2018a) has noted, when the FEA is read as a theory of the adaptive behaviour of biological systems, it does not fall far from autopoietic enactivism. He notes that the FEA and autopoietic enactivism seem to provide complementary stories about biological systems. While autopoietic enactivism sees the system from within—from the perspective of the system’s processes of self-production—the FEA sees it from the outside—from the perspective of its processes of self-preservation (Kirchhoff, 2018a, p. 2521). The marriage between SMEn and the FEA can be considered, then, a promising step in that direction. Rather than enriching the orthodoxy, when paired with the FEA, SMEn serves as a bridge between non-representational and more orthodox approaches within cognitive science.

Bibliography

Adams, F., & Aizawa, K. (2010). *The Bounds of Cognition*. Blackwell Publiching Ltd.

Adams, R. A., Shipp, S., & Friston, K. (2012). Predictions not commands: active inference in the motor system. *Brain Structure and Function*.

Agre, P. E., & Chapman, D. (1987). Pengi: An Implementation of a Theory of Activity. *Proceedings of the Sixth National Conference on Artificial Intelligence*, 268–272.

Aizawa, K., & Adams, F. (2009). Why the Mind Is Still in the Head. In P. Robbins & M. Aydede (Eds.), *The Cambridge Handbook of Situated Cognition* (pp. 78–95). Cambridge University Press.

Alcoff, L. M. (2006). *Visible Identities: Race, Gender, and the Self*. Oxford University Press.

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

Anderson, M., & Chemero, A. (2019). The world well gained: On the epistemic implications of ecological information. In M. Colombo, L. Irvine, & M. Stapleton (Eds.), *Andy Clark and His Critics* (pp. 161–173). Oxford University Press.

Anderson, M. L. (2014). *After phrenology: Neural reuse and the interactive brain.* The MIT Press.

Arango, A. (2019a). From sensorimotor dependencies to perceptual practices: making enactivism social. *Adaptive Behavior*, *27*(1), 31–45. https://doi.org/10.1177/1059712318811897

Arango, A. (2019b). Social enactivism about perception—reply to McGann. *Adaptive Behavior*, *27*(2), 161–162. https://doi.org/10.1177/1059712319835162

Attinger, A., Wang, B., & Keller, G. B. (2017). Visuomotor Coupling Shapes the Functional Development of Mouse Visual Cortex. *Cell*, *169*(7), 1291-1302.e14. https://doi.org/10.1016/j.cell.2017.05.023

Auvray, M., & Myin, E. (2009). Perception with compensatory devices: From sensory substitution to sensorimotor extension. *Cognitive Science*, *33*(6), 1036–1058. https://doi.org/10.1111/j.1551-6709.2009.01040.x

Bach-y-Rita, P. (1996). Nonsynaptic Diffusion Neurotransmission and Brain Plasticity. *The Neuroscientist*, *2*(5), 260–261. https://doi.org/10.1177/107385849600200511

Bach-y-Rita, P., Collins, C. C., Saunders, F. A., White, B., & Scadden, L. (1969). Vision Substitution by Tactile Image Projection. *Nature*, *221*(5184), 963–964. https://doi.org/10.1038/221963a0

Bar, M. (2009). The proactive brain: memory for predictions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521).

Beaton, M. (2016). Sensorimotor direct realism: How we enact our world. *Constructivist Foundations*, *11*(2), 265–276.

Bechtel, W. (2008). *Mental Mechanisms. Philosophical Perspective on Cognitive Neuroscience*. Routledge. Taylor & Francis Group.

Beer, R. D. (2000). Dynamical approaches to cognitive science. *Trends in Cognitive Sciences*, *4*(3), 91–99. https://doi.org/10.1016/S1364-6613(99)01440-0

Berger, J., & Nanay, B. (2016). Relationalism and unconscious perception. *Analysis (United Kingdom)*, *76*(4), 426–433. https://doi.org/10.1093/analys/anw060

Bergson, H. (1988). *Matter and Memory* (N. M. Paul & W. S. Palmer (eds.)). Zone Books.

Bermejo, F., Hüg, M. X., & Di Paolo, E. (2020). Rediscovering Richard Held: Activity and Passivity in Perceptual Learning. *Frontiers in Psychology*, *11*(May), 1–14. https://doi.org/10.3389/fpsyg.2020.00844

Bermúdez, J. L. (1998). *The Paradox of Self-Consciousness*. The MIT Press. Bradford Books. https://doi.org/10.1017/CBO9781107415324.004

Bermúdez, J. L. (2007). From two visual systems to two forms of content? *Psyche*, *13*(2007).

Block, N. (1978). Troubles with Functionalism. *Minnesota Studies in the Philosophy of Science*, *9*, 261–325. https://doi.org/10.7551/mitpress/2111.003.0006

Brancazio, N. (2019). Gender and the senses of agency. *Phenomenology and the Cognitive Sciences*, *18*(2), 425–440. https://doi.org/10.1007/s11097-018-9581-z

Brooks, R. A. (1991). Intelligence Without Reason. In *Computers and Thought* (Vol. 0, Issue

April).

Bruineberg, J., Kiverstein, J., & Rietveld, E. (2018). The anticipating brain is not a scientist: the free-energy principle from an ecological-enactive perspective. *Synthese*, *195*, 2417–2444. https://doi.org/10.1007/s11229-016-1239-1

Buhrmann, T., Di Paolo, E. A., & Barandiaran, X. (2013). A dynamical systems account of sensorimotor contingencies. *Frontiers in Psychology*, *4*(MAY), 1–19. https://doi.org/10.3389/fpsyg.2013.00285

Butler, J. (1989). Sexual Ideology and Phenomenological Description. In J. Allen & I. M. Young (Eds.), *The Thinking Muse. Feminism and Modern French Philosophy* (pp. 85–100). Indiana University Press.

Campbell, J. (2008). Sensorimotor Knowledge and Naïve Realism. *Philosophy and Phenomenological Research*, *LXXVI*(3), 666–673.

Cappuccio, M., & Wheeler, M. (2012). Ground-Level Intelligence: Action-Oriented Representation and the Dynamics of the Background. In Z. Radman (Ed.), *Knowing without Thinking* (pp. 13–36). Palgrave-Macmillan.

Chemero, A. (2000). Anti-Representationalism and the Dynamical Stance. *Philosophy of Science*, *67*(4), 625–647.

Chemero, A. (2009). *Radical Embodied Cognitive Science*. The MIT Press. Bradford Books.

Chua, H. F., Boland, J. E., & Nisbett, R. E. (2005). Cultural variation in eye movements during scene perception. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(35), 12629–12633. https://doi.org/10.1073/pnas.0506162102

Clark, A. (1997). *Being There. Putting Brain, Body, and World Together Again.* The MIT Press. Bradford Books.

Clark, A. (2002). Is Seeing All It Seems? Action, Reason, and the Grand Illusion. *Journal of Consciousness Studies*, *9*(5), 181–202.

Clark, A. (2008). *Supersizing the Mind. Embodiment, Action, and Cognitive Extension*. Oxford University Press. https://doi.org/10.1017/CBO9781107415324.004

Clark, A. (2009). Spreading the joy? Why the machinery of consciousness is (Probably) still in the head. *Mind*, *118*(472), 963–993. https://doi.org/10.1093/mind/fzp110

Clark, A. (2013a). Expecting the World: Perception, Prediction, and the Origins of Human Knowledge. *Journal of Philosophy*, *CX*(9), 469–496.

Clark, A. (2013b). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Clark, A. (2016). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind* (p. 424). https://doi.org/10.1093/acprof:oso/9780190217013.001.0001

Clark, A. (2017a). Busting Out: Predictive Brains, Embodied Minds, and the Puzzle of the Evidentiary Veil. *Nous*, *51*(4), 727–753. https://doi.org/10.1111/nous.12140

Clark, A. (2017b). How to Knit Your Own Markov Blanket : Resisting the Second Law with Metamorphic Minds. *Philosophy and Predictive Coding*, 1–19. https://doi.org/10.15502/9783958573031

Clark, A. (2019). Replies to Critics. In Search of the Embodied, Extended, Enactive, Predictive (EEE- P) Mind. In M. Colombo, L. Irvine, & M. Stapleton (Eds.), *Andy Clark and His Critics* (pp. 266–302). Oxford University Press.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, *58*(January), 7–19. https://doi.org/10.1111/j.1467-9744.2009.01021.x

Clark, A., & Thornton, C. (1997). Trading spaces: Computation, representation, and the limits of uninformed learning. *Behavioral and Brain Sciences*, *20*(1), 57–66. https://doi.org/10.1017/S0140525X97000022

Clark, A., & Toribio, J. (1994). Doing without representing? *Synthese*, *101*(Connectionism and the frontiers of Artificial Intelligence), 401–431.

Clark, A., & Toribio, J. (2001). Sensorimotor Chauvinism? Commentary on O'Reagan, J. Kevin and Noë, Alva, "A Sensorimotor account of vision and Visual Consciousness." *Behavioral and Brain Sciences*, *24*(5), 979–980.

Clavel Vázquez, M. J. (2019). A match made in heaven: predictive approaches to (an unorthodox) sensorimotor enactivism. *Phenomenology and the Cognitive Sciences*, 653–684. https://doi.org/10.1007/s11097-019-09647-0

Colombetti, G. (2013). *The Feeling Body: Affective Science Meets the Enactive Mind*. 288. https://doi.org/10.7551/mitpress/9780262019958.001.0001

Cooke, E. D., & Myin, E. (2011). Is trilled smell possible? How the structure of olfaction determines the phenomenology of smell. *Journal of Consciousness Studies*, *18*(11–12), 59–95.

De Beauvoir, S. (1997). *The Second Sex* (H. M. Parshley (ed.)). Vintage.

De Beauvoir, S. (2004). A Review of The Phenomenology of Perception by Maurice Merleau-Ponty. In M. A. Simons (Ed.), *Philosophical Writings* (pp. 151–164). University of Illinois Press.

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

De Jaegher, H., Di Paolo, E., & Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Sciences*, *14*(10), 441–447. https://doi.org/10.1016/j.tics.2010.06.009

Degenaar, J., & Myin, E. (2014). Representation-hunger reconsidered. *Synthese*, *191*(15), 3639–3648. https://doi.org/10.1007/s11229-014-0484-4

Dennett, D. (1991). *Consciousness Explained*. Back Bay Books.

Di Paolo, E. (2014). The worldly constituents of perceptual presence. *Frontiers in Psychology*, *5*(MAY), 5–6. https://doi.org/10.3389/fpsyg.2014.00450

Di Paolo, E. (2020). Enactive becoming. *Phenomenology and the Cognitive Sciences*. https://doi.org/10.1007/s11097-019-09654-1

Di Paolo, E. A. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, *4*(4), 429–452. https://doi.org/10.1007/s11097-005-9002-y

Di Paolo, E., Barandiaran, X. E., Beaton, M., & Buhrmann, T. (2014). Learning to perceive in the sensorimotor approach: Piaget's theory of equilibration interpreted dynamically. *Frontiers in Human Neuroscience*, *8*(JULY), 1–16. https://doi.org/10.3389/fnhum.2014.00551

Di Paolo, E., Buhrmann, T., & Barandiaran, X. (2017). *Sensorimotor Life: An enactive proposal*. Oxford University Press. https://doi.org/10.1093/acprof

Di Paolo, E., Cuffari, E. C., & De Jaegher, H. (2018). Linguistic Bodies. The continuity between Life and Language. In *Psikologi Perkembangan* (Issue October 2013). The MIT Press. https://doi.org/10.1017/CBO9781107415324.004

DiFrisco, J. (2014). Hylomorphism and the Metabolic Closure Conception of Life. *Acta Biotheoretica*, *62*(4), 499–525. https://doi.org/10.1007/s10441-014-9233-9

Downey, A. (2017). Radical sensorimotor enactivism and predictive processing: Providing a conceptual framework for the scientific study of conscious perception. *Philosophy and Predictive Processing*, 173–188. https://doi.org/10.15502/9783958573123

Downey, A. (2018). Predictive processing and the representation wars: a victory for the eliminativist (via fictionalism). *Synthese*, *195*(12), 5115–5139. https://doi.org/10.1007/s11229-017-1442-8

Drayson, Z. (2017). What is action-oriented perception? *Logic, Methodology and Philosophy of Science: Proceedings of the 15th International Congress*, 1–14.

Egan, F. (2014). How to think about mental content. *Philosophical Studies*, *170*(1), 115–135. https://doi.org/10.1007/s11098-013-0172-0

Einstein, G. (2012). Situated Neuroscience: Exploring Biologies of Diversity. In R. Bluhm, A. J. Jacobson, & H. L. Maibom (Eds.), *Neurofeminims. Issues at the Intersection of Feminist Theory and Cognitive Science* (pp. 145–174). Palgrave-Macmillan.

Flament-Fultot, M. (2016). Counterfactuals versus Constraints. Towards an Implementation Theory of Sensorimotor Mastery. *Journal of Consciousness Studies*, *23*(5–6), 153–176.

Flor, N. V., & Hutchins, E. L. (1991). Analyzing Distributed Cognition in Software Teams: A Case Study of Team Programming During Perfecftive Software Maintenance. In J. Koeneman-Belliveau, T. G. Moher, & S. Roberston (Eds.), *Proceedings of the Fourth Annual*

Workshop on Empirical Studies of Programmers (pp. 35–59). Ablex Publishing.

Franceschini, N., Pichon, J. M., & Blanes, C. (1992). From insect vision to robot vision. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *337*(1281), 283–294. https://doi.org/10.1098/rstb.1992.0106

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622

Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends in Cognitive Sciences*, *13*(7), 293–301. https://doi.org/10.1016/j.tics.2009.04.005

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Friston, K. (2012). Embodied inference and spatial cognition. *Cognitive Process*, *13*. https://doi.org/10.1007/s10339-012-0519-z

Friston, K. (2013). Life as we know it. *Journal of The Royal Society Interface*, *10*(86), 20130475–20130475. https://doi.org/10.1098/rsif.2013.0475

Friston, K., Adams, R. A., Perrinet, L., & Breakspear, M. (2012). Perceptions as hypotheses: Saccades as experiments. *Frontiers in Psychology*, *3*(MAY), 1–20. https://doi.org/10.3389/fpsyg.2012.00151

Friston, K., & Frith, C. (2015). A Duet for one. *Consciousness and Cognition*, *36*, 390–405. https://doi.org/10.1016/j.concog.2014.12.003

Friston, K., & Stephan, K. E. (2007). Free-energy and the brain. *Synthese*, *159*(3), 417–458. https://doi.org/10.1007/s11229-007-9237-y

Froese, T., & Ortiz-Garin, G. U. (2020). Where Is the Action in Perception? An Exploratory Study With a Haptic Sensory Substitution Device. *Frontiers in Psychology*, *11*(April), 1–8. https://doi.org/10.3389/fpsyg.2020.00809

Gallagher, S. (2005). *How the Body Shapes the Mind*. Oxford University Press. https://doi.org/10.1093/0199271941.001.0001

Gallagher, S., & Allen, M. (2018). Active inference, enactivism and the hermeneutics of social cognition. *Synthese*, *195*(6), 2627–2648. https://doi.org/10.1007/s11229-016-1269-8

Gangopadhyay, N. (2010). Experiential blindness revisited: In defence of a case of embodied cognition. *Cognitive Systems Research*, *11*(4), 396–407. https://doi.org/10.1016/j.cogsys.2010.05.003

Gibson, J. J. (1972). A theory of direct perception. In J. R. Royce & W. W. Rozeboom (Eds.), *The psychology of knowing* (pp. 215–240). Gordon & Breach.

Gibson, J. J. (2015). The Ecological Approach to Visual Perception. In *Paths to International Political Economy (Routledge Revivals)*. Taylor & Francis.

Gilbert, C. D., & Li, W. (2013). Top-down influences on visual processing. *Nature Review Neuroscience*, *14*(5). https://doi.org/10.1038/nrn3476.Top-down

Gładziejewski, P. (2016). Predictive coding and representationalism. *Synthese*, *193*(2), 559–582. https://doi.org/10.1007/s11229-015-0762-9

Godfrey-Smith, P. (1994). Spencer and Dewey on Life and Mind. In R. Brooks & P. Maes (Eds.), *Artificial Life IV* (pp. 80–89). The MIT Press.

Gregory, R. L., & Wallace, J. G. (1963). *Recovery from early blindness: A case study. Experimental Psychology Society Monograph No. 2*.

Grice, P. (1962). Some Remarks about the Senses. In *Studies in the Way of Words* (pp. 248–268). Harvard University Press. https://doi.org/10.7551/mitpress/7111.003.0005

Heidegger, M. (1996). *Being and Time* (J. Stambaugh (ed.)). State University of New York Press.

Heidegger, M. (2001). *Zollikon Seminars. Protocols - Conversations - Letters.* (F. Mayr & R. Askav (eds.)). Northwestern University Press.

Held, R., & Hein, A. (1963). Movement-Produced Stimulation in the Development of Visually Guided Behavior. *Journal of Comparative and Physiological Psychology*, *56*(5), 872–876. https://doi.org/10.1037/h0040546

Helmholtz, H. von. (2005). *Treatise on Physiological Optics , Volume III*. Dover. https://books.google.co.uk/books/about/Treatise_on_Physiological_Optics.html?id=cSjEAgAAQBAJ&redir_esc=y

Higgins, J. (2018). Biosocial selfhood: overcoming the ‘body-social problem’ within the individuation of the human self. *Phenomenology and the Cognitive Sciences*, *17*(3), 433–454. https://doi.org/10.1007/s11097-017-9514-2

Hohwy, J. (2013). *The predictive mind*. Oxford University Press. https://doi.org/10.1017/CBO9781107415324.004

Hohwy, J. (2016). The self-evidencing brain. *Nous*, *50*(2), 259–285. https://doi.org/10.1111/nous.12062

Hohwy, J. (2017). How to Entrain Your Evil Demon. In T. Metzinger & W. Wiese (Eds.), *Philosophy and predictive coding* (pp. 1–15). MIND Group. https://doi.org/10.15502/9783958573048

Hohwy, J. (2019). Quick’n’Lean or Slown and Rich. In M. Colombo, E. Irvine, & M. Stapleton (Eds.), *Andy Clark and His Critics* (pp. 191–205). Oxford University Press.

Huang, Y., & Rao, R. P. N. (2011). Predictive coding. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(5), 580–593. https://doi.org/10.1002/wcs.142

Hurley, S. (1998). *Consciousness in Action*. Harvard University Press.

Hurley, S. (2001). Perception and action: alternative views. *Synthese*, *129*, 3–40.

Hurley, S. (2010). The Varieties of Externalism. In R. Menary (Ed.), *The Extended Mind*. https://doi.org/10.7551/mitpress/9780262014038.001.0001

Hurley, S., & Noë, A. (2003). Neural Plasticity and Consciousness. *Biology and Philosophy*, *18*(September), 131–168. https://doi.org/10.3389/fpsyg.2011.00191

Husserl, E. (1989). *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy. Second Book. Studies in the Phenomenology of Constitution* (R. Rojcewicz & A. Schuwer (trans.) (eds.)). Kluwer Academic Publishers.

Husserl, E. (1991). *On the phenomenology of the consciousness of internal time (1893-1917)* (J. B. (Trans. . Brough (ed.)). Kluwer Academic Publishers. https://doi.org/10.2307/j.ctvh4zhv9

Hutto, D. D., & Myin, E. (2012). *Radicalizing Enactivism: Basic Minds without Content*. MIT Press.

Hutto, D. D., & Myin, E. (2017). *Evolving Enactivism. Basic Minds Meet Content*. The MIT Press. https://doi.org/10.1017/CBO9781107415324.004

Jacob, P., & Jeannerod, M. (2003). *Ways of seeing. The scope and limits of visual cognition*. Oxford University Press.

Jacobson, A. J. (2012). Seeing as a Social Phenomenon: Feminist Theory and the Cognitive Sciences. In R. Bluhm, A. J. Jacobson, & H. L. Maibom (Eds.), *Neurofeminism. Issues at the Intersection of Feminist Theory and Cognitive Science* (pp. 216–229). Palgrave-Macmillan.

Jeannerod, M., & Jacob, P. (2005). Visual cognition: A new look at the two-visual systems model. *Neuropsychologia*, *43*(2 SPEC. ISS.), 301–312. https://doi.org/10.1016/j.neuropsychologia.2004.11.016

Jenkins, K. (2018). Toward an Account of Gender Identity. *Ergo, an Open Access Journal of Philosophy*, *5*(20200523), 713–744. https://doi.org/10.3998/ergo.12405314.0005.027

Jonas, H. (1966). *The Phenomenon of Life. Toward a Philosophical Biology*. Harper & Row Publishers. Reprinted Northwestern University Press, 2001.

Kiefer, A., & Hohwy, J. (2018). Content and misrepresentation in hierarchical generative models. *Synthese*, *195*(6), 2387–2415. https://doi.org/10.1007/s11229-017-1435-7

Kirchhoff, M. D. (2018a). Autopoiesis, free energy, and the life–mind continuity thesis. *Synthese*, *195*(6), 2519–2540. https://doi.org/10.1007/s11229-016-1100-6

Kirchhoff, M. D. (2018b). Predictive processing, perceiving and imagining: Is to perceive to imagine, or something close to it? *Philosophical Studies*, *175*(3), 751–767. https://doi.org/10.1007/s11098-017-0891-8

Kirchhoff, M. D., & Kiverstein, J. (2019a). Extended Consciousness and Predictive Processing. In *Extended Consciousness and Predictive Processing*. Routledge. https://doi.org/10.4324/9781315150420

Kirchhoff, M. D., & Kiverstein, J. (2019b). How to determine the boundaries of the mind: a Markov blanket proposal. *Synthese*, *0123456789*. https://doi.org/10.1007/s11229-019-02370-y

Kiverstein, J. (2010). Sensorimotor knowledge and the contents of experience. In N. Gangopadhyay, M. Madary, & F. Spicer (Eds.), *Perception, Action, and Consciousness: Sensorimotor Dynamics and Two Visual Systems* (pp. 259–271). Oxford University Press. https://doi.org/10.1093/acprof

Kiverstein, J. (2020). Free Energy and the Self: An Ecological–Enactive Interpretation. *Topoi*, *39*(3), 559–574. https://doi.org/10.1007/s11245-018-9561-5

Kiverstein, J., & Rietveld, E. (2018). Reconceiving representation-hungry cognition: an ecological-enactive proposal. *Adaptive Behavior*, *26*(4), 147–163. https://doi.org/10.1177/1059712318772778

Klinnert, M. D., Campos, J. J., Sorce, J. F., Emde, R. N., & Svejda, M. (1983). Emotions as Behaviour Regulators: Social Referencing in Infancy. In R. Plutchik & H. Kellerman (Eds.), *Emotions in Early Development* (pp. 57–86). Elsevier. https://doi.org/10.1016/b978-0-12-558702-0.50009-1

Kwisthout, J., & van Rooij, I. (2013). Bridging the gap between theory and practice of approximate Bayesian inference. *Cognitive Systems Research*, *24*, 2–8. https://doi.org/10.1016/j.cogsys.2012.12.008

Kwisthout, J., & van Rooij, I. (2020). Computational Resource Demands of a Predictive Bayesian Brain. *Computational Brain & Behavior*, *3*(2), 174–188. https://doi.org/10.1007/s42113-019-00032-3

Kyselo, M. (2014). The body social: An enactive approach to the self. *Frontiers in Psychology*, *5*(SEP), 1–16. https://doi.org/10.3389/fpsyg.2014.00986

Lakoff, G., & Johnson, M. (1999). *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought*. Basic Books.

Lohner, S. (2020). Find Your Blind Spot! A visual science project from Science Buddies. *Scientific American*. https://www.scientificamerican.com/article/find-your-blind-spot/

Loughlin, V. (2018a). Going wide: extended mind and Wittgenstein. *Adaptive Behavior*, *26*(6), 275–283. https://doi.org/10.1177/1059712318770387

Loughlin, V. (2018b). Sensorimotor theory, cognitive access and the ‘absolute’ explanatory gap.

Phenomenology and the Cognitive Sciences, *17*, 611–627. https://doi.org/10.1007/s11097-017-9543-x

Loughlin, V., & Zahidi, K. (2017). What Is Left of the Active Externalism Debate? *European Journal of Philosophy*, *25*(4), 1614–1639. https://doi.org/10.1111/ejop.12249

Mach, E. (1959). *The Analysis of Sensation* (C. M. Williams (Trans.) (ed.)). Dover.

Macpherson, F. (2011). Individuating the senses. In F. Macpherson (Ed.), *The Senses: Classical and Contemporary Philosophical Perspectives*. Oxford University Press.

Macpherson, F. (2015). The space of sensory modalities. In D. Stokes, M. Matthen, & S. Biggs (Eds.), *Perception and its Modalities*. Oxford University Press.

Macpherson, F. (2017). The relationship between cognitive penetration and predictive coding. *Consciousness and Cognition*, *47*, 6–16. https://doi.org/10.1016/j.concog.2016.04.001

Malafouris, L. (2004). The Cognitive Basis of Material Engagement: Where Brain, Body and Culture Conflate. *Rethinking Materiality: The Engagement of Mind with the Material World*, 53–62.

Martinez-Conde, S., Krauzlis, R., Miller, J. M., Morrone, Concetta Williams, D., & Kowler, E. (2008). Eye Movements and the Perception of a Clear and Stable Visual World. *Journal of Vision*, *8*(14), 1.

Martínez Quintero, A., & De Jaegher, H. (2020). Pregnant Agencies: Movement and Participation in Maternal – Fetal Interactions. *Frontiers in Psychology*, *11*, 1–16. https://doi.org/10.3389/fpsyg.2020.01977

Masrour, F. (2019). On the Possibility of Hallucinations. *Mind*, *0*(February). https://doi.org/10.1093/mind/fzy088

McCauley, R. N., & Henrich, J. (2006). Susceptibility to the Müller-Lyer Illusion, theory-neutral observation, and the diachronic penetrability of the visual input system. *Philosophical Psychology*, *19*(1), 79–101. https://doi.org/10.1080/09515080500462347

McDowell, J. (1994). The Content of Perceptual Experience. *The Philosophical Quarterly*, *44*(175), 190–205.

McGann, M. (2019). A little more social enaction than that: comment on Arango (2019). *Adaptive Behavior*, *27*(2), 157–159. https://doi.org/10.1177/1059712319830650

Merleau-Ponty, M. (1963). *The Structure of Behavior* (L. A. (trans. . Fisher (ed.)). Beacon Press.

Merleau-Ponty, M. (2002). *Phenomenology of Perception* (C. (Trans. . Smith (ed.)). Routledge.

Miłkowski, M. (2016). Computation and Multiple Realizability. In V. Müller (Ed.), *Fundamental Issues of Artificial Intelligence. Synthese Library (Studies in Epistemology, Logic, Methodology, and Philosophy of Science), vol. 376* (pp. 29–41). Springer. https://doi.org/https://doi.org/10.1007/978-3-319-26485-1_3

Milkowski, M., Clowes, R., Rucinska, Z., Przegalinska, A., Zawidzki, T., Krueger, J., Gies, A., McGann, M., Afeltowicz, L., Wachowski, W., Stjernberg, F., Loughlin, V., & Hohol, M. (2018). From wide cognition to mechanisms: A silent revolution. *Frontiers in Psychology*, *9*(DEC), 1–17. https://doi.org/10.3389/fpsyg.2018.02393

Millar, B. (2020). Towards a sensorimotor approach to flavour and smell. *Mind and Language*, *June 2019*, 1–20. https://doi.org/10.1111/mila.12275

Milner, D. A. (2008). Conscious and unconscious visual processing in the human brain. In L. Weiskrantz & M. Davies (Eds.), *Frontiers of Consciousness: Chichele Lectures* (pp. 169–214). Oxford University Press.

Milner, D. A., & Goodale, M. A. (2006). *The Visual Brain in Action* (2nd Editio). Oxford University Press. https://doi.org/10.1017/CBO9781107415324.004

Milner, D. A., & Goodale, M. A. (2008). Two visual systems re-viewed. *Neuropsychologia*, *46*(3), 774–785. https://doi.org/10.1016/j.neuropsychologia.2007.10.005

Milner, D. A., & Goodale, M. A. (2010). Cortical visual systems for perception and action. In F. Gangopadhyay, Nivedita; Madary, Michael; Spicer (Ed.), *Perception, Action, and Consciousness: Sensorimotor Dynamics and Two Visual Systems* (pp. 71–94). Oxford University Press. https://doi.org/10.1093/acprof

Morales, J., Bax, A., & Firestone, C. (2020). Sustained representation of perspectival shape. *Proceedings of the National Academy of Sciences*, *117*(26), 202000715. https://doi.org/10.1073/pnas.2000715117

Murphy, C. M., & Messner, D. J. (1977). Mothers, infants and pointing : A study of a gesture. In H. R. Schaffer (Ed.), *Studies in mother-infant interaction: Proceedings of the Loch Lomond symposium, Ross Priory, University of Strathclyde, September, 1975*. Academic Press. https://catalog.hathitrust.org/Record/000138543?

Myin, E. (2016). Perception as Something We Do. *Journal of Consciousness Studies*, *23*(5–6), 80–104.

Myin, E., Cooke, E., & Zahidi, K. (2014). Morphing Senses. In Dustin Stokes, M. Matthen, & S. Biggs (Eds.), *Perception and its Modalities* (pp. 394–409). Oxford University Press. https://doi.org/10.1093/acprof

Myin, E., & Loughlin, V. (2018). Sensorimotor and Enactive Approaches to Consciousness. In R. J. Gennaro (Ed.), *The Routledge Handbook of Consciousness* (pp. 202–215). Routledge.

Nanay, B. (2015). The Representationalism versus Relationalism Debate: Explanatory Contextualism about Perception. *European Journal of Philosophy*, *23*(2), 321–336. https://doi.org/10.1111/ejop.12085

Nanay, B. (2018). The Importance of Amodal Completion in Everyday Perception. *I-Perception*, *9*(4). https://doi.org/10.1177/2041669518788887

Nazir, T. A., & O'Regan, J. K. (1990). Some results on translation invariance in the human visual system. *Spatial Vision*, *5*(2), 81–100.

Newen, A., de Bruin, L., & Gallagher, S. (2018). 4E Cognition. Historical Roots, Key Concepts, and Central Issues. In A. Newen, L. de Bruin, & S. Gallagher (Eds.), *The Oxford Handbook of 4E Cognition* (pp. 3–15). Oxford University Press.

Newson, J., & Newson, E. (1975). Intersubjectivity and the transmission of culture: On the social origins of symbolic functioning. - PsycNET. *Bulletin of the British Psychological Society*, *28*, 437–446. https://psycnet.apa.org/record/1979-03234-001

Noë, A. (2004). *Action in Perception*. The MIT Press. https://doi.org/10.1068/p5878

Noë, A. (2007). The critique of pure phenomenology. *Phenomenology and the Cognitive Sciences*, *6*(1–2), 231–245. https://doi.org/10.1007/s11097-006-9043-x

Noë, A. (2008). Precis of Action In Perception. *Philosophy and Phenomenological Research*, *LXXVI*(3), 660–665.

Noë, A. (2009). *Out of Our Heads. Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness*. Hill and Wang.

Noë, A. (2010). Vision without representation. In N. Gangopadhyay, M. Madary, & F. Spicer (Eds.), *Perception, Action, and Consciousness: Sensorimotor Dynamics and Two Visual Systems* (pp. 246–256). Oxford University Press. https://doi.org/10.1093/acprof

Noë, A. (2012). *Varieties of Presence*. Harvard University Press.

Nuñez, R. (1999). Could the Future Taste Purple? *Journal of Consciousness Studies*, *6*(11–12), 41–60.

O'Regan, J. K. (1992). Solving the "real" mysteries of visual perception: the world as an outside memory. *Canadian Journal of Psychology*, *46*(3), 461–488. https://doi.org/10.1037/h0084327

O'Regan, J. K. (2011). *Why Red Doesn't Sound Like a Bell: Understanding the Feel of Consciousness*. Oxford University Press. https://doi.org/10.1093/acprof

O'Regan, J. K., & Degenaar, J. (2014). Predictive processing, perceptual presence, and sensorimotor theory. In *Cognitive Neuroscience* (Vol. 5, Issue 2, pp. 130–131). Psychology Press Ltd. https://doi.org/10.1080/17588928.2014.907256

O'Regan, J. K., Myin, E., & Noë, A. (2004). Towards an Analytic Phenomenology: The Concepts of "Bodiliness" and "Grabbiness." In A. Carsetti (Ed.), *Seeing, Thinking and Knowing. Meaning and Self-Organisation in Visual Cognition and Thought* (pp. 103–114). Kluwer Academic Publishers. https://doi.org/10.1007/1-4020-2081-3

O'Regan, J. K., Myin, E., & Noë, A. (2005a). Sensory consciousness explained (better) in terms of "corporality" and "alerting capacity." *Phenomenology and the Cognitive Sciences*, *4*(4), 369–387. https://doi.org/10.1007/s11097-005-9000-0

O'Regan, J. K., Myin, E., & Noë, A. (2005b). Skill, corporality and alerting capacity in an account of sensory consciousness. *Progress in Brain Research*, *150*, 55–68. https://doi.org/10.1016/S0079-6123(05)50005-0

O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, *24*, 939–1031. https://doi.org/10.1017/S0140525X01000115

O'Regan, J. K., Rensink, R. A., & Clark, J. J. (1999). Change-blindness as a result of "mudsplashes." *Nature*, *398*(6722), 34. https://doi.org/10.1038/17953

Oksala, J. (2006). A phenomenology of gender. *Continental Philosophy Review*, *39*(3), 229–244. https://doi.org/10.1007/s11007-006-9025-2

Oliver, M. (2009). Understanding disability : from theory to practice. In *Palgrave Macmillan*.

Olkowski, D. E. (2017). Using our intuition. Creating the future phenomenological plane of thought. In H. A. Fielding & D. E. Olkowski (Eds.), *Feminist Phenomenology Futures* (pp. 3–20). Indiana University Press.

Orlandi, N. (2014). *The Innocent Eye: Why Vision is Not a Cognitive Process*. Oxford University Press.

Orlandi, N. (2016). Bayesian Perception Is Ecological Perception. *Philosophical Topics*, *44*(2), 255–278. https://doi.org/10.5840/philtopics201644226

Peacocke, C. (1983). *Sense and Content*. Oxford University Press.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems. Networks of Plausible Inference*. Morgan Kaufmann. https://doi.org/https://doi.org/10.1016/C2009-0-27609-4

Prinz, J. (2006). Putting the brakes on Enactive Perception. *Psyche*, *12*(1), 1–19. https://doi.org/10.18632/oncotarget.5242

Prinz, J. (2009). Is Consciousness Embodied? In P. Robbins & M. Aydede (Eds.), *Cambridge Handbook of Situated Cognition* (pp. 419–437). Cambridge University Press.

Pylyshyn, Z. W. (2000). Situating vision in the world. *Trends in Cognitive Sciences*, *4*(5), 197–207. https://doi.org/10.1016/S1364-6613(00)01477-7

Pylyshyn, Z. W. (2001). Seeing, acting, and knowing. *Behavioral and Brain Sciences*, *24*(5), 999.

Ramsey, W. (2007). *Representation Reconsidered*. Cambridge University Press. https://doi.org/10.1080/09515089.2010.529261

Ramsey, W. (2015). Must cognition be representational? *Synthese*, *December 2014*.

https://doi.org/10.1007/s11229-014-0644-6

Ramstead, M. J. D., Kirchhoff, M. D., & Friston, K. (2019). A tale of two densities: active inference is enactive inference. *Adaptive Behavior*, *28*(4), 225–239. https://doi.org/10.1177/1059712319862774

Rauss, K., & Pourtois, G. (2013). What is bottom-up and what is top-down in predictive coding. *Frontiers in Psychology*, *4*(MAY), 1–8. https://doi.org/10.3389/fpsyg.2013.00276

Roepstorff, A., Niewöhner, J., & Beck, S. (2010). Enculturing brains through patterned practices. *Neural Networks*, *23*(8–9), 1051–1059. https://doi.org/10.1016/j.neunet.2010.08.002

Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature*, *253*(5489), 265–266. https://doi.org/10.1038/253265a0

Scheier, C., & Pfeifer, R. (1998). Exploiting embodiment for category learning. In R. Pfeifer, B. Blumberg, J.-A. Meyer, & S. W. Wilson (Eds.), *From Animals to Aminats 5. Proceedings of the Fifth International Conference on Simulation of Adaptive Behaviour* (pp. 32–37). The MIT Press.

Schellenberg, S. (2007). Action and self-location in perception. *Mind*, *116*(463), 603–631. https://doi.org/10.1093/mind/fzm603

Segall, M. H., Campbell, D. T., & Herskovits, M. J. (1966). *The influence of culture on visual perception*. Bobbs-Merrill Co. https://catalog.hathitrust.org/Record/000662263?

Seth, A. K. (2014). A predictive processing theory of sensorimotor contingencies: Explaining the puzzle of perceptual presence and its absence in synesthesia. *Cognitive Neuroscience*, *5*(2), 97–118. https://doi.org/10.1080/17588928.2013.877880

Seth, A. K. (2015). The cybernetic bayesian brain. From Interoceptive Inference to Sensorimotor Contingencies. *Open MIND*, *35*, 1–24. https://doi.org/10.15502/9783958570108

Shuler, M. G., & Bear, M. F. (2006). Reward Timing in the Primary Visual Cortex. *Science*, *311*, 1606–1610. https://doi.org/10.7551/mitpress/8876.003.0036

Silverman, D. (2016a). Representationalism and the Sensorimotor Theory. *Constructivist Foundations*, *11*(2), 282–284.

Silverman, D. (2016b). Tying the Knot: Why Representationalists should Endorse the Sensorimotor Theory of Conscious Feel. *Philosophical Quarterly*, *66*(263), 280–301. https://doi.org/10.1093/pq/pqv097

Silverman, D. (2018). Bodily skill and internal representation in sensorimotor perception. *Phenomenology and the Cognitive Sciences*, *17*(1), 157–173. https://doi.org/10.1007/s11097-017-9503-5

Simondon, G. (1995). *L'individu et sa genèse physico-biologique*. Jérôme Millon.

Simondon, G. (2005). *L'individuation à la lumière des notions de forme et d'information.* Jérôme Millon.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*, *28*(9), 1059–1074. https://doi.org/10.1068/p281059

Sims, A. (2017). The problems with prediction. The Dark Room Problem and the Scope Dispute. *Philosophy and Predictive Processing*, *23*, 373–390. https://doi.org/10.15502/9783958573246

Srinivasan, M. V., Laughlin, S. B., & Dubs, A. (1982). Predictive coding: a fresh view of inhibition in the retina. *Proceedings of the Royal Society London, B 216*(Ratliff 1965), 427–459.

Stapleton, M. (2013). Steps to a " Properly Embodied" cognitive science. *Cognitive Systems Research*, *22–23*, 1–11. https://doi.org/10.1016/j.cogsys.2012.05.001

Stern, D. N. (1985). *The interpersonal world of the infant*. Basic Books.

Strawson, P. F. (1979). Perception and its Objects. In G. F. Macdonald (Ed.), *Perception and Identity. Essays Presented to A.J. Ayer with his Replies to them* (pp. 41–60). Palgrave.

Taylor, P. C. (2004). *Race: a philosophical introduction*. Polity.

Thompson, E. (2005). Sensorimotor subjectivity and the enactive approach to experience. *Phenomenology and the Cognitive Sciences*, *4*(4), 407–427. https://doi.org/10.1007/s11097-005-9003-x

Thompson, E. (2007). *Mind in life. Biology, phenomenology, and the sciences of mind*. Harvard University Press. https://doi.org/10.1016/S0301-4770(08)61427-X

Thompson, E., & Varela, F. J. (2001). Radical embodiment: Neural dynamics and consciousness. *Trends in Cognitive Sciences*, *5*(10), 418–425. https://doi.org/10.1016/S1364-6613(00)01750-2

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The Embodied Mind. Cognitive Science and Human Experience*. The MIT Press. https://doi.org/10.1111/j.1468-0149.1965.tb01386.x

Vernazzani, A. (2019). The structure of sensorimotor explanation. *Synthese*, *196*, 4527–4553. https://doi.org/10.1007/s11229-017-1664-9

Walk, R. D., Shepherd, J. D., & Miller, D. R. (1988). Attention and the depth perception of kittens. *Bulletin of the Psychonomic Society*, *26*(3), 248–251. https://doi.org/10.3758/BF03337301

Walsh, K. S., McGovern, D. P., Clark, A., & O'Connell, R. G. (2020). Evaluating the neurophysiological evidence for predictive processing as a model of perception. *Annals of the New York Academy of Sciences*, *1464*(1), 242–268. https://doi.org/10.1111/nyas.14321

Ward, D. (2012). Why don't synaesthetic colours adapt away? *Philosophical Studies*, *159*(1), 123–138. https://doi.org/10.1007/s11098-010-9693-y

Ward, D., Roberts, T., & Clark, A. (2011). Knowing what we can do: Actions, intentions, and the construction of phenomenal experience. *Synthese*, *181*(3), 375–394. https://doi.org/10.1007/s11229-010-9714-6

Ward, D., & Stapleton, M. (2012). Es are Good: Cognition as Enacted , Embodied , Embedded , Affective and Extended. In F. Paglieri (Ed.), *Consciousness in Interaction: The role of the natural and social context in shaping consciousness* (pp. 89–104). John Benjamins Publishing Company.

Webb, B. (1996). A Cricket Robot. *Scientific American*, *275*(6), 94–99. https://doi.org/10.1017/CBO9781107415324.004

Wehrle, M. (2020). Being a body and having a body. The twofold temporality of embodied intentionality. *Phenomenology and the Cognitive Sciences*, *19*(3), 499–521. https://doi.org/10.1007/s11097-019-09610-z

Wheeler, M. (2005). *Reconstructing the Cognitive World. The Next Step*. The MIT Press. Bradford Books.

Wheeler, M. (2010a). In Defense of Extended Functionalism. In R. Menary (Ed.), *The Extended Mind* (pp. 245–270). The MIT Press. https://doi.org/10.1080/05568640903420913

Wheeler, M. (2010b). Minds, Things, and Materiality. In L. Malafouris & C. Renfrew (Eds.), *The Cognitive Life of Things: Recasting the Boundaries of the Mind* (pp. 29–38). McDonald Institute for Archaeological Research.

Wheeler, M. (2011). Mind in Life or Life in Mind? Making Sense of Deep Continuity. *Journal of Consciousness Studies*, *18*(5), 148–168.

Wheeler, M. (1997). Cognition's Coming Home: the Reunion of Life and Mind. In P. Husbands & I. Harvey (Eds.), *Fourth European Conference on Artificial Life* (pp. 10–19). The MIT Press.

Wiese, W. (2017a). Action Is Enabled by Systematic Misrepresentations. *Erkenntnis*, *82*(6), 1233–1252. https://doi.org/10.1007/s10670-016-9867-x

Wiese, W. (2017b). What are the contents of representations in predictive processing? *Phenomenology and the Cognitive Sciences*, *16*(4), 715–736. https://doi.org/10.1007/s11097-016-9472-0

Wiese, W., & Metzinger, T. (2017). Vanilla PP for Philosophers : A Primer on Predictive Processing. In T. Metzinger & W. Wiese (Eds.), *Philosophy and predictive processing* (Vol. 1, pp. 1–18). MIND Group. https://doi.org/10.15502/9783958573024

Williams, D. (2018). Predictive Processing and the Representation Wars. *Minds and Machines*, *28*(1), 141–172. https://doi.org/10.1007/s11023-017-9441-6

Williams, D., & Colling, L. (2018). From symbols to icons: the return of resemblance in the cognitive neuroscience revolution. *Synthese*, *195*(5), 1941–1967. https://doi.org/10.1007/s11229-017-1578-6

Young, I. M. (1980). Throwing Like a Girl: A Phenomenology of Feminine Body Comportment Motility and Spatiality. *Human Studies*, *3*, 137–156.

Zahavi, D. (2008). *Subjectivity and Selfhood: Investigating the First-Person Perspective*. The MIT Press.

Zahavi, D. (2010). Minimal self and narrative self: a distinction in need of refinement. In T. Fuchs, H. C. Sattel, & P. Henningsen (Eds.), *The Embodied Self: Dimensions, Coherence and Disorders* (pp. 3–11). Schattauer.