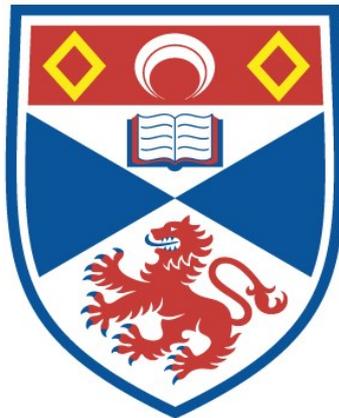


**BEING AND THINKING IN THE SOCIAL WORLD:  
PHENOMENOLOGICAL ILLUMINATIONS OF SOCIAL COGNITION  
AND HUMAN SELFHOOD**

**Joe Higgins**

**A Thesis Submitted for the Degree of PhD  
at the  
University of St Andrews**



**2017**

**Full metadata for this item is available in  
St Andrews Research Repository  
at:**

**<http://research-repository.st-andrews.ac.uk/>**

**Please use this identifier to cite or link to this item:**

**<http://hdl.handle.net/10023/10640>**

**This item is protected by original copyright**

# Being and Thinking in the Social World: phenomenological illuminations of social cognition and human selfhood

Joe Higgins



University of  
St Andrews

This thesis is submitted in partial fulfilment for the degree of PhD  
at the  
University of St Andrews

30<sup>th</sup> March 2017



**1. Candidate's declarations:**

I, **Joe Higgins**, hereby certify that this thesis, which is approximately **87500** 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 a higher degree.

I was admitted as a research student in **Sept 2013** and as a candidate for the degree of **PhD in Philosophy** in **Sept 2013**; the higher study for which this is a record was carried out in the University of St Andrews between **2013** and **2017**.

Date:                      signature of candidate:

**2. 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 Philosophy** in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date:                      signature of supervisor:

**3. Permission for publication:**

In submitting this thesis to the University of St Andrews I understand that I am 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. I also understand 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 my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

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

**PRINTED COPY**

a) No embargo on print copy

**ELECTRONIC COPY**

a) No embargo on electronic copy

**ABSTRACT AND TITLE EMBARGOES**

*An embargo on the full text copy of your thesis in the electronic and printed formats will be granted automatically in the first instance. This embargo includes the abstract and title except that the title will be used in the graduation booklet.*

If you have selected an embargo option indicate below if you wish to allow the thesis abstract and/or title to be published. If you do not complete the section below the title and abstract will remain embargoed along with the text of the thesis.

a) I agree to the title and abstract being published                      **YES**

Date                      signature of candidate

signature of supervisor

*Please note initial embargos can be requested for a maximum of five years. An embargo on a thesis submitted to the Faculty of Science or Medicine is rarely granted for more than two years in the first instance, without good justification. The Library will not lift an embargo before confirming with the student and supervisor that they do not intend to request a continuation. In the absence of an agreed response from both student and supervisor, the Head of School will be consulted. Please note that the total period of an embargo, including any continuation, is not expected to exceed ten years.*

*Where part of a thesis is to be embargoed, please specify the part and the reason.*



**Being and Thinking in the Social World: phenomenological illuminations of social cognition  
and human selfhood**

**ABSTRACT:**

At least since the time of Aristotle, it has been widely accepted that “man is by nature a social animal”. We eat, sleep, talk, laugh, cry, love, fight and create in ways that integrally depend on others and the social norms that we collectively generate and maintain.

Yet in spite of the widely accepted importance of human sociality in underlying our daily activities, its exact manifestation and function is consistently overlooked by many academic disciplines. Cognitive science, for example, regularly neglects the manner in which social interactions and interactively generated norms canalise and *constitute* our cognitive processes. Without the inescapable ubiquity of dynamic social norms, any given agent simply could not cognise *as a human*.

In this thesis, I aim to use a range of insights – from phenomenology, social psychology, neuroscience, cultural anthropology and gender studies – to clarify the role of sociality for human life. More specifically, the thesis can be broadly separated into three parts. I begin (chapters 1 and 2) with a broad explanation of how human agents are fundamentally tied to worldly entities and other agents in a way that characterises their ontological existence. In chapters 3 and 4, I criticise two recent and much-discussed theories of social cognition – namely, we-mode cognition and participatory sense-making – for failing to make intelligible the social constitution of human existence. In the later chapters (5-7), I then propose foundations for a more satisfactory theory of social cognition, as well as explicating a view of human selfhood as ‘biosocial’, such that even the autonomy of biological bodies is socially codified from a human perspective.

Taken together, the aforementioned chapters should contribute to calls for a new direction in social cognitive science, whilst also yielding novel insights into the nature of human selfhood.



## Contents

- Introduction	1
- Chapter 1: Heideggerian Cognitive Science and the Nature of Everyday Existence	7
- Chapter 2: Encountering Others	31
- Chapter 3: Interactive Social Cognition: the Cognitive We-mode	53
- Chapter 4: Interactive Social Cognition: Participatory Sense-making	73
- Chapter 5: The Pervasive Social Constitution of Human Existence	103
- Chapter 6: Finding the Social within the Social: a New Framework for (Anthro-)Social Cognition	131
- Chapter 7: The Biosocial Self	167
- Conclusion	197



# Introduction



## Introduction

What is human cognition? What is a human *subject*? How do such subjects interact with one another? And what is the ‘world’ in which these subjects live?

These are just some of the questions to which this thesis bears relevance. My hope is to shed light on the fundamental role of sociality – that is, social processes and structures – within human life. The motivation for addressing the fundamental role of sociality within human life is simply that many academic disciplines under-appreciate (at best), or completely overlook (at worst), its importance. Cognitive science, for example, regularly underestimates the manner in which social interactions and socially generated norms<sup>[1]</sup> canalise and constitute our cognitive processes. A similar underestimation is found in disciplines that are more directly socially oriented, such as anthropology, political studies, economics, gender studies and pedagogy. The scope and diversity of these disciplines hopefully makes it clear that the systematic neglect I am claiming is prevalent with regards to human sociality has far-reaching and practical implications beyond the confines of academia. To ignore or misunderstand the fundamental role that sociality plays in human life is to ignore or misunderstand *who* we are and, therefore, how we should live and how we should structure those systems on which we integrally depend.

In very broad terms, there are two key questions that underlie much of the content of this thesis:

- i. How, and to what extent, is human selfhood constituted by social processes and social structures?
- ii. How should we conceive of human social cognition?

In order to answer these questions, the thesis will progress as follows:

- The first two chapters lay the foundations for those that follow, focusing on human subjects’ fundamental connection to worldly entities and other agents in a way that characterises their existence. Chapter 1 will expound various aspects of mind and existence that I believe need to be accepted for a proper understanding of human subjects to be able to flourish. This initially involves an appreciation of the recent progression away from ‘Cartesian’ cognitive science and towards ‘Heideggerian’ cognitive science; that is, a progression away from individualistic, de-contextualised and cognitively isolated subjects towards enworlded, context-sensitive and constitutively social subjects. Subsequent claims address in detail the contextually embedded manner in which humans engage with worldly entities and the

---

<sup>[1]</sup> Throughout the thesis, it will become clear that I am using ‘norms’ in a very broad sense, to capture those principles by which human behaviour (and cognition) is deemed appropriate. My view is that these norms are generally implicit and always socially permeated, such that they simply encapsulate *what one does*, as much as *what one ought to do*.

omnipresence of social norms in these engagements with worldly entities. My intention is that these claims will help to frame the arguments that are presented in the following chapters.

- Chapter 2 considers how the enworlded, context-sensitive and constitutively social subjects that are expounded in chapter 1 are phenomenologically transformed during the incipient moments of encountering one another. This transformation involves the *reconfiguration* of a subject's phenomenological world, such that the subject's experience is dependent on another's (or others') presence, and the *reification* of the subject's existence, such that the subject becomes (non-reflectively) self-aware of her own agential presence. Part of the motivation for analysing this transformation is to pave the way for a claim that is fully developed in chapter 6: that human-human encounters differ importantly from other encounters, to the extent that the discipline of social cognition should be reconceived as *anthro-social* cognition. Chapter 2 ends by further developing the notion of the social-normative constitution of human existence that is introduced in the opening chapter.
- Across chapters 3 and 4, the analysis shifts somewhat away from phenomenology and towards the field of social cognition. My aim is to consider certain prominent current theories of social cognition and explain why a new approach to this nascent discipline is required. After a brief overview of 'orthodox' theories of social cognition, chapter 3 focuses on *we-mode cognition* (Gallotti and Frith, 2013a, 2013b) and argues that it (i). is incoherent in its presentation of "first-person plural" cognition, and (ii). fails to appreciate the unique transformative potential of other (physically co-present) humans.
- Chapter 4 continues the critical approach to current theories of social cognition, this time focusing on the revolutionary enactivist theory of *participatory sense-making* (De Jaegher and Di Paolo, 2007). Whilst participatory sense-making has many theoretical benefits, it is ambiguous in its presentation as a 'social' theory. This ambiguity will be expounded in virtue of the fact that the theory seeks to define itself as an explanans for how humans socially cognise, yet it fails to satisfactorily delineate 'social' processes from 'non-social' processes. A potential consequence of this is that participatory sense-making simply captures the dynamics that are involved in all cognitive acts, rather than capturing the unique dynamics of social engagements. As chapter 5 will show, there is something broadly correct about the idea that all cognition is social, but this is not how De Jaegher & Di Paolo intend the theory of participatory sense-making to be interpreted.
- Picking up where the criticisms of participatory sense-making end, chapter 5 looks at the pervasive constitution of human existence by social norms. In short, the claim is

that everything humans think, feel, or do is ‘social’. No act or thought can be de-contextualised from the norms of the social world; indeed, the very world that humans experience is only manifest within a social framework. In many ways (and as hinted at in chapter 2), this deep inherence of sociality within human life is nothing new and has taken on several guises within the phenomenological tradition (e.g. Husserl, 1952/1989, 1973; Heidegger, 1927/1962; Merleau-Ponty, 1945/1978; Scheler, 1913/2008). However, the presentation I put forward here differs from these in that I explicitly play out our social constitution in terms of normative structures that are actively modulated by individuals’ behaviour, drawing (briefly) on sociological and evolutionary theories for support.

- Whilst chapter 5 addresses the notion that humans (including all human cognition) are, in a very important way, social through and through, chapter 6 puts forward the claim that there is something unique about our cognitive experiences when engaging with other physically co-present humans. ‘Social cognition’ thus still merits its place as a distinct discipline within the ‘all-cognition-is-social’ framework that is an outcome of chapter 5. However, the ‘social cognition’ that I expound threads together a series of new disciplinary ideas. Drawing, in part, on the phenomenological transformations that are discussed in chapter 2, I firstly make a case for the field of social cognition to be confined to human-human engagements, rather than engagements between humans and non-human entities, such that *anthro-social cognition* is a more satisfactory title. Within *anthro-social cognition*, my second claim is that phenomenology should form the bedrock for social cognitive investigations. This is because (i). phenomenology allows for a fundamentally subjective approach to social cognition, without confining subjects to the kind of individualistic and de-contextualised characterisation that they are given by Cartesian cognitive science, and (ii). phenomenology appreciates the uniqueness of humans’ experiential nature and thus corroborates the chapter’s first claim. Developing this latter point, my third claim is that the notion of human subjects as enworlded, context-sensitive and constitutively social should be carried through to all aspects of social cognition, including scientific studies. On its most extreme interpretation, this would involve revolutionising the manner in which scientific studies of human social engagements are conducted. Yet even on a more reserved interpretation, it involves a significant shift towards greater interdisciplinarity, such that subjects are not always ‘blank’ experimental participants, but are appreciated in virtue of gender, age, ethnicity, cultural belonging and physicality.
- In the final chapter, I approach the matter of human selfhood head-on, placing my analysis as a counter-position to Kyselo’s (2014) recent enactivist account of selfhood.

The reason for engaging with Kyselo's work is that I agree with her premise that there is a need to remedy the problem of how bodily and social processes relate to one another within the individuation of the human self. After highlighting flaws in Kyselo's work, I propose that humans occupy a unique experiential domain of rationality due to the innate communicatory potential of our bodies. As such, our bodies should be conceived as 'linguistic' and as generating meaningful norms in interaction with others from the earliest moments of life. In other words, our bodies are never 'merely physical', but are always socially laden and socially expressive. The conclusion of this claim is that human selfhood is 'biosocial', incumbent on our ongoing modulation of bodily generated social norms and socially generated bodily norms.

The final result will be that we will have a full-blooded picture of human sociality, both at the level of the individual subject and at the level of individual subjects engaging with one another. As the brief chapter summaries hopefully indicate, the approach that I am taking in this thesis is both multi- and inter-disciplinary. In particular, I demonstrate how phenomenological and cognitive-scientific insights can complement one another. To some, this may seem controversial, in that phenomenology and science are often considered incompatible. Indeed, it would be wrong to completely ignore the fact that there is ongoing debate regarding this issue (see Kiverstein and Wheeler (2012) and Petitot, Varela, Pachoud and Roy (1999) for a variety of views, along with Gallagher and Zahavi (2008), Ratcliffe (2003, 2006, 2013) and Wheeler (2005, 2013)). However, few would deny that phenomenology and cognitive science can be mutually illuminating, even if one cannot have a strict 'science of phenomenology'. As far as a broad overview of this thesis is concerned, the specific mutual illumination that is provided can be put in the following terms: on the one hand, the scientific field of social cognition generally approaches sociality from Cartesian underpinnings and thus cries out for a phenomenological (i.e. 'Heideggerian') reformulation; on the other hand, existing phenomenologically sympathetic approaches to cognitive science (e.g. Dreyfus, 2007; Gallagher, 2005; Gallagher & Zahavi, 2008; Thompson, 2007; Varela, Thompson and Rosch, 1991; Wheeler and Di Paolo, 2011; Wheeler, 2005, 2013) tend to under-appreciate the social dimensions of phenomenology. My combination of phenomenology and cognitive science should therefore benefit both, producing cognitive science that is properly sensitive to the human subject and phenomenology that is robustly ensocialled.

Aside from these two fields, I also draw on the likes of cultural anthropology, evolutionary biology and gender studies in grounding and expounding my arguments. In some sense, I thus hope that the methodology of this thesis is an instance of practising what I preach: creating a body of work that depends on others, just as every human self does.

**Chapter 1 – Heideggerian**  
**Cognitive Science and the**  
**Nature of Everyday**  
**Existence**



## **Heideggerian Cognitive Science and the Nature of Everyday Existence**

### **1. Introduction**

There is such diversity in theories regarding the human mind and its fundamental role in human existence that it is a near-impossible task to provide a comprehensive overview of its many conceptions across many academic circles. Indeed, I will make no attempt to provide such an overview. Instead, this opening chapter will concern itself with laying on the table certain aspects of the mind and its role in everyday existence that are foundationally relevant to claims that I will make later in the thesis. As my intent here is merely to complete some groundwork from which later claims can be better understood, the ideas that are covered in this chapter will mostly be mere snippets of deeper philosophical issues. Whilst this may initially seem careless, it will hopefully become apparent that I am simply selecting constructive details that are appropriate to the architecture of my overall cause.

The foundational aspects of mind and existence that I wish to briefly outline will be split across four sections: firstly, I will discuss the recent turn from the orthodoxy of ‘Cartesian’ cognitive science to ‘Heideggerian’ cognitive science. Section 3 will be the longest in this chapter as I discuss how a subject may encounter everyday features of the world, working from within the framework of Heideggerian cognitive science; this will include an in-depth elucidation of the unreflective, yet importantly *mindful* nature of our everyday coping with the world. Subsequently, I will elucidate the complementarity of a subject and her environment. Lastly, I will touch upon the idea that a subject implicitly engages with social norms during everyday dealings with the world.

Each of these sections is a preparatory step to answers that will be given later (in subsequent chapters) to the central questions of the thesis, namely: how is human selfhood constituted by social processes? And how should we conceive of human social cognition?

### **2. Cognitive Science: from Descartes to Heidegger**

Cognitive science is, in many ways, a patchwork discipline of more established and clearly defined fields of study. It is typically taken to embrace neuroscience, psychology, artificial intelligence, linguistics and philosophy, although this list is by no means definitive. In spite of this hotchpotch make-up, cognitive science is one of society’s most exciting and relevant academic subjects, perhaps partly because of its inherently interdisciplinary nature and wide-reaching scope. It is also a relatively ‘new’ discipline, having only properly emerged under the title of ‘cognitive science’ in the 1970s, and it thus propels itself forward on a wave of academic enthusiasm and innovation. However, like any nascent discipline, it has faced certain teething problems and, since the 1990s, there has been a persistent forking of opinion that represents if not a revolution, then at least a schism.

As it coalesced into a viable field of study, cognitive science became fundamentally entrenched in the psychological notion of *cognitivism*, which states that mental processes are the result of computational mechanisms acting on neural representations, with these processes being abstractable as a system from our bodily and world-embedded behaviour. This definition yields two key insights. Firstly, cognitivism rests on the complementary notions of *representationalism*, which claims that cognising systems develop, store and process representations of worldly information, and *computationalism*, which claims that these representations are processed computationally (that is, through state changes that are detailed in a formal algorithmic specification such that “the causal structure of the system mirrors the formal structure of the computation” (Chalmers, 1994, p. 323)). Secondly, cognitivism offers an account of cognition as ‘de-worlded’. The notion of ‘de-worlded’ cognition will be explained shortly, but the rough idea is that cognitive processes are limited to neural machinery that is explanatorily independent – in the sense that explanatory value lies uniquely with the computational processing of neural representations – from the rest of the contextualising world. This second insight is particularly important, as it is this notion of a ‘de-worlded’ cognising subject that this chapter (and several later chapters) seeks to overcome.

Whilst there were various non-cognitivist approaches to cognitive science,<sup>[2]</sup> the two notions of representationalism and computationalism came to form the orthodoxy within the field (see Clark (2001) and Margolis, Samuels and Stich (2012) for comprehensive introductions to cognitive science). Due, at least in part, to the fact that such orthodox cognitive science endorses the computational manipulation of representational states, it is sometimes referred to as *Cartesian cognitive science*, drawing on the common conception of Descartes (1641/1996, 1644/1983) as the father of the view that the mind is a unified system of representations (see Brook (2007, pp. 5–8) for more on Descartes’ representationalism, Haugeland (1985) on an interpretation of Descartes as a computationalist and Wheeler (1995, 1996, 2005) on the commitment of orthodox cognitive science to a Cartesian account of mind). Whilst earlier philosophers may have extolled a similar view of the mind as a representational system – for example, one could ascribe such a view to Plato and/or Aristotle (Rakova, 2006, p. 161) – it is primarily due to Descartes’ enduring notion of the mind as an independent ontological domain that traditional cognitive scientists endorse a ‘sandwich model of cognition’, in which cognition is a process that is sandwiched between perceptual inputs and action outputs (Hurley, 1998). In parallel with the ontological independence of the internal mind that Descartes posited, orthodox cognitive science attempts to retain the explanatory independence of the internal (i.e. neural) mind.

---

<sup>[2]</sup> Connectionism would be a noteworthy theory here (see Bechtel and Abrahamsen (1991) and Clark and Lutz (1992)).

In recent years, however, a ‘Heideggerian’ opposition to this traditional Cartesian cognitive science has emerged. On a narrow interpretation, ‘Heideggerian cognitive science’ refers to those cognitive scientific accounts that are directly motivated by Heidegger’s thoughts on human existence (see Wheeler (2005), Dreyfus (2007) and Kiverstein and Wheeler (2012)). I am employing the term more broadly than this, however, as loosely synonymous with ‘anti-Cartesian cognitive science’, ‘phenomenological cognitive science’ or ‘4EA cognitive science’. Ultimately, any of these terms could be seen as problematic, but I favour the term ‘Heideggerian’ for two reasons. Firstly, this initial chapter draws on some specific Heideggerian notions; notably, our enworlded existence, our seamless engagement with worldly entities as being *for* some specific task and our pervasive social constitution. Secondly, the specific application of Heideggerian ideas to cognitive science provides fertile ground for tackling an overarching problem that this thesis seeks to address (see Introduction). That is, whereas cognitive science that is rooted in Cartesianism presents a wrongly ‘de-worlded’ framework for social cognition, the positive accounts of more phenomenologically oriented cognitive science often overlook the fundamentally social nature of being human (Dreyfus (2007a), Wheeler (2005) and the collection put together by Kiverstein and Wheeler (2012) – with the exception of Gallagher and Sete-Jacobsen (2012) and Talero (2012) – could all be accused of such an error, in spite of addressing the importance of social human nature during critical accounts of Cartesian cognitive science). As socialisation is a fundamental ingredient of Heidegger’s (1927/1962) notion of ‘Being’, a properly Heideggerian framework should be one in which subjects’ worldly embeddedness is inherently social, which is an idea that I will endorse later in the thesis (chapters 5 and 7 in particular). More than the other existing labels for cognitive science, ‘Heideggerian’ thus serves to emphasise the social imbuelement of cognitive science (and existence more generally) that I believe is vitally important.

Having said this, it is important to stress that my appropriation of Heidegger’s name is not an endorsement of his philosophical positions unless explicitly stated as such, nor is it to suggest that all of my subsequent discussions and claims should be read through a Heideggerian lens. Indeed, throughout later chapters, the phenomenological hues of cognitive science will not focus solely on Heidegger, but will be expounded through references to Merleau-Ponty, Sartre, Henry and Jonas, amongst others. I am thus using the term ‘Heideggerian cognitive science’ at present as the best available for capturing the plethora of anti-Cartesian cognitive theories that have recently emerged in a manner that is suggestive of key issues – such as worldly embeddedness and socialisation – that are central to this thesis.

Terminology aside, the recent turn away from orthodox ‘Cartesian cognitive science’ has replaced the idea of mind as a kind of central processing unit that sequentially receives

inputs, computationally processes them and delivers causally formalised outputs with the idea of mind as a dynamic phenomenon that depends non-trivially – sometimes constitutively – on its surrounding world (including the body and an environmental niche). In other words, Heideggerian cognitive science strives to deliver a more satisfactory alignment between cognition and our existential nature as living bodily beings who are embedded in specific socio-cultural worlds. As such, it encompasses various views of the mind as embodied, embedded, extended, enactive and affective (Kiverstein, 2012). The differences and compatibilities between these theoretical views have been hotly debated (see Clark (2007), Kiverstein and Clark (2009), Menary (2007), Rowlands (2010), Rupert (2009), Thompson and Stapleton (2009), Ward and Stapleton (2012), and Wheeler (2011c) for a range of relevant views), but the common thread that loosely stitches them together is nicely captured by Haugeland (1998):

If we are to understand mind as the locus of intelligence, we cannot follow Descartes in regarding it as separable in principle from the body and the world[...] Broader approaches, freed of that prejudicial commitment, can look again at perception and action, at skillful involvement with public equipment and social organization, and see not principled separation but all sorts of close coupling and functional unity[...] Mind, therefore, is not incidentally but intimately embodied and intimately embedded in its world (pp. 236–237).

Of course, the extent of “close coupling”, “functional unity”, ‘intimate embodiment’ and ‘intimate embeddedness’ is where one can find scope for a plethora of divergent opinions. Nonetheless, there is a general anti-Cartesian commitment to properly situating cognition as an achievement that belongs, in part, to the ‘external’ natural world that justifies placing many views together under an umbrella phrase.

Perhaps more than anyone else, the truly Heideggerian nature of this gradual revolution in cognitive science can be traced to the critical work of Hubert Dreyfus (1991, 1992, 2001, 2007a, 2007b, 2007c, 2012; Dreyfus & Kelly, 2007). Dreyfus initially came to the fore through his attack on the orthodox approach to artificial intelligence (1991, 1992, 2007a), but he certainly seems to view his critique as applying to orthodox cognitive science more generally. Drawing on Heidegger’s (1927/1962) phenomenological notion of *being-in-the-world*,<sup>[3]</sup> Dreyfus contends that systems which aim to replicate real-world cognition should not be programmed to computationally respond to stimuli, but should instead mirror the organismic capacity for practical know-how that is based in historical familiarity with context-dependent meaning. Whereas a computational system needs to use formalised

---

<sup>[3]</sup> For Heidegger, the primary human mode of Being is *Dasein* (‘there-being’), which is characterised by (i). the capability to inquire into the nature of Being through engagement with the phenomenal world: “Dasein...is ontically distinguished by the fact that, in its very Being...Being is an *issue* for it” (Heidegger, 1927/1962, p. 32), and (ii). the fact that humans are always ‘in-the-world’, engaged with worldly equipment and existential structures not as withdrawn ‘objects’, but as potentially ‘transparent’ entities which contribute meaningfully to the mode of Being. As such, at any point of analysis, Dasein is always already embedded in a world that is inherently meaningful, which results in every situation offering a range of possibilities that will matter to Dasein to varying degrees.

representations to firstly construct a contextual model and then select an appropriate response from a gargantuan ‘playbook’ of propositional rules, a biological cognising system already has an evolutionarily and developmentally refined set of skills that allow for appropriate choices to be made in open-ended, everyday situations (Kiverstein, 2012, pp. 2-5). The problem for computational systems, which employ context-independent representations, is that they fall into a regress in which they are perennially searching for further contextual rules to ‘frame’ their current situation, along with subsequent appropriate ‘frames’ as the situation progresses (this is appropriately called the *frame problem*; see Dennett (1978), Fodor (1983), Shanahan (2016) and Wheeler (2005)). Systems in the ilk of humans, on the other hand, need not reconstruct contextual significance from (context-independent) formal rules and states because we employ non-propositional, practical know-how that is developed over time. As Dreyfus (1992) puts it:

When a person has enough experience to make him or her an expert in any domain, the field of experience becomes structured so that one directly experiences which events and things are relevant and how they are relevant.[...] Objects appear to an involved participant not in isolation and with context-free properties but as things that solicit responses by their significance.[...] In general, human beings who have vast experience in the natural and social world have a direct sense of how things are done and what to expect. (p. xxviii/xxix)

With this emphatic statement we find that the Cartesian need for formal computation is jettisoned, along with the input-output processing of ‘box and arrow’ cognitivist psychology and the positing of an immense contextual database of propositional knowledge. Instead, we have pervasive contextual understanding of relevance that draws on our embodied and embedded nature and endows us with a “direct sense” of how to act in open-ended situations.

This notion of having a direct sense of how to act once we are experienced in the world seems to match nicely with our intuitions regarding our experience of the world. Unlike the ostensibly laboured sequencing of computational processes, the idea of responding in an almost instinctive manner aligns more comfortably with the smooth and unreflective way in which we tend to act out our everyday lives. Indeed, it is perhaps more clear now why Heidegger is an appropriate exponent of this turn away from Cartesian cognitive science. Along with the other existentialists of the 19<sup>th</sup> and 20<sup>th</sup> centuries, Heidegger extolled a view of humans (who are, of course, our canons of cognitive systems) as beings who are not disembodied manipulators of context-independent representations, but are contextually embedded and dynamic enactors of a meaningful world.

It is in this existential subsoil that the subsequent topics of this thesis – both scientific and phenomenological – will be rooted.

### 3. Encountering Everyday Entities and Coping with the World

Within Heideggerian cognitive science (as we will see in more detail later on), a cognising subject is always meaningfully *in-the-world* and thus acts in accordance with one's historically sedimented repertoire of skills. Such skills typically result in smooth and unreflective behaviour, rather than tackling everyday tasks as complex problems that need to be solved. One way of looking at this smooth and unreflective activity is through the existentialist idea that we encounter entities as being *for* something. For example, Heidegger (1976/2010) claims:

What is first of all 'given' [...] is the 'for writing', the 'for going in and out', the 'for illuminating', the 'for sitting'. That is, writing, going-in-and-out, sitting, and the like are what we are a priori involved with. What we know when we 'know our way around' and what we learn are these 'for-what's'" (p. 144).

So as we cope smoothly and skilfully with our environment, we do not engage with physical entities as complex puzzles to be solved, but rather respond to each entity's 'for-what' (Dreyfus, 2007a). When I need to write something down, for instance, I do not need to analyse the scene before me, deciphering the functional purposes of the entities surrounding me and concluding that the plastic, cylindrical, ink-producing object is best suited to the task. Instead, I see a pen and move to pick it up without any complex analysis; I can enter into the act of writing without ever seemingly 'thinking' about retrieving the pen, or moving my hand in a certain way to make my scribbling legible. The pen is simply encountered as offering the possibility of writing – of being 'for writing', just as a cup of tea is 'for drinking', a football is 'for kicking' and a book is 'for reading'.

In more recent literature, this notion of skilfully encountering entities as being *for* something has been developed via Gibson's (1979) ecological notion of *affordances*. For example, Rietveld (2008a, 2008b, 2012a, 2012b, 2013; Rietveld and Kiverstein, 2014), who is largely motivated by Merleau-Ponty's (1945/2012) notion of 'motor intentionality', describes how we have "a form of unreflective *embodied* intelligence" (Rietveld, 2012a, p. 207) that is captured by our responsiveness to affordances. As we shall see shortly, Gibson is concerned primarily with perceptual engagement with the world, but through the likes of Rietveld (see also Dreyfus (2007a, 2007b, 2007c; Dreyfus & Kelly, 2007) and Gallagher (2005)) his notion of affordances has come to form a bridge between traditional phenomenology and modern, phenomenologically oriented cognitive science.<sup>[4]</sup> Through this modern application,

---

<sup>[4]</sup> There are of course nuanced differences between Gibson's (1979) perceptual account of affordances and the more recent phenomenological engagements with affordances. However, as Rietveld generally does, it is possible to conceive of affordances broadly as "an organism's possibilities for action provided by the environment" (2012b, p. 109). It is this broad interpretation that I am primarily concerned with (see Rietveld & Kiverstein (2014) for a conceptual framework in which affordances are defined by agential abilities with specific contexts).

affordances (as we will see throughout this chapter) become a highly useful concept for considering how we engage with everyday worldly entities.

In many ways, Gibson's ecological theory of perception can be viewed as one side of a small skirmish within the wider argument between 'Heideggerian' and 'Cartesian' stances in cognitive science. For Gibson, there was a need to break away from the orthodox representation-processing theories of perception, which relied primarily on physical optics and proposed that vision's 'starting point' is a static retinal array (e.g. Gregory (1966, 1970)). In place of this, Gibson proposed that visual perception is *non-inferential*. In his novel model, there is no need for mediating involvement from a representation-building brain because the information received from the "*medium, substances, and the surfaces*[...of the...]*terrestrial environment*" (Gibson, 1979, p. 16) is sufficiently rich for an active perceiver to interact directly with the environment. In contrast with orthodox theories' dependence on a reconstructable physical reality, Gibson's approach employs an *ecological* framework, in which "*animal and environment* make an inseparable pair" (ibid., p. 8; this idea will be returned to in section 4). It also makes use of ecological optics – the "*ambient optic array*" (ibid., p. 51) of light which specifies "surfaces, their composition, texture, colour, and layout" (ibid., p. 52) – whereas the physical optics of traditional approaches suggest that radiant light can at most specify atomic information. The idea here is that the ambient light of ecological optics conveys information about "gross properties, not[...] atomic properties" (ibid.). With these gross properties, a perceiver receives the necessary structural information to specify the environment and act upon it without inference; the orthodox alternatives require the brain to infer environmental structure from the atomic picture received by the retina. What this amounts to is that fact that the structural information that an active perceiver receives from the ambient medium, substances, and surfaces will *afford* certain ways of acting:

The *affordances* of the environment are what it *offers* the animal, what it *provides* or *furnishes*[...]I mean by it something that refers to both the environment and the animal in a way that no existing term does. It implies the complementarity of the animal and the environment. (Gibson, 1979, p. 127)

In this short passage (and its preceding description), we find a renaissance of the existentialist notion of subjects' *being-in-the-world* (albeit in a different context and with different connotations). The perceiving animal is not isolated from its surroundings, but is embedded in a large, information-rich environment. The animal is thus endowed with a unique ecological position from which certain things matter to it depending on the niche in which it is found, and these relevant things will *afford* the animal certain ways to behave according to its current needs. Just as the Heideggerian notion of *in-the-world* subjects

provides opposition to Cartesian cognitivism, so, too, does the Gibsonian theory of ecological perception oppose orthodox theories of perception.

### 3.1. Mindless or Mindful? How do we Cope with the World?

An essential aspect of perceiving and acting in a non-cognitivist manner, which has already been touched upon and which nicely aligns with the notion of our responding to affordances, is that one is exploiting the informational depth of one's familiarity with the world (in which one is always already embedded), rather than having to analyse objective physical features of the world and then construct a representation of these features. This lack of deliberative analysis is self-evident in our everyday dealings with things; we do not struggle to decipher our surroundings, but behave, for the most part, as accomplished experts. A useful way of viewing this is through the world of sport, in which we find expert performers executing highly complex tasks with incredible speed, accuracy and contextual awareness. Professional athletes, for example, often talk of being 'in the zone', inhabiting a level of performance in which there is an "absence of irrelevant thought" (Smith, 2002, p. 24), and an effortless "kind of concentrated blankness" (Smith, 2004, p. 117). At these times, one seems able to execute complex and skilled actions with boundless ease, performing with a calmness, clarity and adroitness that exceeds that of teammates and opponents. This is as true of complex, rule-abundant sports such as football, tennis and cricket, as it is of simpler sports such as running. The decorated ultramarathoner Scott Jurek describes these instances of immersion 'in the moment':

Running backs speak of the game slowing down until all the other players are moving with almost cartoon sluggishness[...]. Basketball players testify that the hoop at which they're shooting not only seems larger but *is* larger. Runners speak of being absorbed into the universe[...] the sensation is one of effortlessness. (Jurek and Friedman, 2012, pp. 137/138)

Later, Jurek talks of entering "that place beyond thinking", in which one is "immersed in the present moment where nothing else matters" (ibid., p. 182). Such instances seem devoid of explicit thought, or awareness of anything beyond the present situation; there is simply the fluent flow of experience being played out by the perfected bodily movements of expert performers. We get an insight into this abandonment of thought and choice during post-match (or post-race) interviews, when players and athletes are entirely unable to describe how they played a crucial pass, how they held their nerve at the pivotal moment, how they beat two men with a feint of the shoulder, or how they executed the perfect tactical race. Rather than the vivid and in-depth analysis that we crave as spectators and fans, we are often simply presented with a response of "Um, I'm not sure *how*, I just kind of *did it*". Importantly, this being 'in the zone' is not limited to sporting professionals or moments of

personal transcendence in a sporting arena. We can also consider our phenomenological understanding of quotidian acts from the perspective of being ‘in the zone’. For instance, we readily accept and shake, without hesitation, the hand of a new acquaintance; we are able to dress in the morning without deliberating about how to put on our clothes; and I can eat my dinner without considering it a cognitively taxing task. In all of these cases, the common factor is the analytical effortlessness – the lack of deliberation or reflection – that one experiences.

Now, at this point, it is necessary to tread carefully over some philosophical eggshells. In the above exposition, I have elucidated the idea that human subjects – who are always already meaningfully in-the-world at any given time – typically engage with the world in such a way that everyday behaviour is dominated by “episodes where the activities of a skilful individual unfold adequately without reflection on his or her part” (Rietveld, 2010, p. 203). For Dreyfus, these acts tend to be “neither caused like a reflex, nor done for a reason” (Dreyfus, 2007a, p. 1143). Indeed, as far as Dreyfus is concerned, when we are fully absorbed in an action and responding to affordances so as to perform in a skilled and unreflective manner, we find that:

There is no place in the phenomenology of fully absorbed coping for mindfulness. In flow,[...] there are only attractive and repulsive forces drawing appropriate activity out of an active body. (Dreyfus, 2007c, p. 374)

Whilst this claim accords with the lack of deliberation and an apparent abandonment of real-time reflection which typify the sporting and everyday examples mentioned earlier, it perhaps takes things a step too far in its complete rejection of “mindfulness”. Indeed, Dreyfus (2007b) goes as far as to say that “mindedness is the enemy of embodied coping” (p. 353). On an initial reading, this may seem completely justified and, taking into account the discussion thus far, a fairly straightforward claim. As we have already made apparent, athletes will seemingly act with a kind of practised impulsiveness rather than with carefully calculated movements, and the (affordance-invoking) dressing, walking, writing, reading, eating and drinking of everyday life seem to be a far cry from cognitively measured behaviours. As Dreyfus (2012) puts it, “if you try to reflect on the intelligibility of the situation, that is, if you try to *think about why* things are going so well[...]you will at best perform competently” (p. 2, italics added). The worst thing a person can do when coping smoothly with the world, is try to make determinate the manifestation and maintenance of this smoothness (ibid.). There simply seems to be no room for mindfulness in skilled and unreflective activity.

However, the worry is whether the accepted concept of skilled, non-deliberative and unreflective action can be satisfactorily assimilated into a description of ‘mindless’ activity –

activity in which there is *no operative self-awareness or subjective monitoring of a situation's particularities* (Dreyfus, 2007c, p. 373). If we consider simple motor actions in isolation, then it is easy to sympathise with the idea that there is a 'mindlessness' to our performances. It seems highly plausible that the positing of cognitive activity when taking a sip of water, flicking a light switch or walking across a room is effectively obsolete; we find in such cases that the execution of the act is entirely given over to bodily mechanisms (Merleau-Ponty, 1945/2012, p. 55). However, once we turn to more complex acts, along with more temporally extended activities, it seems misleading to propose an absence of mindfulness. To illustrate this claim, let's reassess the actions of sporting professionals 'in the zone'. These sportsmen and women are able to execute highly complex acts to an exceptional standard with great ease, assuredness and clarity. But does such ease and effortlessness require a lack of mindful activity? Take, for example, tennis players. Tennis players spend the vast majority of their youth and a large portion of their adult lives hitting tennis balls. The techniques involved become 'second nature' so that a top-spin forehand, sliced backhand, drive volley or smash can be hit with minimal difficulty. On Dreyfus's view, these dynamic actions have become ingrained in the players' experiential repertoire, so that the analytical, reflective brain can simply defer responsibility to practised bodily movements. Sheets-Johnstone (1999) refers to these experientially inbuilt dynamic movements as 'kinaesthetic memory': "a kinetic dynamics unfolds that is at once both familiar and yet quintessentially tailored kinetically to the particular situation at hand" (p. 21). The problem Dreyfus faces is that such fluent dynamic movements do indeed seem to be "tailored" to specific situations. It may be correct to say that kinaesthetic memory drives our frequently practised actions (e.g. hitting a top-spin forehand) at the expense of self-referential awareness, deliberation and reflection, but how does this kinaesthetic memory account for the *tailoring* to specific situations? That is, how can kinaesthetic memory account for the great diversity of the *same action* (hitting top-spin forehands, for example)? No two top-spin forehands are exactly alike, and so how can a player mindlessly *adapt* to the continually changing ebb and flow of a tennis match? The bodily activity of absorbed coping cannot be solely responsible for the fine-tuned adjustments and trickery involved in competitive sport. A champion tennis player will, generally speaking, be capable of hitting forehands with (apparent) consummate ease, but hitting *this* forehand in *this* specific context will no doubt require some cognitive attention to the particularities of the situation. Having a "a direct sense of how things are done" (Dreyfus, 1992, p. xxix) will only take us so far. Once we realise that the majority of actions – and, indeed, varied instances of the same generic action – are context-dependently unique, we must acknowledge that there must be at least *some* operative mindfulness in the kinds of activities that are as complex as those found in sport. As Sutton, McIlwain, Christensen and Geeves (2011) explain:

[Expert performers] like to entrust grooved action sequences to the body, to the habitual routines of kinaesthetic memory. But[...] they also want to be able to bring all of their experience to bear in the moment, to bring memory and movement together, with thought and action cooperating instead of competing. (p. 80)

Even the most well-practised and undemanding actions, such as a professional tennis player's top-spin forehand, will require the ongoing tracking and accommodating of the context's "unique constellation" of salient features (ibid.). Of course, this is not to say that these actions involve reflective cognition, self-referential awareness or "intellectual instructions to the body" (ibid., p. 78), but there is an operative awareness of some sort – a kind of mindful monitoring of the situation that maintains the absorbed, flexible nature of the action. As Fridland (forthcoming) argues, our skilled performances can be flexible and versatile precisely because of the *minded* nature of our expertise. Whilst embodied motor routines undoubtedly play a part in our expertise, Fridland suggests that we also employ a kind of operative cognitive control by parsing these motor routines into smaller segments and thereby maintaining the potential for flexibility and variation. Through this expert segmentation of our automatized bodily routines, we can lithely adapt our activity to the unique needs of present performance, meaning that "skilled actions...are minded almost all the way down" (ibid., p. 1). Contrary to Dreyfus's view, skilled and unreflective action seemingly can be mindful; not a kind of "heavy-duty computationalism", but a "dynamical, fleet-footed, improvisatory, collaborative" kind of mindfulness (Sutton et al., 2011, p. 78).

In order to reinforce this idea of skilled and unreflective *mindful* action, let's consider the 'everyday' act of hammering a nail into a block of wood (cf. Heidegger, 1927/1962). As a fairly simple action, hammering is something that most of us who have done even a little DIY would consider ourselves experts at. At the very least, after one has been hammering for some time – when putting up a block of shelves, for example – one would generally seem to be properly absorbed in the activity. Hitting a nail two, three, four times into a block of wood before moving to the next nail can be a seemingly fluent and cognitively effortless task. If, however, as one lands an initial hammer-blow, a nail becomes slightly skewed from its ideally perpendicular position, then an adjustment is required. Rather than needing to analyse the situation and calculate the appropriate angle at which to hit the nail in order to re-align it, one can simply make an on-the-fly *adaptation* to one's movement. For Dreyfus, this adaptation would be explained by the subject's 'direct sense' of how the nail should be hammered, but, as we have already noted, this sense seems to require a mindful monitoring of the situation – not a specific monitoring of oneself as a subject, but a monitoring *of the action*. Any adjustments will therefore arise neither from propositional beliefs, nor mindless bodily-governed intuitions, but rather from a mindful form of absorbed coping which is open

to both environmental fluctuations (such as the nail's skewedness) and cognitive "nudges and triggers" (Wheeler, 2005, p. 229). Thus, the smooth engagement in the act of hammering is not disrupted by the nail's off-centre position because an absorbed, yet mindful, adaptation to the kinetic dynamics of the situation can be made. It is crucial that this mindfulness is not considered a reflective or deliberative kind of cognitive activity, but it is still "in the realm of the psychological" (Sutton et al., 2011, p. 78). The key to acting in a skilled and unreflective manner thus does not have to be cutting off engagement with intellect and awareness, but achieving an effortless "*flexibility* in linking thought and action, knowledge and motion" (Sutton, 2007, p. 779). Experience and expertise allow us to effortlessly, but mindfully, adapt to the fluctuating world. As Sutton et al. (2011) put it, "[g]enuinely thinking on one's feet is still a form of thinking" (p. 79).

Dreyfus would no doubt argue that the maintenance of one's optimal grip on a situation, even when achieved via dynamic adaptations, is still a mindless activity. That is, on Dreyfus's view, any awareness is achieved solely through "a steady flow of skillful activity in response to one's sense of the situation", such that "when one's situation deviates from some optimal body-environment relationship, one's activity takes one closer to that optimum and thereby relieves the 'tension' of the deviation" (Dreyfus, 2002, p. 378). For Dreyfus, then, any dynamic adaptations arise from *circumspective awareness* – a bodily know-how and subjective 'sense' of how a situation should unfold. This circumspective awareness is meant to cover any kind of operative awareness, and leaves no room for mindful input.

In response to this, we can approach the idea of mindful, skilled and unreflective action from another angle. In our everyday lives, a significant proportion of time is taken up by the activities of paid employment. Teachers teach, lawyers practise law, painters paint, and chefs cook. Each job brings with it a (partly unspoken) set of norms which determine appropriate behaviour, as well as incorporating oft-encountered sets of affordances. This results in subjects behaving in a domain-specific manner. What I want to argue is that this domain-specific behaviour will involve not just operative awareness, but also a configurative kind of mindfulness that is still not self-referential or reflective. The following passage from Sartre (1943/1984), describing a waiter at work, will help to illustrate this point:

Let us consider this waiter in the café. His movement is quick and forward, a little too precise, a little too rapid. He comes towards the patrons with a step a little too quick. He bends forward a little too eagerly; his voice, his eyes express an interest a little too solicitous for the order of the customer[...] He applies himself to chaining his movements as if they were mechanisms, the one regulating the other[...] He is playing at *being* a waiter in a café. (p. 59)

Sartre, here, is in the process of making a complex existential claim regarding 'bad faith'; that is, he is claiming that the waiter is suppressing his freedom by loading his behaviour

with waiter-esque mannerisms (i.e. ‘a little *too* precise...a little *too* eagerly’ etc.). However, the complexities of bad faith need not concern us. Instead, there are two aspects of the passage that are highly relevant to our current discussion: (i). the waiter is behaving in a very specific manner; he is performing the specialised actions *of a waiter* – of a domain expert coping appropriately with his work environment, (ii). these expert actions are completed “as if they were mechanisms”, that is, in a manner which we can safely assume is skilled and unreflective. What is interesting about this is that the specific behaviour of the waiter – let’s call him Mr. X – will be notably different from his everyday behaviour when he is not *being a waiter*. When Mr. X is at home watching TV, reading a book, or listening to the radio, his actions are completely unremarkable, in the sense that we could not single out Mr. X’s actions (broadly construed) as being ostensibly different from any other person’s. But once Mr. X is at work in the café, he becomes definable by his actions – we can single him out *as a waiter*. The fluency and expert skill with which Mr. X performs his duties as a waiter rule out the possibility that he is explicitly controlling his actions in order to behave in a ‘waiter-appropriate’ manner. However, it seems equally unlikely that his being a waiter is a mindless act. Dreyfus would most likely argue that the waiter’s domain expertise and the affordances provided by the context of the café result in a bringing forth of actions that are appropriate for a waiter: Mr. X’s expert waiter-behaviour is simply solicited by his environmental surroundings. Unfortunately, this argument is only partially satisfactory. I say this because having “a direct sense of how things are done” seems insufficient to explain Mr.X’s specific behaviour. Rather, his specific behaviour involves a suppression of certain traits and movements, an exaggeration of others, and continual adaptations to the charm, deference, tranquillity or efficiency required by various patrons, colleagues and managers. Mr. X is not simply ‘going with the flow’ as he works in the café, and letting “attractive and repulsive forces” draw out appropriate behaviour. He is in fact establishing his *own* benchmarks and working against what may be an attractive option (and towards more repulsive ones) in order to behave in a waiter-appropriate way. He is adopting a *specific attitude* which acts as a subsoil for his skilled and unreflective actions. What’s more, Mr. X may be capable of adopting this specific attitude – this waiter-expertise – to a variety of scenarios. When, for example, Mr. X is entertaining friends in his home, he may draw on his expert serving skills “as if they were mechanisms” (Sartre, 1943/1984, p. 59), just as he does when at work. Further afield, Mr. X may make use of the politeness and deference that his employment requires in a variety of daily tasks – giving directions to a stranger, talking to shop clerks, or paying bills over the phone. In each of these instances, Mr. X is not *being a waiter* in either the Sartrean or commonsense understanding of the phrase, but he is making use of aspects of his waiter-expertise in different domains, with the same effortless that is typical of his work in the café. That is, Mr. X is *configuring* the situation so as to be capable of behaving in

a domain-specific way (see Cappuccio and Wheeler (2012) for more on this notion of contextual configuration). Without positing some pre-figurative mindful input, it seems that a lot is being asked of Dreyfus's 'sense' of a situation in order to explain this. More than merely being a kind of comportment that is brought forth by environmental surroundings, such a specific attitude can thus be *mindfully* adopted not through an explicit choice or propositional state, but through Mr. X's expert understanding of when the attitude is appropriate.

In a similar vein to the 'waiter-attitude' that is variably adopted by Mr. X, Hochschild (1983) studied how flight attendants are trained in a kind of 'emotional dissonance' in which they "induce or suppress feeling in order to sustain the outward countenance that produces the proper states of mind in others" (p. 7). In other words, the flight attendant adopts a specific attitude of "gracious manners and warm personal service" (ibid., p. 93) as a way of tailoring her behaviour to the situation of her employment, even if it requires stifling certain emotional reactions and exaggerating others. As an expert in the domain of airline service, we can assume that the flight attendant carries out her tasks in a skilled and unreflective manner, but whilst doing so she is still *mindfully* behaving in accordance with her specific work-appropriate attitude. She is not continually making explicit decisions to act in a certain way, and she is still fluently responding to affordances, but the situation has been mindfully configured (and is mindfully adjusted in different circumstances) so as to bring forth situationally appropriate responses. This, along with the behaviour of Mr. X, can in fact be conceived of as an extension of the adaptive mindful dynamics we have discussed thus far. That is, in much the same way that a tennis player adapts to each unique shot and a carpenter adapts to the subtleties of hammering, so too can a waiter or flight attendant flexibly adapt an expertly practised psychological attitude to appropriate situations. This capacity to configure one's situation by flexibly adapting a specific attitude can be thought of as a toolkit for skilfully coping with the world that can be appropriately realised in relevant situations. Subjects are then 'at home' in a familiar context and able to unreflectively act in an appropriate manner, say, *as a waiter* or *as a flight attendant*. The mindfulness in question is now more than just a (non-self-referential) operative awareness regarding the particularities of a specific action (such as hitting a top-spin forehand). It also encompasses a broader operative awareness which amounts to the configuration of a situation into a 'familiar world', so that affordances bring forth a skilled and unreflective response, in accordance with a mindfully maintained psychological attitude.

Initially, then, we find that the mindfulness that is at stake here is in fact highly compatible with much of Dreyfus's position. It does not reject outright the potential for actions in which the mind is obsolete, such as grasping a cup, having a drink, or walking across a room. Nor does it diminish in any way the presence of those more complex instances in which

“everything becomes easier, confidence rises, time slows down, and the mind, which usually monitors performance, is quieted” (Dreyfus, 2007c, p. 373). The point of divergence comes from the fact that Dreyfus implies that ‘monitoring’ during some activity means ‘monitoring oneself as a subject’, and therefore having a reflective and self-referential kind of awareness. What’s more, for Dreyfus, any awareness *of the situation* is ‘mindless’. But neither of these claims need be the case. Firstly, one can monitor performance through a *mindful* operative awareness of the situation, not as a self-referential subject, but as a constituent of an unfolding event with an unreflective, mindful capacity to tailor appropriate actions to the situation. This mindfulness comes in the form of adaptive inputs to skilled and unreflective actions, and promotes a significant divergence from the Dreyfusian picture of mindless embodied coping. Secondly, one can also expertly configure one’s situation through the mindful maintenance of an operative attitude, which results in unreflective behaviour of a domain-specific kind. This mindfulness can be necessary in a pre-figurative sense to enable skilled and unreflective activity of a specific kind to take place. Although closely intertwined and not conflicting, it is notable that these two forms of mindfulness may have different consequences, differing as they do in terms of when they influence an active process (real-time and pre-figurative), and in terms of the extent of disagreement with Dreyfus’s approach to skilled activity (there may be room for marrying the configurative kind of mindfulness with Dreyfus’s framework, but there is direct conflict between ongoing mindful awareness and a mindless ‘embodied sense’ of how things should be). In both instances, aspects of Dreyfus’s work can be highly illuminating, but without any mindful input his reliance on a mindless ‘sense’ of a situation’s optimal balance and what is appropriate ultimately produces a phenomenological framework that distorts the explanatory story. It thus seems that the mind can be quieted, but not silenced entirely. It is still needed to contribute to the adaptations which are inherent in skilled and unreflective activity.

We thus have a more detailed picture of how *in-the-world* subjects engage with their everyday surroundings. This involves an emphatic rejection of Cartesian cognitivism in favour of historically sedimented skills of embodied and embedded subjects, but it does not evict the mind entirely. There is no longer a reliance on the computational processing of representations as per the orthodox Cartesian model, yet researchers chasing human-like cognition should not dash too rapidly to the domain of implicit ‘sense’. Contextual understanding still requires mindful inputs and the outstanding question, which will be addressed in later chapters, is how these inputs (and the context itself) are both grounded and processed.

#### **4. The Complementarity of Subject and Environment**

In the previous section, we have seen how in-the-world subjects, who help to form the basis for Heideggerian cognitive science, engage with everyday entities as being *for* something, or, in Gibsonian (1979) terms, as *affording* certain possibilities for action. A key feature of this structural aspect of everyday existence is the “complementarity of the animal and the environment” (ibid., p. 127). In this section, I will look at the notion of complementarity a little more closely, as it provides an opportunity for considering our *in-the-world* nature in more detail and thus sheds further light on the motivation for shunning a Cartesian picture of cognition.

As noted already, it is essential to Heidegger’s (1927/1962) phenomenological account of human existence that we are always ‘in-the-world’ and frequently engaging with worldly entities and other people not as withdrawn objects, but as ‘transparent’ equipment which contributes to our meaningful mode of existence.<sup>[5]</sup> On this view, we are not faced with (nor are we) physically delineated objects and subjects, but instead inhabit a phenomenal world in which, for the most part, we unreflectively act in such a way that “we pour ourselves out into [external objects] and assimilate them as parts of our existence” (Polanyi, 1964, p. 54). Thus, when hammering, we move into a “primordial” relationship with the hammer in which it is encountered “unveiledly” as a *transparent* entity (Heidegger, 1927/1962, p. 98). The hammer is not an isolated entity with context-independent properties, but a constituent of an ongoing activity. This specific kind of ‘in-the-worldness’, which is characterised by the existential transparency that has been described, also encapsulates our being *absorbed* in an activity. Developing Heidegger, Wheeler describes such absorption as follows:

[W]hile engaged in trouble-free hammering, the skilled carpenter has no conscious recognition of the hammer, the nails, or the work-bench, *in the way that one would if one simply stood back and thought about them*. Tools-in-use become phenomenologically transparent.[...] The carpenter becomes absorbed in his activity in such a way that he has no awareness of himself as a subject over and against a world of objects[. ...] [T]he awareness that is present[...] is non-subject-object in form. (Wheeler, 2011a)

This passage nicely reinforces the idea of being ‘in the zone’ and acting in an effortless manner that lacks explicit thought, deliberation or reflection. It also illustrates how our being *absorbed* in an activity is typical of our skilled and unreflective action: one loses self-referential awareness and the notion of a subject-object dichotomy. In such instances, the ongoing flow of experience is all that matters, and this is key to understanding how we fluently and consistently respond to affordances without needing to constantly re-analyse the world. Because we are always already in-the-world that matters to us, we can existentially

---

<sup>[5]</sup> Such ‘transparent’ engagement with the world typifies Heidegger’s (1927/1962) phenomenological category of the “ready-to-hand”. However, we can also engage with the “present-at-hand” and “unready-to-hand”, as well as simultaneously inhabiting multiple other phenomenological categories that will bear relevance to these modes of worldly engagement.

“accept” worldly entities by engaging with them in a transparent, non-subject-object manner (Polanyi, 1964, p. 59). This engagement with the world means that the orthodox theoretical picture of an isolated subject needing to analyse the objective world is simply the wrong way to consider the issue; with regards to those occasions in which one is absorbed in an activity, one is (with the potential for some mindful input) merely a subsystemic component in an unfolding worldly event. When viewed this way, it becomes easier to understand how environmental entities can so frequently and consistently *bring forth* appropriate responses from a subject. Due to the “complementarity” (p. 127) of subject and environment that Gibson (1979) mentions, we can now see that the expert’s bodily movements, their mindful adjustments, and worldly affordances are all dynamic processual constituents of an overarching mind-body-world system. As I touched upon in section 2, this perspective obviates the presence of a dualistic mind/world interface and emphasises our worldly embeddedness, such that we *react to* and *interact with* the environment, rather than analysing it and *acting on* it in a removed manner. In other words, discontinuity between a subject and her environment is existentially dismissible during absorbed activity (Haugeland, 1998). Thus, when an affordance brings forth a response, it is not phenomenologically detached from us, nor are we isolated subjects; the affordance becomes a constituent of our phenomenological experience – it is a *relation* between a subject and the environment rather than a context-independent objective offering. During absorbed behaviour we therefore find that the unfolding activity – viewed as a phenomenologically lived experience – is constituted by the organism (brain and body) and world, and the existentially transparent organism-world interface.

Of course, it is not the case that every engagement with worldly entities is a smooth and unreflective activity in which equipment is encountered as ‘unveiledly transparent’. Many everyday scenarios will require conscious focus and effortful reasoning. The point is that the complementarity of a subject and her environment results in a *vast number* of everyday activities being accomplished in a smooth and unreflective way (with such activities being prone to the kinds of mindful inputs that have been discussed earlier). Even as a subject is engaged in an overtly analytical task such as playing chess, she can still expertly have a sip of water, check her watch or adjust her clothing without having to explicitly reflect on the execution of these acts. Part of the reason that many everyday tasks are executed unreflectively is that a subject’s skilled bodily nature is always *fundamentally* pre-figured by overarching subjective concerns (Rietveld, 2008a). One can state this by saying that there is *purposiveness* to subjective behaviour (that is, behaviour is explicable and conceivable only as being causally grounded in accordance with purposes (Kant, 1790/2000, p. 220)). In virtue of having a specific purposive stance towards the world, subjects’ “responsiveness[...] is *concernful*, in the sense that it takes into account what matters to us” (Rietveld, 2012a, p.

212).<sup>[6]</sup> Thus, while worldly entities tend to afford specific responses in accordance with some socially defined role – for example, a pen affords ‘writing’, a book affords ‘reading’ and a football affords ‘kicking’ – it is also integral to the structure of everyday existence that this response-drawing accords with a subject’s situational purposiveness. For instance, although a hammer typically affords ‘hammering’, its role can be refined or altered according to specific circumstances; it could thus be employed as a metal forger, a wood-splitter (in conjunction with a chisel), a missile, a theatrical prop or a mountain climbing aid. Exactly what the hammer affords depends on the circumstantial disposition of a given subject in her given environment.

With this understanding, we can further refine our notion of the complementarity between a subject and her environment. It is not the case that worldly entities afford fixed possibilities for action, nor that they are encountered through context-independent properties; instead, entities ‘show up’ in subjective experience according to their relevance to circumstantial purposiveness. Such purposiveness can of course ‘switch’ across and within any context as attention can be diverted in myriad ways, but the connection between subject and environment remains a complementary one due to the environment’s flexibility in meeting subjective needs, interests and concerns. In the next section, I will further clarify this notion of nesting any given situation within an overarching network of subjective needs and concerns.

The brief ruminations of the preceding paragraphs have fleshed out further the nature of everyday existence within a broadly ‘Heideggerian’ framework. In particular, the complementarity of subject and environment helps to elucidate a specific aspect of the notion of being in-the-world. This is all the more relevant as it subserves the phenomenologically inspired theory of *enactivism*, which will be discussed at several points in later chapters. In a similar vein to a Heideggerian subject being embedded in a meaningful world, in which a subject is purposively disposed and environmentally sensitive, enactivism claims that a subject is always “bringing forth” a world by actively generating meaning through environmental interactions (Varela, Thompson and Rosch, 1991; Capra, 1996; Thompson, 2007; De Jaegher and Di Paolo, 2007). Both theoretical approaches thus reject the idea of an isolated subject who computationally represents a world through discrete propositions and instead endorse the idea of an embedded subject who enacts a meaningful world through purposively configured and historically sedimented expertise.<sup>[7]</sup> This notion of

---

<sup>[6]</sup> Whilst I am referring to *purposiveness*, Rietveld gives a unique description of a similar notion that he refers to as our *task-directedness* (Rietveld, 2008a, p. 342). That is, a subject’s expertise in a given situation exhibits itself a “task-directed system of possible actions [that] unifies many domain-specific capacities” (ibid.). So in any given domain, a subject’s task-directedness ensures that environmental entities afford possibilities for action that are relevant to a subject’s present “needs, projects, or interests” (ibid.). By this definition, a subject’s behaviour is always explicable as causally grounded in situational purposes, as it is through the notion of purposiveness.

<sup>[7]</sup> Aligning the affordance-led notion of complementarity between an agent and her environment with enactivism is a controversial move for some, in that the original formulation of enactivism is explicitly declared to be incompatible with Gibson’s (1979) ecological theory of perception (Varela et al., 1991, pp. 203-204). However, this incompatibility is due to the fact that Varela et al. interpret Gibson as claiming

embedded, in-the-world beings is henceforth how the term ‘subject’ (or ‘agent’) should be understood throughout the thesis.

### 5. **The ‘Social Environment’ and the Questions it Raises**

Thus far, we have looked at how Heideggerian ‘subjects’ – that is, in-the-world beings who enact their existence through co-determined subjectivity – encounter worldly entities during everyday activities. We know that such beings frequently engage with the world in a smooth, yet mindful manner, and that worldly entities can be transparent constituents of such beings’ ontology. However, what has not been considered is how such subjects engage with the social world; that is, how they interact with other subjects and social structures. As I mentioned at the outset, this chapter is primarily laying groundwork for subsequent chapters, and the matter of social encounters is particularly prevalent in later material. Prior to this later material, there are two important points connected to aforementioned issues that are worth briefly discussing.

Firstly, there is the fact that the expert know-how of daily activities is nested within networks of further contextual understanding, which inevitably taper towards social features of existence. In order to appreciate this, let us return to the idea that we encounter entities as having a ‘for-what’ quality (Heidegger, 1976/2010; Dreyfus, 2007a). In virtue of being for something, each entity is always *involved* in some activity that is performed by a subject, and is thus said to realise “an *involvement*” (Heidegger, 1927/1962, p. 115). That is, environmental entities are encountered as meaningfully relevant due to a “conceptually prior” involvement-structure (Wheeler, 2005, p. 146). Returning to the hammer, for example, we can say that it is only ‘for hammering’ if it is present in an involvement in which there are things to hammer (such as nails and wood) and a purpose for hammering which is relevant to a subject. Wheeler (2005) describes these *involvements* as “the roles that equipmental entities play – the ways in which they are involved – in the human agent’s normal patterns of activity” (p. 145). They are not independent structures, but are interconnected so as to form a holistic “network of referential significance” (ibid.). So the context of any given activity is part of a unique involvement-structure, which is in turn part of a system of further involvement-structures, which holistically connect to form a “totality of involvements” (Heidegger, 1927/1962, p. 121). This totality of involvements is essentially what Heidegger means by the “world” which we all find ourselves inescapably *in*. So we find

---

that perception is the direct detection of features within an independent environment (ibid., p. 204). Regardless of whether this interpretation is correct, I am using a broad notion of affordances (see, in particular, footnote [4]), rather than endorsing all of Gibson’s views on the matter. On this broad interpretation, affordances have a distinctively enactivist quality, which is that they are a *relation* between an agent and her world. As such, they can be used to capture the idea of *co-determination* between an agent and her environment, such that what an environment affords is dependent on the agent’s sensorimotor capacities and her history of coupling to the environment. This view is unquestionably something that most enactivists would welcome. The key move is that affordances are not simply an agent ‘picking up’ environmental features, but instead capture the (bidirectional) relation between the agent and environment.

that even in the simplest events, such as an entity affording ‘hammering’, there is a global context of significance in play:

*with* this thing[...] which we accordingly call a “hammer”, there is an involvement in hammering; with hammering, there is an involvement in making something fast, there is an involvement in protection against bad weather; and this protection ‘is’ for the sake of providing shelter[...] that is to say, for the sake of a possibility of Dasein’s Being. (ibid., p. 116)

As a subject engages in hammering in order to build a shelter, they are thus not only realising an afforded-action, but also realising a specific way of *Being-in-the-world*. Without delving too deeply into the complexities that are in play here, we can appropriate this notion of involvement-structures in order to confront the social-normative embeddedness of everyday existence.

Consider, for example, the following (paraphrased) description of the non-uniform referential links which make up Heidegger’s interconnected involvement-structures:

A subject can work *with* a computer (a relation that Heidegger calls a “with-which”), *in* the practical context of an office (an “in-which”), *in order to* write a book (an “in-order-to”), which is aimed *toward* presenting some philosophical analysis (a “towards-this”), *for the sake of* academic research, which is *for the sake of* being an academic (a “for-the-sake-of-which”). (Wheeler, 2005, p. 147)

Wheeler uses this to exemplify how Heidegger’s contextual structure will ““bottom out” in involvements which are for-the-sake-of-which in form” (ibid., p. 148). The importance of this is that engaging with everyday entities, which entails the realisation of an involvement-structure, is (in virtue of ‘bottoming out’ at being *for the sake of* something), in a fundamental way, “a concrete act of human projection[...] in which a human agent interprets itself in terms of certain behavioural norms” (ibid.). In other words, as a subject projects herself into an everyday activity, she will “interpret herself in definite ways” which are identified by “certain normatively constrained, public ways of behaving” (ibid., p. 122). So as a subject is expertly acting in a domain-specific manner, their skilled and unreflective behaviour is subsumed into a situated social normativity that is appropriate *for the sake of being* a certain way. Think, for example, of the waiter (Mr. X) who was discussed in section 3. He goes about his work with an attitude and movements that are generally expected of waiters. In this way, he is behaving in accordance with socially normative constraints (manifest in what is socially expected of a waiter), whilst also realising a system of involvement-structures which confronts the relation of acting *for-the-sake-of-being-a-waiter*. The waiter is encountering and engaging with a field of normatively-laden relevant affordances which facilitate self-interpretation *as a waiter*.

Now, if we take this idea and extrapolate it to all other situations (not just that of being a waiter), we find that the 'bottom' of our involvement-networks is inevitably social. Whether one's purposive engagement is ultimately for the sake of being a parent, a partner, a sports champion, a professional, a leader, a priest, an anarchist, or a waiter, one's fundamental relation to the world is inescapably social. This becomes clearer when we consider the normative roots of being a good parent, partner, champion, leader, or whatever. These roots are not individualistically manifest (if they were, there would be no consensus regarding typical behaviours across and within societies), but are socially generated and maintained through ongoing interactions. How one acts for-the-sake-of-being-a-parent or for-the-sake-of-being-a-partner is normatively regulated by a collective aggregation of what makes a good parent or good partner. In the most basic existential sense conceivable, relating to the world for-the-sake-of-being-a-human is itself a social relation, as individual humanness (normatively speaking) only makes sense from within the world of humanity (there will be more on this in chapters 2 and 5). In this way, "the human agent's everyday world is, in the first instance, and of its very essence, a shared world" (Wheeler, 2005, p. 149; Heidegger, 1927/1962).

By applying this underpinning social relationality at the behavioural level of engagement with worldly entities, we can uncover a final (social) insight into everyday existence. Recall, for a final time, the Heideggerian notion of encountering entities as 'for-what's'. This is a relatively straightforward feature of how the world is experienced, considering entities in and of themselves (albeit within given contexts). However, once the social normativity of involvement-networks is in play, it is perhaps now more precise to describe an entity as 'for-what-for-us'. This is due to the fact that no action can be viewed as a strictly individual accomplishment because every experience is saturated by sociality. A hammer, for instance, would seemingly be 'for-hammering' whether or not others are present. Yet the very act of hammering only makes sense in the social world; it only becomes meaningful in virtue of being, in part, an expressive act which reinforces one's own understanding of hammering whilst also informing the understanding of others. To hammer is to modulate the collective normative understanding of the activity of hammering. At a 'deeper' level, my act of hammering will modulate the normative understanding of what it is to be a good builder, or partner, or whatever worldly relation I am enacting. In this way, any pragmatic engagement with the contextual environment both takes its meaning from being *in* the social world and is inherently meaning-conveying *for* the social world. Even when hammering in a completely isolated fashion (far away from the presence of others), the act of hammering is still socially constituted because social meaning pervades even our most solitary instrumental copings with the world.

Through this brief exposition of the social nature of our ‘totality of involvements’ with the world, it is hopefully apparent that the social dimension of everyday existence cannot be ignored. Indeed, the particularities of the social dimension of everyday existence will go on to dominate the bulk of this thesis and will thus soon be analysed in much greater detail.

## **6. Conclusion**

In this chapter, several preparatory steps have been taken to claims that will be made later in the thesis about human existence (specifically when interacting with others). In reverse order, the key points that have been elucidated can be summarised as follows: (a). there is a normative social dimension to all human activity, (b). we are not isolated subjects, but are meaningfully *in-the-world* in such a way that there is complementarity between our existence and features of our environmental surroundings, (c). our everyday behaviour is dominated by smooth and unreflective engagement with the world, which, although non-deliberative, is still partially mindful, and (d). all of these preceding claims are underpinned by the turn towards Heideggerian cognitive science.

With these foundational points established, we can now progress to considering the social nature of our existence in more detail.

# **Chapter 2 – Encountering Others**



## Encountering Others

### **1. Introduction**

In the previous chapter, I outlined the recent ‘Heideggerian turn’ in cognitive science and how this paradigm shift has impacted on our understanding of how cognitive subjects engage with the everyday world. For the most part, the theoretical and practical progress that has been made through (broadly) ‘Heideggerian’ approaches to cognitive science has been impressive (see Arkin (1998), Brooks (1991), Colombetti (2014), Clark (1997, 2008), Damasio (1994, 2010), Gallagher (2005), Haugeland (1998), Varela et al. (1991) and Wheeler (2005) for a variety of positive steps in this regard). Yet there remains a relatively under-developed cavity in this otherwise strengthening field; namely, how does this paradigm shift impinge on the social world?

There are two perspectives from which I will address this question during the remainder of this thesis, each of which gravitates towards the other in a pincer-like movement. Firstly, in this chapter, I will look at a key feature of the human social world – how we encounter others – from a predominantly phenomenological perspective, which will also yield a better understanding of our ontological nature. In chapters 3 and 4, I will then look at the predominantly scientific field of social cognition and how current theories of social cognition lack a true appreciation of the consequences of the ongoing shift in cognitive science. The remaining chapters will combine preceding phenomenological and scientific insights in order to shed further light on the nature of human existence.

The first step for this chapter will be to consider the straightforward sense in which humans encountering other humans differs from humans encountering non-human entities. The straightforward difference will then be refined across sections 3 and 4, considering the *re-configuring* capacity of others (section 3) and the capacity of others to *reify one’s self* (section 4). Section 5 will then build on the earlier sections to elucidate the manner in which human existence is fundamentally intertwined with the lives of others.

### **2. The Relevance of Others**

One of the claims of the previous chapter was that in many everyday situations we engage with worldly entities in a smooth and unreflective, yet still partially mindful manner. The ease and smoothness of such engagements are largely due to an agent’s contextual expertise in coping with what certain entities are *for*; that is, what they *afford* (Gibson, 1979; Dreyfus, 2007a; Rietveld, 2008a, 2008b). As an agent encounters some worldly entity, the entity will offer up specific action possibilities in accordance with the agent’s expertise, needs and present disposition. If an agent wishes to build a shelter, for example, a hammer (or even just a hammer-esque object) will be encountered as a manipulable entity with the present functional purpose of being ‘for hammering’. The agent’s act of hammering is thus an

uncovering of the hammer's "specific 'manipulability'" (Heidegger, 1927/1962, p. 98), as well as uncovering the agent's present disposition and desire to hammer. The entity and agent achieve a kind of existential harmony through the hammer's suitability to satisfying the agent's present *task-directedness* (Rietveld, 2008a, p. 342) and, reciprocally, the agent's suitability to acting on what the hammer affords. What we find in such a scenario is that the 'harmony' between agent and worldly entity is entirely straightforward: the agent wants to hammer in order to build a shelter, and the hammer affords hammering. Notably, a comparable harmony can be achieved in instances where the hammer's manipulability is distanced from its socially determined role as a 'thing for hammering'. Given the appropriate situational circumstances, incorporating both the environmental context and the given subject's present needs and disposition, a hammer could be harmoniously engaged with through any number of instrumental roles within various pragmatic activities. Although the hammer can be said to have a *specific* manipulability (as 'a thing for hammering'), it can also be put to a vast array of further functional uses, and these further uses can be accommodated in skilled and unreflective activity with an ease that is comparable to that of the typically expected purpose of hammering. In other words, in spite of its socially accepted name and 'specific manipulability', a hammer doesn't *only* afford hammering. Nor is any other entity limited to a single action. A book, for example, is generally considered to afford 'reading', but given the appropriate circumstances it could also be used (in a skilled, smooth and unreflective manner) as a paperweight, a cushion, a building block, or a door-stop, amongst many other things. What we find, then, is that even the simplest entities, with seemingly straightforward instrumental roles, can in fact be highly variable in the manner and number of ways in which they can be manipulated. If we find this inherent complexity with basic everyday items such as hammers and books, what exactly do we encounter with the emotional and functional supernovae that are other people?

The first thing to point out is that other people do not have a straightforward 'specific manipulability' in the same way that most non-human entities do. Unlike objects such as hammers or books, we are never able to truly engage with other people as some sort of functional tool; instead, we are always latently aware of their being another person and therein being more than an inanimate and lifeless 'thing'. To quickly exemplify this, consider ordering a meal in a restaurant. As I place my order with the waiter, I maintain an awareness that she is not just a 'tool' for taking my order, but also has the potential to afford a plethora of intricate and varied actions. If I start to choke, I am aware that the waiter could potentially help me, or, if I need directions, she could provide them. If I need comforting or feel the need to vent a build-up of anger at someone, then my expert knowledge of human capacities means that I am aware of the waiter being a potential candidate, alongside an awareness that I can offer the same possibilities for her. Thus, even if the waiter is fulfilling the socially

determined functional role of a waiter, and I am treating her as such, there remains an awareness that she is not *just* a waiter in an exhaustive sense, but is also (simply in virtue of being a fellow human being) a potential profusion of further action possibilities. This is not to say that humans are unique in moving beyond a purely functional role – a sentimental object such as a family heirloom or gift from one’s partner can obviously be more than functional too. The point is that although non-human entities are sometimes more than functional, humans cannot be anything but more than functional. As Sartre puts it, the waiter can never *be* a waiter in the sense that an inkwell *is* an inkwell, or a glass *is* a glass (1943/1984, p. 59).<sup>[8]</sup>

When an agent encounters another, we thus find that an unrivalled degree of potential complexity is imported to the situation. Non-human objects, such as smartphones, laptops or pets, have the potential to afford a great many actions, but even these will pale in comparison to those afforded by another person. Encountering another person within some situational context is, from a functional perspective, akin to encountering a highly developed Swiss army knife of practicality; and alongside this, we have the emotional perspective from which another person can offer up “possibilities such as conversation, companionship, consolation, love, humiliation, pride and shame” (Ratcliffe, 2013a, p. 159). A person, then, simply does not have a *specific* socially determined and accepted ‘manipulability’. Rather, other people are highly manipulable (and, as we shall see shortly, *manipulating*) smörgåsbords of action possibilities. Any situation in which another person is present brings with it a potential for variability and complexity that simply isn’t attained with non-human worldly entities.

To put this in a little more detail, recall the “mutual relationship” that is established between a smoothly coping agent and her environment (Gallese and Sinigaglia, 2011; p. 122; Rietveld, 2012a). The environment ‘opens up’ in an accommodating manner towards the present needs and disposition of the agent, such that there is a purposive tendency towards entities that *matter* to the agent (Rietveld, 2012a, p. 212). When this idea is applied to other humans, I believe the unsurprising outcome is that the presence of (an)other human(s) within a given phenomenological context will *always* be relevant to the present needs, interests and disposition of an agent. Others always impinge, at least in some minimal way, on an agent’s present worldly engagement (there will be more on this throughout the subsequent sections). It is thus not only that other humans embody an unrivalled level of complexity, but that they are also uniquely *relevant* to our everyday existence. Moreover,

---

<sup>[8]</sup> In Sartrean terms, the waiter is primarily a ‘being-for-itself’ (*pour soi*) while the inkwell is a ‘being-in-itself’ (*en soi*) (Sartre, 1943/1984). The idea here is that objects such as the inkwell simply *are* – they are passive, inanimate and determinate. Sartre describes such entities as descriptively exhaustible through their *facticity*; that is, their determinate characteristics. Entities such as humans, however, transcend such facticity in virtue of being descriptively inexhaustible through their determinate characteristics (although facticity remains an important feature of our existence) (*ibid.*). As the chapter progresses, this distinction will be seen to meld nicely with features of human existence that become evident in our encounters with others.

there is the crucial fact that other people are subject to their *own* needs, interests, dispositions, and capacity to manipulate the environment. Whereas my encountering a hammer, book or inkwell is predominantly one-sided, in that the objects themselves have no ‘choice’ in how they are engaged with, the same cannot be said of my encountering another person. Other persons have their own *autonomy*<sup>[9]</sup> and can thus influence and adapt to situations on their own terms. This autonomous nature of other people means that an element of unpredictability enters the situation when agents encounter one another, as the environment can be subject to either (or both) agent’s manipulations. In terms of affordances, the “permanent features” of inanimate entities will flow in and out of relevance in accordance with a subject’s needs and disposition, but the autonomous “temporary state[s]” (Gibson, 1979, p. 42) of other individuals are something that an agent will always be potentiated to respond to. That is, one’s understanding of others includes an awareness of their autonomous capacities that one may potentially act on, or have to react to. This awareness of others’ autonomy can be captured through the notion that others always have the potential to *solicit*<sup>[10]</sup> one’s attention, such that another’s presence is always a relevant intrusion into the phenomenological scene (with the nature of this intrusion being elucidated in the subsequent sections).

When considering the encountering of others, we thus have to step away from the basic conceptual descriptions that can be applied to our encountering of non-human entities, simply because human autonomy brings with it complexity and relevance that non-human entities do not currently achieve. A quick scan across cognitive-scientific literature wholeheartedly endorses such a view. Evolutionary psychologists, for example, would be swift to point out that our attunement to the relevance and complexity of others is to be expected “because we have adapted, over generations and within our lifetimes, to be sensitive to each others’ smell, sight, behaviors, creations, emotions, and thoughts” (Theiner, Allen and Goldstone, 2010, p. 380). At a neurobiological level, there is evidence that the release of specific hormones such as oxytocin will increase trust between humans (Kosfeld, Heinrichs, Zak, Fischbacher and Fehr, 2005; Theodoridou, Rowe, Penton-Voak and Rogers, 2009), suggesting that the relevance attached to the perceived presence of another can be

---

<sup>[9]</sup> ‘Autonomy’ is a contentious term as it is so closely tied to the notion of freedom and thus encroaches upon the ‘free will debate’ (see Bratman (1979), Frankfurt (1988), Van Inwagen (1983) and Wolf (1990) for a range of views). In chapter 4, autonomy will be given a technical definition in accordance with the theory of *enactivism*. Here, however, I am using it in the straightforward sense of agents (seemingly) displaying free will and self-governance, thereby being capable of making their *own* decisions and acting on the world of their own volition.

<sup>[10]</sup> *Solicitations* are affordances that complete an active process by drawing forth a response from an agent (see Dreyfus & Kelly (2007, p.52) for an excellent explanation). They are not merely affordances which we are poised to respond to, but actually result in an alteration to active behaviour. To say that an entity has “the potential to *solicit* one’s attention” is to classify it as an affordance which is “on the horizon of my experience summoning my attention as potentially (not merely possibly) relevant to the current situation” (Dreyfus, 2007a, p. 28).

Whilst solicitations are not unique to responding to the other persons – for example, Merleau-Ponty (1964), Rietveld (2012b) and Dreyfus & Kelly (2007) employ the term to describe how inanimate entities call forth appropriate actions from us – it is a useful concept for capturing the fact that other persons are always *relevant* to a given situation.

influenced intranasally, as well as through the more standard phenomenological routes of sight, touch and hearing. Furthermore, recent studies have shown that we (as humans) are keenly attuned to the dynamics of biological motion, in that we can pick it out from other motion types (Scholl and Tremoulet, 2000), and we can even decipher gender and emotion purely from motion dynamics (Kozlowski and Cutting, 1977; Dittrich, Troscianko, Lea and Morgan, 1996). Such keen attunement suggests that we can be readied for more trait-specific activities (related to another's gender, emotional state or physical capacity) within our awareness of others' holistic relevance, courtesy of something as basic as another's movement. Forays such as these into other disciplines are clearly worthy of detailed discussion themselves. For present purposes, however, their significance is simply to add weight to the idea that the phenomenological relevance of others can not only be vastly varied, but we can also be attuned (and potentiated to respond) to their presence through various fine-grained perceptual means.

Across the next two sections, I will unpack this straightforward notion of humans being more *complex* and *relevant* than non-human entities through more detailed phenomenological analyses, focusing on the *(re-)configuring* and *reifying* impacts of others.

### **3. The (Re-)Configuring Presence of Others**

As noted in the previous section, part of the reason that humans stand out in terms of what they afford is that they are always more than mere functional tools. Encapsulated within this notion of being 'more than mere functionality' are two key features of human ontology: *autonomy* (as previously mentioned) and *affectivity* (having moods and emotions such that one can affect others and be affected oneself (Deleuze and Guattari, 1980/1987)). These two features are inherently interrelated in that one's autonomous behaviour and affective disposition are reciprocally dependent on one another, in the sense that a human could not be non-affectively autonomous. However, when considered one at a time, each of these ontological features can shed light on how the presence of others induces a *(re-)configuration* of a given agent's phenomenological landscape.

Firstly, consider that as an agent engages with some worldly entity, there are always other entities on the 'horizon' of the agent's experience that have the potential to solicit one's attention and produce other action possibilities (Cappuccio & Wheeler, 2012; Dreyfus & Kelly, 2007; Rietveld, 2008a, 2012a, 2012b). For example, as I am stood in my kitchen reading a cookery book, I am still aware of the kettle affording 'boiling water'. That is, the kettle still offers a relevant *possibility* for action, even though I am not currently responding to it and, indeed, may not respond to it at any point in the near future. A simple way of putting this is that I am *potentiated to respond* to the kettle's affordance (Rietveld, 2012a), with such potentiation involving a *readiness to act* (Wrathall, 2000; Dreyfus, 2000;

Rietveld, 2008b, 2012b). Dreyfus (2000) describes this readiness as “something that is not an activity but is more active than a capacity” (p. 339). In other words, whilst I am reading a cookery book in my kitchen, my relationship to the kettle is more than a mere dormant capacity – in that my worldly expertise endows me with an awareness of the kettle’s possibilities for action – but less than actual activity – in that I am not presently acting on it. Instead, I am somewhere between these two poles: beyond being merely capable of responding to the kettle’s affordances and yet not sufficiently drawn by its ‘allure’ to presently respond to it. I am in some sense poised to respond to the kettle should circumstances change in a way that results in it aligning with my present needs and disposition. Now, when it comes to other humans, our affective nature means that I am not only potentiated to respond to pragmatic possibilities for overt action (in the simple way that I am potentiated towards an ordinary kettle), but I am also potentiated towards the *affective* possibilities offered by another. That is, the possibilities brought forth (or potentially brought forth) by another person will permanently have an intensified dimension to their ‘allure’ because a given agent will always be *affectively potentiated* to respond to them. Of course, one could claim that human agents are always affectively potentiated to non-human objects as well as other persons (e.g. Damasio, 1994, 2010; Colombetti, 2014), in the sense that human autonomy is always affectively laden (as I noted at the outset of this section). However, the affective potentiation towards other persons differs in kind from potentiation towards non-human entities for reasons that will be accumulatively developed throughout the remainder of the chapter.

Initially, there are two ways to view this affective potentiation towards others. In a straightforward sense that aligns with the considerations of the previous section, we can say that the presence of another in the phenomenological scene can affect us in a greater variety of emotional ways than non-human entities typically do; that is, others have the potential to make us sad, happy, ashamed, proud, jealous or excited etc. However, in a deeper sense that begins to move us towards what is unique about encountering others, this affective potentiation is a shift in one’s phenomenological world. Once another has been encountered, an agent is no longer engaging with the world from a ‘private’ stance, but is instead subject to the affectivity of the other(s). There is thus a holistic transformation to a phenomenological perspective that is open to the direct affective influence of others. In a particularly profound way, such a transformation is exhibited by people in love (De Jaegher, 2013), where all aspects of the world may come to be seen, in part, through the mindset of one’s partner. Indeed, the vicissitudes of love capture the “simultaneous enjoyment and suffering of one’s own life[...] as modulated, transformed, moved, upturned by another” (ibid., p. 127). In such a case, one may be affectively transformed to the extent that there is a complete dispositional overhaul. However, even mere acquaintances may transform one’s world through a

subjective readying towards others' potential for affective modulation. If, for example, I am preparing dinner in a communal kitchen and another person enters the area, I become readied for a potential social interaction: to exchange pleasantries, share cooking equipment, converse about any number of matters, or simply to move around one another in a mutually non-antagonistic performance of indifference. Even if no noticeable interaction unfolds, I am implicitly transformed by the *mere potential* of a social encounter – by the simple bodily presence of the other within the tangible vicinity of my present phenomenological world. For Michel Henry (1963/1973, 1965/1975, 1990/2008), this transforming potentiation is the root of what it is to experience another *as another*. He describes it as

a desire seeking out some sort of response or nonresponse, an emotion before the reciprocity of this desire, a feeling of presence or absence, solitude, love, hate, resentment, boredom, forgiveness, exaltation, sorrow, joy, or wonder. (Henry, 1990/2008, pp. 103–104)

In other words, we harbour a fundamental affective connection with others that transforms our world when another enters the phenomenological scene. This connection with others – as fellow affective beings – leaves us affectively poised for significant modulation even prior to any developed social interaction taking place. A mere passing-by of one another or sharing of some space results in a phenomenological transformation in virtue of others' affectivity and the affiliate affective potentiation that this induces in a given agent. For the likes of Henry, such inter-affectivity is a fundamental aspect of human ontology (see De Jaegher (2013) for an interpretation of Henry that is relevant to modern approaches to intersubjectivity) and, indeed, a highly similar claim will be made in section 5.

An analogous transformational power to that described above can be elucidated by approaching the same issue from the connected perspective of others' autonomy. Due to the capacity of others to act on and manipulate the world, their presence shifts an agent's phenomenological outlook away from the kind of subjective governance that the agent has when alone. If I am alone with a cup of tea then the cup offers the possibility for 'tea drinking' in a very clear-cut manner. However, once another enters the phenomenological scene, the possibility is now one of 'tea drinking *in the presence of another*'. My phenomenological landscape becomes open to penetration by the other's needs and disposition, such that the action possibilities offered by present entities are no longer truly *mine* in an immediate phenomenological sense, but are coloured by the other's connection to the immediate environment and the possibilities for action that match with their purposiveness. The other therefore indirectly robs me of the 'mineness' of the situation and renders the context an interpersonal domain – not merely in the sense of there being two persons present, but in the sense of all domain entities being denotative of the other's existence. What this amounts to is a *re-configuration* of the phenomenological landscape,

blanketing the situation in other-referentiality. Merleau-Ponty (1945/2012) captures this re-configuration with the idea that in another's presence "objects[...] receive a new layer of signification: they are no longer merely what I could do with them, they are also what this behaviour is about to do with them" (p. 369). With the onset of this 'new signification', the other "tear[s] me away from my immediate inwardness" (ibid., 1964, p. 136), importing me to a phenomenological world that is enshrouded in the other's presence. Whilst non-human entities may have profound phenomenological impact, they tend not to categorically import the same kind of isotropic configurability that the presence of another does.

What we find with the presence of others, then, is that their influence is not limited to their complex manipulability or their arousal of occurrent emotions. Instead, the other re-configures the nature of the whole phenomenological landscape, such that a given agent must relinquish independent control over the situation and address a world that is directly imbued with interpersonal significance.<sup>[11]</sup> Such (re-)configuration occurs in the mere presence of another, in that we are potentiated to respond to their affectively autonomous actions and appraisals. In encountering another's presence, we thus encounter "a transformation in the experience of possibilities that is constitutive of our sense of others" (Ratcliffe, 2009, p. 209).

Yet again, it is worthwhile noting that we can find empirical support for this phenomenological claim. Samson, Apperly, Braithwaite, Andrews and Bodley Scott (2010) conducted a series of experiments, one of which involved a subject who was placed in a room with an avatar that could see just two of the four objects that the subject could see. The subject then had to respond as quickly as possible to numerous prompts about whether the objects (considered one at a time) could be seen from their perspective and the avatar's perspective. They found that the *mere suggested presence of another person* interfered with participants' response times and error rates by creating an "altercentric intrusion" (ibid., p. 15) in judgement. These altercentric intrusions arose "even when participants were given a clear opportunity to ignore the other person" (ibid., p. 29). Samson et al.'s conclusion includes the claim that even the mere suggested presence of another person interferes with an agent's personal perception of environmental features. That is, the other's presence could not simply be discounted as situationally irrelevant as it affected the agent's ability to say which objects they could see. In other words, there is a case for arguing that the other (i.e. the avatar) has *re-configured* the agent's context such that the agent's (egocentric) perspective is now (altercentrically) imbued with the presence of the other. At least in certain circumstances, then, there is psychological backing for the idea that another's presence

---

<sup>[11]</sup> Of course, in a very important sense the world is *always* imbued in interpersonal significance (see chapter 5). However, the interpersonal significance that is referred to here is that of reference to a *non-specific other*, not the general sense of sociality that pervades our existence.

influences the manner in which an agent encounters their environment, transforming the agent's experience of possibilities.

Two potential caveats regarding the re-configuring potential of others need to be swiftly addressed. Firstly, with regards to Samson et al.'s study, there is the fact that the 'other' in question is an avatar. In light of the earlier comments regarding the affectivity and autonomy of others, along with subsequent claims that will be made in section 5 and chapter 6 regarding the uniqueness of human nature, this is relevant in that the avatar is merely a weak tracing of a human 'other' from a phenomenological perspective. Nonetheless, the study still nicely illustrates a general aspect (i.e. the taking into account of another's perspective) of the re-configuring phenomenon that I have been describing. Secondly, it is worth noting that in exceptional circumstances a non-human object may holistically re-configure the phenomenological landscape: for example, a highly religious Catholic may find their phenomenological landscape re-configured by the presence of a crucifix, so that all present entities refer (in part) to the crucifix and its religious connotations. Similarly, a family heirloom may re-configure the phenomenological landscape in reference to one's ancestry. However, I would suggest that such cases are rare (within the scope of all non-human entities that are encountered throughout a day), whereas the re-configuration of the phenomenological landscape is a permanent characteristic of human presence. What's more (and as noted previously), the uniqueness of human nature will be further expounded in section 5 and in chapter 6.

In the next section, more detail will be added to this idea of re-configuration through the notion of *reification of the self*.

#### **4. Others and the Reification of One's Self**

Recall that the approach that has been taken thus far regarding human engagement with everyday entities involves a kind of pairing between an agent's present needs and disposition and what the manipulability of a worldly entity can afford. In general, this approach relies on the premise of everyday engagement with the world implicating the perspective of an *acting* agent. So, for example, understanding the phenomenology of an event in which hammering occurs relies, in part, on understanding an agent's actual engagement in the act of hammering. Importantly, there is another, complementary perspective that we can take: that of the *act-able* subject. This perspective considers one's (potential and realised) *capacity* to act in certain ways, rather than considering the actions themselves. So understanding the phenomenology of an event in which hammering occurs also relies, in part, on understanding that a subject *can* in fact manipulate a hammer in an appropriate way. Thus, for a subject to encounter a worldly entity as a thing 'for hammering', they have *to be able* to hammer – to experience themselves as 'a being that is capable of hammering'. One's

readiness to engage the hammer therefore naturally entails an unreflective awareness of one's capacity to carry out the necessary movements for hammering. Without such awareness, the hammer simply wouldn't afford hammering; the act of hammering wouldn't be part of the experiential repertoire or basic motor capabilities that make up one's expertise for acting in the everyday world. We thus find that, as Gallese & Sinigaglia (2011) put it, "in perceiving something as graspable, throwable or kickable, we are experiencing ourselves as bodies that can grasp, throw or kick" (p. 127). Indeed, for Gallese & Sinigaglia, this kind of self-experience is an integral phenomenological constituent of our active living:

our experience of the surrounding things cannot but be accompanied by the experience of ourselves as a bodily 'power' for action, that is, as the variety of action possibilities belonging to our own motor repertoire (ibid.).

In engaging with and acting on the world, we thus have a sense of self that is "given to us as a manifold of possibilities for action" (ibid., p. 130). Notably, this sense of self is not a deliberative assessment of one's capabilities, nor a concentrated kind of introspection; it is, rather, an implicit appreciation of one's own body as capable of performing a certain act. In order to substantiate such a bodily sense of self, Gallese & Sinigaglia draw on Noë's (2009) position that one's body can be a 'power for action' whether or not one is explicitly attending to it, such that "my arms can be present to me now, even though I am not now thinking of them: the feeling of their presence comes down to such things as my sense that the coffee cup on the table is within reach" (p. 77). This accords nicely with our commonsense conception of our bodies-in-action: when I move to shake someone's hand, have a drink of water or control a football, my attention tends to be on the other person, the glass of water or the ball, not on my own body (Gallagher, 2005, p. 26).<sup>[12]</sup> The idea is thus that the experience of one's body as capable of acting in some way does not entail explicit choice or analysis, but is simply a constitutive experiential aspect of one's expert coping with the world. Referring this idea back to our everyday responsiveness to affordances, one can say that the 'allure' of certain worldly entities is simultaneously accompanied by the 'allure' of one's own body as capable of engaging with these entities.

The question that drops out of this with regards to present concerns is what exactly is the bodily 'power' for action that we experience when we encounter another person? That is, if

---

<sup>[12]</sup> Notably, our bodies are not always simply on 'automatic pilot' below our conscious level of attention. As Gallagher (2005) explains, consciously attending to one's body can be crucial to learning new skills or overcoming some challenging situation. For example, learning a specific tennis stroke will involve attentive monitoring and correcting of bodily movements; similarly, "movement along a narrow ledge above a deep precipice may involve a large amount of willed conscious control based on the perception of[...] limbs" (ibid., p. 27). However, even in these cases of directing one's attention to one's body there will remain bodily capacities that are non-attentively employed in order to make the consciously willed bodily activity possible. In certain circumstances the body can thus be both a 'power' and 'target' for action.

one encounters oneself as ‘capable of hammering’ when one encounters a hammer, what is the self-capability one realises when encountering another person?

The first step to answering these questions is to return to the idea that another’s affective autonomy can re-configure a given agent’s phenomenological landscape. Upon encountering another, one’s world takes on a layer of affective reference, as well as being open to the other’s autonomous (positive or negative) activity. As Sartre (1943/1984) puts it, the other can become an “indispensable mediator between myself and me” (p. 222). That is, my actions come to be played out in an interpersonal domain such that my expert coping with the world becomes coping “*before somebody*” (ibid., p. 221). Experiences are no longer uniquely ‘mine’, but are mediated by the other’s presence and their configuration of the entities surrounding me. In other words, “[t]he Other as the synthetic unity of his experiences[...] comes to organise *my* experience” (1943/1984, p. 226).

This description of the other’s mediating presence leads us to the phenomenon of *reification of one’s self* that I believe occurs when one autonomous agent encounters another. That is, the other’s capacity to appraise an agent (and their actions), or wrest control of entities within the agent’s phenomenological landscape, forces the agent into a kind of non-reflective self-awareness. It makes one aware of one’s existence as an entity that can be perceived (and therefore judged, helped, hindered, threatened, loved etc.) by another agent, rather than just being aware of oneself via one’s immediate engagement with the world. The other person therefore ‘mediates’ between oneself as an existentially transparent perceiving agent – who can smoothly engage with worldly entities in a way that resists a phenomenological subject/world dichotomy – and an existentially ‘opaque’ perceivable agent – who is conspicuous via one’s autonomous presence. Before another person, an agent ceases to be phenomenologically ‘transparent’ from his/her own perspective and instead becomes a distinct worldly entity: perceivable, appraisable and capable of affording responses in much the same way that others do for them. When two (or more) autonomous agents reciprocally encounter one another, each is thus encountering herself too, finding their agential self to be reified by the other’s presence.

Put another way, we can say that the other’s mediation of the phenomenological landscape – rendering it an interpersonal domain – makes one aware of one’s *own* autonomous presence, as well as that of the other. As the situation becomes indicative of the other, so that entities afford actions of the form ‘*x before another*’, so, too, is one aware of the fact that the shared phenomenological landscape will now show up to the other as indicative of oneself. One’s own presence is reified as one is made acutely aware of one’s perceptibility and appraisability, along with one’s ability to perceive and appraise. Before another, one thus no longer ‘simply lives’ one’s experiences, but partly lives them ‘as one appears before the other’; reciprocally, one is aware of one’s own power as an autonomous agent before whom the other

must live their own experiences. As such, we find that “[o]pposite the Other and confronting the Other, each one asserts his right of being individual” (Sartre, 1943/1984, p. 236).

By applying this idea to the wider context of the chapter, we can propose that a key feature of encountering others is experiencing one’s own ‘encounterability’ – one’s reified existence as an autonomous being. This encounterability encapsulates the intertwined nature of one’s power to act on the world and, simultaneously, one’s susceptibility to the power of another. At the heart of this is the basic idea that when encountering others we confront the fact that they can “have a perspective not just upon the world of objects, but upon us too” (Zahavi, 2014, p. 193).

Crucially, however, the notion of reification of one’s self that emerges from the reciprocal encountering of another autonomous agent need not imply an analytical kind of introspection or contemplative self-reference. Instead, the performances of one’s reified self can still be played out in the smooth and unreflective manner that typifies one’s engagement in the everyday world. To explicate this idea, we can employ Seigel’s (2005) distinction between *reflective* and *reflexive* self-awareness. For Seigel (2005), “[r]eflectivity refers to the ingredients of intellectual self-awareness in selfhood, the contribution made to it by the mind” (p. 12), such that any reflection “establish[es] a “second-order” relationship to the contents of experience” (ibid.). In this sense, reflection is an abstraction from experiential thinking or acting, an explicit consideration of real-time embodied behaviour. In contrast, “something is reflexive if it simply doubles or reinforces its origin” (ibid.) in the manner of a mirror reflecting the world. There is nothing self-directed or attentive about reflexivity, it is simply an automatic or involuntary response that concretises a phenomenon (in this case, one’s experiential existence). Whereas reflective self-awareness is “active”, “intentional and purposive”, reflexive self-awareness is “passive” and “unwilled” (ibid., p. 12-13). Although Seigel goes on to lament the fact that “in many instances it is difficult to say which [reflectivity or reflexivity] is at work” (ibid., p. 13), his distinction comes in highly useful at the present juncture. What we find is that Seigel’s *reflexive self-awareness* is a natural descriptor for the reification of one’s self that is induced by the phenomenological presence of another person. Before another, one’s presently active (autonomous) self becomes existentially ‘opaque’ and conspicuous, not in the sense of being reflected upon from a ‘second-order’ stance, but in the sense of being mirrored by the autonomous presence of the other. A given subject can thus still act in an unreflective, non-self-referential manner before another, but their implicit understanding of their agential self – their capacity to act in the present circumstances – becomes passively reinforced.

Another way to justify the unreflective reification of one’s self that is experienced in the presence of others is by appropriating Gallagher’s (2000, 2005) notions of ‘the sense of agency’ and ‘the sense of ownership’ in motor action. For Gallagher, both of these ‘senses’ are

“closely related aspects of minimal self-awareness” (2000, p. 16): the sense of ownership is “the sense that it is my body that is moving”, and the sense of agency is “the sense that I am the initiator or source of the action” (ibid.). In the vast majority of cases, both ownership and agency come as a kind of package deal – I am implicitly aware that it is *my own body* that is moving and I am implicitly aware that *I* (as an agent) intended it to move in such a way. However, in the case of involuntary action, or in the actions of some schizophrenic patients (Frith, 1992), it is possible to have a sense of ownership – knowing it is one’s own body that is moving – but not a sense of agency – no “sense of causing or controlling the movement” (Gallagher, 2000, p. 16). Importantly, howsoever agency and ownership are manifest, the self-reference that is implied is “immediate and non-observational”, in that “I do not go through some cognitive process in which I try to match up first-person experience with some known criterion in order to judge the experience to be my own” (Gallagher, 2000, p. 15). Rather, it is a form of unreflective access to one’s ‘self’ that involves no contemplative act of cognition.

If we unite this distinction between self-ownership and self-agency with earlier discussions about others’ ‘configuring’ abilities, then it follows that one’s sense of agency will be (re-)configured in line with one’s transformed world. Although I am still the initiator and source of my actions when I act before another, my actions are no longer solely *mine* in the same way that they are in the other’s absence. They become ‘my actions in an interpersonal domain’ – my actions *before another*. My sense of agency is thus suffused with the other’s presence. I still intend and control my actions, but every speech, gesture or posture is now indicative of the other – my actions are played out in a world which is *his* or *hers* as well as *mine*.<sup>[13]</sup> More than just being a causal motivator of behaviour, the other’s presence (as discussed earlier) *mediates* my actions, so that they are not my actions in the same way that they are when I experience myself as phenomenologically alone. Such mediation is evident in even the most commonplace behaviour: an everyday action such as eating a sandwich or drinking tea will entail a referential awareness of another so long as they are present in the phenomenological situation. In other words, if another is present, I do not act purely on my own terms: I initiate and control my actions, but do so in accordance with the other’s autonomy, as if we have a fleeting circumstantial rulebook drawn up between us.

---

<sup>[13]</sup> Describing the manner in which one’s sense of agency is modulated as partly losing the ‘mine-ness’ of one’s agency, or as it being suffused with another’s presence, may suggest that it is always a diminishment of one’s agency that occurs before another. During the incipient moments of non-specific social encounters, I think this is generally the right stance to take. However, it is worth noting that the alteration of one’s sense of agency may become much more complex once the full scale of one’s intersubjective constitution is considered (see chapters 5 and 7 for more on intersubjective constitution). For instance, if one were to dominate an interaction with another, then it is highly plausible that one’s sense of agency would in fact be strengthened. If one were to take on a subservient role in interaction, on the other hand, one’s sense of agency would perhaps be significantly diluted. This variation grows once one considers the likes of loving dyads, in which one’s authorship of one’s actions may be contingent on one’s partner (i.e. one’s agential self is filtered through one’s partner). These are speculative claims, but it is notable that once one moves the phenomenological analysis to specific persons and specific interactions, the modulation of one’s sense of agency will succumb to nuanced transformations.

This mediation of one's self-agency clearly adds further conviction to the earlier claim regarding the re-configuring effect of others. At the same time, I believe the presence of another can also heighten one's sense of ownership. As the other wrests partial control of the phenomenological domain, thereby altering my sense of agency, they also reinforce my implicit understanding that it is *my* body that is moving. That is, the sense of ownership of one's active existence (presently constituted by one's actions) is exacerbated by the other's autonomous capacity to encounter oneself. One's ownership of one's actions thus becomes steeled against the other's potential for appraisal and involvement. As agential control is modified, we find that one's sense of ownership is heightened, therefore supporting the reciprocal phenomenon of reification of one's self.

In order to not misappropriate Gallagher's (2000, 2005) agency/ownership distinction, it is important to point out that one could question my preceding analysis on the grounds that even in another's presence, willed action is still emphatically part of one's subjective consciousness. So even as others re-configure the phenomenological landscape, one could argue that this is merely a contextual transformation and one's sense of agency (i.e. one's sense of initiating one's actions and thoughts) is unchanged per se. However, I believe that Gallagher construes the sense of agency as more flexible than this. Consider, for example, the suggestion that there is a certain lack of a sense of agency if one has a melody stuck in one's head when one would rather be contemplating something else (Gallagher, 2005, p. 174). As noted earlier, schizophrenia also seems to disrupt an agent's sense of agency, with schizophrenics feeling alienation not just from their agency but from many aspects of selfhood (ibid., pp. 173-205). Surrounding these clear-cut examples, it seems fair to assume that there are more blurry fluctuations in one's sense of agency. My claim here is simply that the presence of another modulates – perhaps in a very minimal way – one's sense of authoring one's actions. The 'I' that acts before another is no longer an 'I' of circumstantial privacy, but one of interpersonal consideration. As such, one's sense of agency before another is a sense of oneself as an *agent-before-another* (which, in the terms of Heideggerian 'involvement structures' that were discussed in the last chapter, is a kind of parenthetical addendum to one's relation to the world, such that one acts for-the-sake-of-being-X(-before-another)). At the same time, however, one tends to become more bodily self-aware before another than one generally is when alone, such that one's sense of ownership is heightened.

Whether viewed through the lens of Seigel's (2005) *reflexive self-awareness*, or via Gallagher's (2000, 2005) aspects of minimal self-awareness (the *sense of ownership* and the *sense of agency*), it is the mediatory *reification of one's self* that is induced by others that I believe is phenomenologically relevant. Encounters with certain non-human entities may affect a given agent in a powerful and pervasive way; they may even hold a significance that re-configures the agent's phenomenological landscape as denotative of their presence. Yet it is

only encountering other humans that permanently imports these phenomenological transformations, along with the reification of one's self.

### **5. The Fundamental Intersubjectivity of Human Ontology**

Across the previous three sections I have outlined some phenomenological features which justify distinguishing human encounters with other humans from human encounters with non-human entities. An underlying assumption of this claim is that humans are themselves distinct in some way from non-human entities. In many ways, this is a completely unremarkable statement, but the exact manner of this unique 'human-ness' is a matter of much debate. For example, human distinctness has been variously attributed to rationality (Aristotle, 2002), use of language (Gadamer, 1960/1992; McDowell, 1994; Maturana, 1978, 1995, 2002), radical freedom (Kierkegaard, 1844/1980, 1847/2009, 1849/1983), inviolable dignity (Maritain, 1947/1985; Wojtyła, 1969/1979), unpredictable innovation (Lieberman, 2013), empathy (Scheler, 1913/2008; Stein, 1917/2008) or having a theory of mind (Pagel, 2012). For my own part, to the extent that we may want to say that there is in fact a unique human 'nature', I favour the view that such a nature is distinct because of the *social normativity* that pervades all human experience. That is, I believe that we (humans) intersubjectively generate and modulate unique sets of norms that reciprocally generate and modulate our overt actions and 'inner' cognition. Of course, one could search for an underlying cognitive mechanism or transcendental quality that makes such normativity possible, but a pursuit of this kind would take us beyond the experiential 'level' of human existence to a neural description or a description that focuses on the conditions of possibility of our existence. Underlying approaches such as these tend to address certain aspects of the unique nature of human existence whilst downplaying others,<sup>[14]</sup> whereas a normative picture of human existence comprehensively captures a balanced view of human nature (as we will see later on). In chapters 5-7, I will demonstrate that human normativity is a socio-historically rooted phenomenon that is both evolutionarily and developmentally manifest through the expressivity of interacting human bodies. That is, we are not unique in merely engaging with norms of certain kinds, but also in being nascently *disposed* towards such (unique) normative engagement. Due to this character of normativity, I believe it provides an explanatorily satisfactory account of the unique nature of human existence without necessitating an appeal to further qualities, whilst maintaining a focus on all features that are relevant to the uniqueness of our being.

---

<sup>[14]</sup> For example, an account of human nature that focuses on transcendental qualities of our existence may fail to satisfactorily address our concrete physicality or the interactive dynamics that evidently occur in daily encounters. Similarly, an account that focuses on neural mechanisms may downplay the role of the body, or, conversely, may treat our sociality as a mere context that provides causal inputs for cognitive processing. Approaches such as these fail to fully account for the multi-faceted nature of human selfhood and thereby produce unbalanced notions of human existence. In chapter 7, I will address this issue and demonstrate how a focus on experiential normativity can satisfactorily capture a notion of human existence that is constituted by both the biological body and the social world.

At this point, the introduction of normativity serves the purpose of facilitating an explanation of how humans are ontologically connected to one another, and this ontological connection is itself an explanation for why humans always matter to other humans, why they re-configure one another's phenomenological landscapes and why they reify one another's existences. To say that humans are ontologically connected is to say that they belong to and live through the same normative world. For an agent to be considered 'human',<sup>[15]</sup> it is not enough to merely look and sound human – a being must also engage with our commonly shared networks of social norms. With this idea, 'human' and 'humanity' must be considered as two sides of a single ontological phenomenon: one cannot be human without constitutively partaking in the normativity of humanity, which is generated and maintained by various human collectives (again, there will be much more on this in chapters 5-7). 'Human-ness', from an existential perspective, is thus a question of enacting collectively shared norms.

Within the phenomenological tradition, the notion of being connected to one another in a fundamental manner is well established, although few phenomenologists explicate the idea in terms of social norms. In general, the idea is put forward as an existential relation to the world, rather than through the kinds of physical encounters that I have discussed thus far. So, for Husserl (1973), each of us is "a member of a we-community in the broadest sense — a community that has its tradition and that for its part is connected in a novel manner with the generative subjects, the closest and the most distant ancestors" (p. 223). In this way, we each belong to customs and traditions that are constituted by others; we are fundamentally party to a social world (ibid., 1952/1989, 1973). Heidegger, too, recognises that the world of a human agent is not private, but social (1927/1962). Indeed, as far as Heidegger is concerned, all worldly entities refer essentially to others because (a). to be a human agent is to always be *in-the-world* (ibid.; see also the previous chapter), and (b). the 'world' in question is not one of brute physicality, but one of social mores and meaning (Heidegger, 1927/1962). In Merleau-Ponty's works, we find the complementary claim that the phenomenal field of human agency is always a conjoint system of "self-other-things" (1945/2012, p. 56), and the likes of Scheler (1913/2008) and Schütz (1932/1967) describe (respectively) how we are empathetically connected to one another and how we are disposed to entering into shared contexts of 'we-relationships'. We have also already seen how Henry claims that we are always affectively connected to others such that "only in community is the Self possible" (2003, p. 159). Naturally, there are differences between each of these accounts and, in several cases, outright criticisms by one author of another. However, Zahavi (2001) convenes the broad idea that they all espouse as follows:

---

<sup>[15]</sup> 'Human', here (and unless stated otherwise), does not strictly align with the biological definition of *homo sapien*. Rather, it is an existential notion of 'human', such that to be a human agent is to partake in the lived world of humanity.

the phenomenologists would argue that[...] it is not possible simply to insert intersubjectivity somewhere within an already established ontology; rather, the three regions ‘self’, ‘others’, and ‘world’ belong together; they reciprocally illuminate one another, and can only be understood in their interconnection. (p. 152)

We thus find that in the same way that an agent is always existentially *in-the-world* – as suggested in the previous chapter through analysis of our ability to expertly and ‘transparently’ cope with worldly entities – she is also always *in-the-world-with-others*. The manner in which we encounter others in a distinct way from non-human entities is therefore a consequence of others’ foundational role in our existence.

Within these various accounts of how humans share a fundamental connection to one another, I am going to briefly focus on Sartre’s (1943/1984) account of shame as a means of revealing our ontological unity.

Sartre describes a situation in which an agent makes “an awkward or vulgar gesture” (Sartre, 1943/1984, p. 221). Initially, this gesture is simply ‘lived’ by the agent; it is neither blamed nor judged for its vulgarity or awkwardness. However, the agent then realises that another is present and has observed the gesture, and the agent consequently feels ashamed (ibid.). The other transforms the agent’s present phenomenological experience of the world; namely, the agent’s experience of the gesture that he has just performed. Without the other’s spectatorial presence, the experience of shame simply would not arise. As Sartre explains, “[Shame] is in its primary structure shame *before somebody*” (ibid.), and the agent is ashamed of himself *as he appears* to the other (ibid., p. 222). Zahavi (2014) emphasises this with the claim that shame “presupposes the intervention of the other[...] because that of which I am ashamed is only constituted in and through my encounter with the other” (p. 213). It is only through our fundamental connection to others that shame can be induced, transforming our phenomenological world in terms of how possibilities for action manifest themselves. More than this, however, Zahavi’s claim also emphasises the transformation of how one presently relates to oneself; that is, the shame is *of oneself* and such a self-relation is only manifest through the other’s observance of the “awkward or vulgar gesture”. As one feels ashamed of one’s gesture, one’s relation to oneself is transformed to “this degraded, fixed, and dependent being which I am for the Other” (Sartre, 1943/1984, p. 288). In this way, one’s existence is emphatically mediated by the other when in the throes of shame (ibid.).

Shame, then, exemplifies our ontological connection to others – our affective disposition to the autonomous influence of another’s presence. Referring back to the ideas of sections 2-4, we can see that shame nicely exemplifies the phenomenological transformations that were elucidated. For example, the mediating presence of others that shame encapsulates illustrates the relevance that others import to a situation and the manner in which they can re-configure the phenomenological landscape. Perhaps more interestingly, it also illustrates the reification

of oneself that I outlined in section 4. For example, as one feels shame before another, the world shifts from a resolutely egocentric point of view – one is no longer merely a transparently living agent, but also a perceivable and appraisable agent (Sartre, 1943/1984; Zahavi, 2014). The other transports a given agent to a ‘shared’ world, in which every (present) agent’s perspective and ‘power’ is apprehended alongside one’s own. One’s existence is thus reified in the sense that a specific dimension of one’s selfhood – a fundamentally intersubjective dimension – is revealed. As Nietzsche (1881/1997) dramatically puts it:

The feeling ‘I am the mid-point of the world!’ arises very strongly if one is suddenly overcome with shame; one then stands there as though confused in the midst of a surging sea and feels dazzled as though by a great eye which gazes upon us and through us from all sides. (p. 166)

When ashamed, one is thus no longer simply engaged in one’s own projects from a strictly egocentric perspective; instead, one is thrust into the ‘surging sea’ of interpersonal activity, exposed and susceptible to the behaviour of others. In terms of Gallagher’s sense of agency and sense of ownership, Nietzsche’s words seem to capture the notion that before others one’s agency is shot through with “gazes upon us and through us from all sides”, thereby modulating the egocentricity of one’s agency, whilst simultaneously crystallising one’s sense of ownership as a bodily, perceivable being.

A particularly interesting application of shame, which complements the idea that it reveals our ontological connectedness to one another, is Honneth’s (2001) view that shame can arise through the inattention of another. For example, one may feel ashamed if a friend mistakenly neglects to greet you at a party, or, worse, you are simply overlooked because your presence is deemed insufficient or irrelevant to some situation (*ibid.*). In such ways, shame does not have to be induced by another’s perceiving presence, but also by another’s capacity to intentionally ignore their perception of a given agent. For Honneth (2001, pp. 114-116), such non-perception can render an agent ‘socially invisible’, therein disturbing the agent’s typical self-relation of *being-an-other*. That is, the full-blown recognition of oneself as an affectively autonomous agent is in some sense denied, potentially resulting in psychological disorders of selfhood (see Lynch (1979), Hendin (1982), Baumeister and Leary (2000), and Baumeister, DeWall, Ciarocco and Twenge (2005)). The suggestion here is that shame not only highlights how others mediate an agent’s existence, but also how others’ wilful manipulation of this mediation can disrupt an agent’s selfhood.

Honneth’s view moves the analysis of shame from one of implicit recognition of/by others to intentional recognition (or the lack thereof). In this sense it strays slightly from our current ontological concerns. Nonetheless, it adds weight to the fundamental importance of our connection to others. Indeed, the key idea from the preceding elucidation of shame is that it highlights our ontological connectedness to one another – our ability to transform, reify and

reveal aspects of existence for one another. The phenomenological transformations that were elucidated in sections 2-4 are thus encapsulated by the implicit acceptance of another *as another*. Without this connection at the base of human ontology, there is no clearly discerned framework within which the phenomenological transformations that another imports can unfold. We are uniquely sensitive to the encountering of others because our existential relationship to others is unique; indeed, others are an integral part of our agential ontology.

## **6. Conclusion**

In this chapter I have taken steps towards elucidating the phenomenological nature of our encounters with others. This is obviously an incredibly rich domain of inquiry and I have thus confined my account to some non-reflective agential transformations that take place during the incipient moments of encounters. The outcome that I have reached is that humans share a fundamental ontological connection with other humans (section 5), which results in encounters between humans being distinguishable from human encounters with non-human entities. Due to this common ontological nature, human-human encounters entail phenomenological manifestations of relevance (section 2), re-configuration (section 3) and reification (section 4).

Building on this shared ontological foundation, I will turn my attention in the subsequent chapters (3 and 4) to the field of social cognition. More specifically, I will address the fact that even radical theories of social cognition fail to appreciate the scope and influence of our fundamental connection to one another.



# **Chapter 3 – Interactive** **Social Cognition: the** **Cognitive We-mode**



## Interactive Social Cognition: the Cognitive We-Mode

### **1. Introduction**

Over the course of the remaining chapters, I aim to build on the Heideggerian turn in cognitive science (chapter 1) and the unique phenomenology associated with encountering other persons (chapter 2) by delivering a new construal of human sociality. On the one hand, I aim to capture the nature of our bodily generated (en)socialisation, which makes us truly 'human' (chapters 5 and 7). On the other hand, I will consider the philosophy of social cognition, capturing the key processes that are in place when we understand other persons (chapter 6). This picture of social cognition will focus on the co-presence of other persons, whilst in no way relegating the importance of our collective belonging to permeating social structures and institutions. Although this double-edged picture of sociality is primarily approached through phenomenology and emphasises the importance of the social normative structures and cultural institutions in which we are entrenched, I do not believe the resultant view should be restricted to transcendental discussions, but should be viewed as contributing to and illuminating psychological investigations into life and mind.

In order to make this picture of sociality as robust as possible, I intend to spend considerable time elucidating why, despite considerable progression, social cognitive theories still fall short of being sufficiently 'social'. This elucidation will lay the groundwork for the benefits that can be reaped by taking a new perspective on social cognition.

In this chapter, I will look at Gallotti and Frith's (2013a, 2013b) theory of *we-mode cognition* and will discuss its strengths and weaknesses as an explanans of how we come to understand other persons. Particular attention will be paid to Gallotti & Frith's recent debate with advocates of a competing theory, *participatory sense-making* (De Jaegher and Di Paolo, 2007; Di Paolo and De Jaegher, 2012; Di Paolo, De Jaegher and Gallagher, 2013). The subsequent chapter will then look at participatory sense-making in detail, including the difficulties that it runs into by centralising the concept of *autonomous interactions*. By drawing out the flaws in these interactive social cognitive theories, I will be establishing the need for a more expansive view of sociality and a more precise framework for social cognition. Chapters 5-7 will put forward positive accounts of human sociality, both fully (en)socialising cognition and drawing out the novel processes that unfold when we cognise with other persons.

### **2. The Interactive Turn in Social Cognition**

In order to fully appreciate Gallotti & Frith's (2013a, 2013b) theory of *we-mode cognition*, it will be necessary to very briefly look at the traditional scope of theories of social cognition. This is useful in two respects. Firstly, *we-mode cognition* is proposed as an alternative to 'classical' theories of social cognition. Secondly, the theory of *participatory sense-making*,

which the subsequent chapter deals with, is also proposed as an alternative to classical theories of social cognition. A brief introduction to the traditional field of social cognition thus contextualises the upcoming expositions and their accompanying critiques.

So-called ‘classical’ theories and commentaries regarding social cognition – commonly referred to as ‘theories of mind’ (henceforth ToM) or ‘mindreading theories’ – tend to focus on the attribution of mental states to others as a theoretical or simulative skill. The two dominant ToMs are ‘Theory Theory’ (TT) and ‘Simulation Theory’ (ST), which can be outlined as follows:

- TT claims that person X understands person Y’s behaviour by attributing mental states to Y courtesy of a ‘folk psychological’ theory of human behaviour. So the capacity of X to predict Y’s behaviour “is essentially an exercise in theoretical reasoning” (Ravenscroft, 1997/2016).
- ST claims that person X attributing mental states to person Y simply requires person X to put herself in the “mental shoes” (Goldman, 2005, p.80) of Y. So the capacity of X to predict Y’s behaviour is an exercise in cognitive simulation.

Both theories have in common the idea that attempts to understand others involve an agent having to predict behaviour from a detached, *observational* viewpoint. So, for example, if someone reaches for a cup of water, I observe their actions and then theoretically or simulatively postulate what their mental reasons are for reaching for the cup. Theorising about or simulating another’s actions is thus a two-stage process in which one indirectly makes predictions about the cognitive reasons ‘behind’ the other’s activity. In *some* scenarios, it seems fair to suppose that postulation of this theoretical/simulative ilk is indeed taking place. Others’ behaviour can be puzzling and highly complex, so theorising about motives or putting oneself in another’s shoes emerge as reasonable cognitive strategies. Yet there is an underlying assumption to both TT and ST that sits uneasily with both our intuitions and the kind of phenomenological insights that we have looked at over the previous two chapters. This assumption is that others’ beliefs, desires and intentional states are confined to neural machinery, and thereby concealed from theorising or simulating agents. In other words, our knowledge of others is not direct because we cannot directly access others’ minds. The minds of others are “hidden away, closed in, behind the overt behavior that we can see” (Gallagher, 2005, p.209). There are, here, heavy overtones of the orthodox ‘Cartesian cognitive science’ that was discussed in chapter 1. Rather than being in-the-world agents who are phenomenologically constituted by historically entrenched subjectivity and worldly surroundings, ToMs revert to the notion of agents as isolated beings, separate from the world and thus separate from one another. As isolated beings, the starting point for understanding others must be one of detachment, where the other in question is a puzzle to be solved – a physical manifestation of internally hidden mental states. Such a

perspective trades on the notion that the bodily behaviour of others that one observes is brute physical movement being driven by mental states which remain enclosed inside the other (that is, enclosed neurally). Not only does this assume that agential cognition is 'hidden', it also suggests that it is de-contextualised. Yet this seems counter-intuitive. Think, for example, of an ordinary conversation with a friend. To suggest that one's primary understanding of the friend is through simulating or theorising about the mental states hidden 'behind' the friend's words and movements is a jarring stance to take, implying that one's understanding is somehow removed (i.e. de-contextualised) from the interaction itself. During conversation, one simply isn't observing the other from a detached perspective and inferring contextual parameters; one is *in* the interaction, contributing to the other's thoughts and privy to the subtleties of coordinated activity that emerge between oneself and the other, as well as being familiar with the contextual history of one's relationship with the other. One's access to the other must, to some degree, be direct because one is an integral part of the other's cognitive activity, which is not completely hidden away as per Cartesian cognitive science, but is subject to contextual embeddedness. Even if a devotee of TT and ST were to claim that they could theoretically account for the subtleties of coordinated activity that occur *within* an interaction, any theoretical explanation of contextual nuance that relies on de-contextual subjects as its foundation is, at best, going to be in a precarious position. Gallagher sums up these difficulties as follows:

Both theory theory and simulation theory conceive of communicative interaction between two people as a process that takes place between two Cartesian minds. It assumes that one's understanding involves a retreat into a realm of *theoria* or *simulacra*, into a set of internal mental operations.[...] If, in contrast, we think of communicative interaction as being accomplished in the very action of communication, in the expressive movement of speech, gesture, and the interaction itself, then the idea that the understanding of another person involves an attempt to theorize about an unseen belief, or to 'mind-read', is problematic. (p. 212)

In short, ToMs are grounded in the idea that any given subject is an external observer of others who has *no direct access* to others' minds and must therefore infer or 'mindread' another's mental processes from the behaviour that the other person exhibits. One issue with this approach is that it excludes beings without appropriate theoretical or simulative skills (such as preverbal children or autistic persons) from engaging in the inferential 'mindreading' that is alleged to encapsulate social cognition, as mindreading is considered a 'high-level' mental process. It seems clear, however, that such beings are still able to interact in a meaningfully social way. Secondly, and more relevant to the next two chapters, the focus of mindreading theories on de-contextualised minds engaged in a back-and-forth sequential process of inference seems to largely ignore the dynamic nature of our social interactions and

the unique relations that emerge *between* interacting agents (Fantasia, De Jaegher and Fasulo, 2014).

In response to such traditional approaches to social cognition, there has been a welcome turn towards the *interactive* nature of social cognition, in which the traditionally individualistic and inferential view of social cognition is replaced by an interest in the idea that thinking *about* or *with* others is part of a relational process that involves some sort of psychological togetherness, or the coordination of communicatory intentional activity. Broadly speaking, interactive approaches claim that social cognition involves a non-abstract connection to at least one ‘other’; the claim that follows is generally along the lines of explaining that this non-abstract connection eradicates the need for detached inferential processing of the other’s behaviour (e.g. De Jaegher, 2011; Gallagher, 2008; Wiltshire, Lobato, McConnell and Fiore, 2015; Zahavi, 2011). The emphasis is thus put on the connection between individuals, rather than looking at individuals in isolation from one another. Some of these interactive approaches propose a methodological paradigm shift for the discipline of social cognition, whilst others focus on more specific aspects of social understanding. Two prominent interactive theories are the representational theory of *we-mode cognition* (Gallotti and Frith, 2013a, 2013b), which seeks to expand our understanding of the cognition involved in joint action, and the enactive theory of *participatory sense-making* (De Jaegher and Di Paolo, 2007, 2008; Di Paolo and De Jaegher, 2012; Di Paolo, De Jaegher and Gallagher, 2013), which takes a more revolutionary approach to social cognition in general. Both of these theories claim to move away from psychological individualism (to varying degrees) – that is, the view that psychological states are ontologically confined to individual agents in isolation from their social and physical surroundings – and towards a picture of social cognition that emphasises the relational ‘we’ character of thinking and acting with others. However, they disagree on the manifestation, dynamics, functioning and explanatory power of such interactive thinking and acting.

I will now briefly outline Gallotti & Frith’s theory of we-mode cognition (section 3), the enactive response to it (section 4) and its eventual downfall (section 5).

### **3. The Cognitive We-Mode**

The cognitive we-mode emerges from the view that shared intentionality – the capacity to share the mental states of others – is a unique human characteristic that explains our sociality (Gallotti & Frith, 2013a).<sup>[16]</sup> Although we are given no detailed definition of this capacity, we can take the sharing of a mental state to occur when two (or more) individuals

---

<sup>[16]</sup> For Gallotti and Frith (2013a), shared intentionality entails the sharing of mental states, and shared mental states permit reference to the presence of ‘shared minds’. To speak of minds as ‘shared’, on their view, is thus to consider the intentional activity of agents as directed at some action that will be tackled collectively (as a ‘we’), with each agent accounting for the other’s perspective and possibilities for action along with her own (ibid.). Accordingly, I will refer to shared intentionality, shared mental states and shared minds as entailing one another.

share a common goal (qua intention) and have coordinated action roles if this goal is to be pursued together. Rather than pursuing the shared goal from a strictly individual perspective, the mental content of each involved individual will include the intention(s) of the other participant, such that an 'I' is committed to a plan that a 'we' is executing. For present purposes, this 'we-ness' should not be taken to implicate a division of mental content between participants in action, such that the overall goal of some cognitive task is distributed across participants, with each tackling a specific part. It should be taken to suggest, however, that the participants intend some joint action *collectively* (in a *we-mode*), such that a goal is achieved through non-aggregative mutual collaboration, rather than each participant individually intending their own contribution to a joint action. When two subjects intend to jointly pursue some action together, the claim is that they enter into a mode of *co-representation*, so that aspects of the world are 'shared'; that is, each subject takes into account the other's perspective and potential for joint action along with one's own, in a manner that would not arise if each subject were to individually represent only their own contribution. Another way of putting this is that each individual's cognitive state must realise a 'we-intention', in which the individuals collectively *will* the shared activity and therein distinguish their action-with-others from their individual actions (Roth, 2010). As we shall come to see, proponents of we-mode cognition contend that sharing mental states through *co-representation* amounts to positing mental states of a uniquely collective variety.

The central claim for Gallotti & Frith (2013a) is that the sharing of mental states "can take place in an irreducibly collective mode called the 'we-mode'" (p. 160), which "captures the viewpoint of individuals engaged in social interactions and thereby expands each individual's potential for social understanding and action" (ibid.). Or, as they put it later on:

the we-mode [...]captures the role of interaction for expanding the social-cognitive resources of individuals. Our proposal is that individuals engaged in joint action have a broader understanding of the behaviour of their partners, and thus of the options available for action, by representing aspects of the interactive scene in the we-mode. (ibid., p. 161)

Their assertion is thus that mental states are not invariably confined to an individualistic perspective, but that individual minds can in fact be enriched by involvement in the interactive domain of collective behaviour.

To be more specific, Gallotti & Frith argue that shared mental states can be formed and sustained through individuals adopting a "first-person plural perspective" (ibid., p. 160) of the interactive scene. Such a perspective requires each individual mind to be enabled by another, so that there is an "irreducibly collective" (ibid.) quality to the cognitive we-mode – a quality that will be scrutinised throughout section 4. For an explanation of how the first-person plural cognitive mechanism functions in a satisfactory manner, we are provided with

a description of “what a we-mode process is and how it works” (ibid., p. 163). This description rests on the notion of individuals representing their role in some joint action as collaborative “contributions to something that they are going to pursue together, as a ‘we’” (ibid.). In so doing, each interacting subject enables the other to represent “aspects of the interactive scene in *a distinct psychological attitude* of intending-together, believing-together, desiring-together” (ibid.; my italics). This psychological attitude that drives acting together is thus reciprocally enabled and held by the involved individuals, with it being unique to their interactive domain.

To exemplify this ‘we-ness’, which involves co-representing aspects of the interactive scene, Gallotti & Frith provide the following illustration:

imagine that Mary and John come across a friend in difficulty and offer help. It seems plausible that whatever they mean to do in order to bring about joint assistance, they do it together. In other words, it is because Mary and John see each other as being part of the same ‘group’ that Mary understands John doing his part, and herself doing hers, as contributions to something that they are doing together, rather than just as the result of individual tasks undertaken simultaneously. (2013a, p. 162)

Here, we find that each of the subjects – Mary and John – are considering the interactive scene from a we-perspective, representing what their joint contributions to action can be, and therein enhancing their individual understanding of the social situation. If they were not able to see one another “as being part of the same ‘group’” then Mary and John simply would not find themselves disposed towards thinking about their situation in the cognitive we-mode. Its manifestation relies on their engaging together in joint action.

To consolidate this first-person-yet-shared approach, Gallotti & Frith give a further example that aims to elucidate the concept of subjects ‘co-representing’ the potential actions of their interacting partners. In this example, a subject – Mr. Blue – is sat at a square table which has four mugs on it. Each of these mugs has an ‘affordance salience’ according to how available they each are for action, which in this case we can assume means how comfortably grasp-able they are. If a potential collaborator – Mr. Red – is to sit on a different side of the table then he will elicit co-representations “by inducing Mr. Blue to take into account the perspective of Mr. Red” (ibid., p. 163). This ‘I-and-another’ co-perspective will then change the affordance salience of the mugs accordingly. So, for example, the affordance salience (i.e. the ‘grasp-ability’) of a mug that is on the opposite side of the table to Mr. Blue will be increased (i.e. it will be encountered as ‘more graspable’) once Mr. Red’s perspective is taken into account. The potentially collaborative and interactive presence of Mr. Red results in Mr. Blue no longer thinking purely as an ‘I’; instead, he is poised to think and act in the we-mode by co-representing a situation – such as the grasping of a mug on the opposite side of the table – that may be prone to full-blown “reasoning-as-a-we” (ibid., p. 164). This example

relies on the *collaboration* of Mr. Red, or at least the genuine potential for it. Once, however, this collaboration is taken for granted, Mr. Blue and Mr. Red are poised to think of their actions as playing a role in “*something that they are doing together*” (ibid.; my italics) so that the scene of action is interactively framed by the we-mode process of co-representing the other’s perspective.

This example is noteworthy for clarifying the notion of co-representation. However, as we shall see in section 5, it also highlights two pressing problems for the theory of we-mode cognition: firstly, it is unclear exactly how Mr. Red and Mr. Blue achieve further we-mode processes (and what these processes are), besides co-representation, that Gallotti & Frith (2013a) contend are required for full-blown we-mode cognition; secondly, it is unclear why Mr. Red’s involvement differs from that of a non-human entity (thus justifying we-mode cognition as a uniquely *social* activity).

What we can assert at this stage is that with the manifestation of co-representations by those engaged in joint action (howsoever this is precisely achieved), Gallotti & Frith claim that each subject is no longer cognitively limited by their restrictive individual perspective, but has a collective notion of what can be done together *as a ‘we’*. As such, “cognition is enriched” (ibid.) so that novel possibilities for action become available to the ‘we’ – possibilities that would be unavailable to each individual on their own. In the above example, the collaborative introduction of Mr. Red opens up new possibilities for action to Mr. Blue, as long as Mr. Blue sees the two of them as (at least potentially) jointly engaged in some task (which in this case is the picking up of a mug). Similarly, in the previous example, Mary and John individually would have been unable to offer help to their friend in the same way that they offer it together because without one another, they simply would not be able to think about their possibilities for action in the same way (i.e. in the cognitive we-mode).

With the cognitive we-mode, we are thus provided with a theory of cognising with others that requires social interaction to enable and incite specific psychological capacities and mechanisms (those that manifest co-representations of the interactive scene). When qualifying this theory of social cognition, Gallotti & Frith (2013b) give us the following conclusion:

The *we-mode* comprises no more than a set of psychological propensities and dispositions, described in terms of one’s taking into account the perspective of others, which remain latent until individuals engage in interaction. (p. 304)

As we have seen, these propensities and dispositions can enhance one’s isolated, individual perspective so that new possibilities for thought and action become available to subjects engaged in some interaction as a ‘we’. Gallotti & Frith also claim that this engagement in interaction provides a route to explaining how subjects can have their understanding of one

another and their situation enhanced without relying on theoretical or simulative ascriptions of 'hidden' cognitive states (as is generally considered necessary for classical theories of social cognition (see earlier)). The we-mode facilitates a kind of attachment to another person, with the co-representation of this other's viewpoint resulting in "a 'meeting' of minds" (Gallotti & Frith, 2013a, p. 164), instead of solitary minds that have to rely on the back-and-forth interplay of individualistic inferences. When engaged in some action together, subjects are thus able to think differently by taking into account one another's perspective and therefore moving away from cognition as an isolated 'I'.

#### **4. The Enactive Response to the Cognitive We-mode**

##### **4.1. Individual or Collective?**

The cognitive we-mode undoubtedly has numerous positive attributes. It should be commended for addressing the interactive aspect of social cognition and for attempting to capture and explain the phenomenon of subjects in groups thinking differently to isolated individuals. However, there are some immediate issues to consider.

First and foremost is the issue of coherence. Although I have avoided discussing it thus far, one of the central features of the cognitive we-mode for Gallotti & Frith is their claim that it is a non-reductionist account of social cognition. The motivation behind this claim is that "spectatorial" individualism, in which an agent must observationally theorise about the intentional behaviour of another, is deemed an "unsatisfactory" approach for social psychological research (Gallotti & Frith, 2013a, p. 162). Gallotti & Frith define individualism as the theory "that all sorts of complex behaviours entail properties of the individual as distinct from other levels of functional organization" (ibid., p. 160; Glossary). They contend that any individualistic reliance on the internalised cognition of *isolated observers* is insufficient for explaining the unique cognitive mechanisms that seem to emerge when subjects are engaged in some forms of group or interactive behaviour. Such insufficiency arises from the fact that most individualistic accounts of social cognition tend to consider interacting agents to have "the perspective of an observer *qua* theorist, who represents the decision problem faced by the [other] agents as it appears to him" (ibid., p. 162). In other words, individualistic theories generally suggest that mental states belong exclusively to individuals who can, at best, make third-person theoretical or simulative predictions regarding the internal cognition that motivates the external behaviour of others. Gallotti & Frith consequently deliver we-mode cognition as a non-reductionist alternative to such observational individualistic theories, in that individuals who are poised to interact with others in some collective behaviour actually unveil "novel routes to knowledge of other minds" (ibid.) by cognising from a first-person plural, *we*-perspective, which is irreducible to the first-person singular perspective of non-we-mode (or 'I-mode') cognition. Once in the

we-mode, with its unique we-perspective, individuals are no longer mere observers or third-person parties to some event (and some other individual(s)), but are *participants* in a genuinely collective act.

It is therefore not only the fact that engaging with others in certain joint actions can lead to an enriched understanding of said others that is important for Gallotti & Frith; there is also the fact that the theoretical framework of this enriched understanding is allegedly non-reductionist. Yet the cognitive we-mode is also meant to be “consistent with individualism” (ibid., p. 163). Such consistency with individualism is supposed to rest on the idea that the psychology of an individual who is cognising in the we-mode can be studied and understood by investigating this individual’s capacities and internal mechanisms. Simultaneously, the non-reductionist nature of the theory is said to persist in virtue of requiring social interaction to enable these *distinct* psychological capacities and mechanisms.

#### 4.2. A Subtle Balance or Paradoxical Nonsense?

Gallotti & Frith’s attempted balance of a theory that is non-reductionist yet consistent with individualism is summed up by the problematic description of ‘first-person plural’ cognition. As Di Paolo et al. (2013) point out, there is something intuitively perplexing about the psychology of we-mode cognition being “paradoxically[...] both irreducibly collective and belonging to the individual” (p. 303). If the functionality of a psychological mode can be explained through the representations and mechanisms within each involved subject, then surely the psychological attitudes in question are reducible to the individuals? Indeed, for Di Paolo et al. (2013) the only reasonable explanation is that “the we-mode is definitely not collective (much less irreducibly so)” and, at best, it can carry “trans-individual content, like most intentional attitudes, social or non-social” (p. 303). It must be assumed here that Di Paolo et al. are using ‘trans-individual content’ to convey the straightforward idea that some aspect of the world can be contemplated by two individuals simultaneously. So, for example, if two individuals are looking at the same tree, they could both agree that its leaves are green, resulting in their mental content (regarding the greenness of the leaves) being alike. The relevant mental content would hence be *trans*-individual (i.e. consistent across the two individuals). Di Paolo et al.’s criticism is thus presumably that we-mode cognition simply describes two disparate individuals attending to the same aspect of the world (in virtue of an action plan), with their resultant mental states being trans-individual. This trans-individuality merely indicates individual sameness, rather than genuine collectiveness.

In responding to Di Paolo et al., Gallotti & Frith (2013b) aim to strengthen their stance by putting forward the claim that

it takes two *individuals* to tango, so whatever is involved in group behaviour must include, amongst other things, individuals with mental states that are causally connected with their actions in ways that enable them to act together – as a group. (p. 304)

The causal connectivity between individuals' mental states and their (joint) actions is intended to suggest that we-mode cognition involves more than just trans-individual content. It would seem that the fact that the individuals are acting *together* holds the key for Gallotti & Frith: in so doing, the individuals are not mere passive observers on the outside of the unfolding event; they are *in* the interaction, collaboratively deciding how the event will unfold from the inside.<sup>[17]</sup> Rather than abstracting themselves from the situation as isolated observers, individuals in the cognitive we-mode are seen to take on participatory roles towards some activity that they intend to execute together. This active togetherness is what purportedly gives we-mode cognition its irreducibly collective character; the novel we-mode mechanisms simply wouldn't come into play without the presence of one individual interacting (or at least being seemingly poised to interact) with another. We-mode cognition is, on Gallotti & Frith's view, thus undoubtedly "a claim about properties of individual minds being truly collective in that they need group environments to work" (*ibid.*, p. 305). It is this enabling role of the individuals involved in some joint action, and the resultant novelty of the psychological states that are enabled, that is deemed sufficient for endowing the psychological states with a 'truly collective' character. The theory thus purports to provide a subtle balance between individual psychology and its social conditioning.

For the likes of Di Paolo et al., on the other hand, it is simply paradoxical to posit an 'irreducibly collective' quality to individual psychological states – the fact that the individuals are supposed to be engaged in acting together does not satisfactorily explain how their mental states become 'irreducibly collective'. For instance, even if one were to accept that the notion of causally connected mental states moves we-mode cognition beyond the simple concept of trans-individual mental content, it is unlikely that a participatory sense-making supporter would accept this as a complete story (see section 4.3 and, in particular, the next chapter for an exposition of participatory sense-making). It remains to be shown how the *dynamics* of persons (inter)acting together leads to a genuine progression from individualism. That is, how are individual (we-mode) psychological states "novel" and "distinct" from individual (I-mode) psychological states? Whilst the theory undoubtedly

---

<sup>[17]</sup> Being *in* an interaction is also the reason why Gallotti & Frith believe the cognitive we-mode is incompatible with classical theories of mindreading. They believe that mindreading theories are "the outcome of cognitive processing that occurs in an individual's mind in abstraction from, and as a precondition for, interaction with others" (2013a, p. 160). In comparison, we-mode cognition relies on the ongoing interactive involvement of another individual, such that holding a we-mode perspective is a *relational process* of cognitive enabling.

It is worth noting, however, that as long as the cognitive we-mode endorses *individual* (co-)representations, then it seems to require some kind of 'abstraction' in joint action in order to maintain consistency with individualism (see section 4.3). What's more, there is no reason that TT and ST could not employ the same kind of abstraction, with individuals becoming *actively* involved in some joint activity and enabling distinct psychological dispositions within one another.

emphasises the significance of social context and its ability to induce cognitive processes that are unavailable to isolated subjects, the story of how it is non-reductionist remains unclear from the viewpoint of Di Paolo et al., as co-representations are still individualistically confined to non-group, subjective minds.

#### 4.3. We-Mode Cognition Stands its Ground...by Shifting the Goalposts

It is possible that this confusing friction between Gallotti & Frith and Di Paolo et al. may be better understood at a more fundamental theoretical level. In order to see why this is the case, I will firstly attempt to add clarity to the cognitive we-mode's framework by drawing a comparison with Robert Wilson's (2004, 2005) *social manifestation thesis*.

Wilson introduced this thesis as a development of a line of thinking within the 'collective psychology' literature, which he believes bifurcates in a notable manner:

There are two very different claims about cognition that are made in the collective psychology tradition [...] The first is that groups can have or can be thought of as having minds in something like the sense in which individuals can have minds. [...] The second] is the idea that individuals engage in some forms of cognition only insofar as they constitute part of a social group. (p. 3)

At first glance, it should be apparent that we-mode cognition is steadfastly aligned with Wilson's (2005) interpretation of the second claim made within the collective psychology tradition. This second claim embraces his *social manifestation thesis* and in many ways, as Gallotti & Frith (2013b) eventually acknowledge, we-mode cognition is in fact a version of Wilson's (2004, 2005) *social manifestation thesis*. The central claim, in Wilson's (2004) words, is "that individuals have properties, including psychological properties, that are manifest *only when those individuals form part of a group of a certain type*" (p. 281; my italics). Rather than positing group-level properties that are "emergent from and thus not reducible to the psychology of the individuals in [the group]" (ibid., p. 268), the social manifestation thesis contends that the cognitive properties of interest belong to those individuals within a group, not the group itself. To exemplify this approach we can consider the pop-psychological notion of 'mob mentality', in which "persons who act as part of a group get "caught up" in the excitement of the mob" and so do not behave as they would if they were acting alone (Harvard Law Review, 1995, p. 1111). History and daily life are rife with occurrences of this 'mob mentality', from the extreme fervour of events such as the Salem Witch Trials, to the weekly frenzy of fans in football stadia. What connects these events is the almost hypnotic effect that being in a social collective can have on otherwise normal persons, often making them behave in a radically out-of-character manner. Mob mentality, although an ill-defined phenomenon, thus captures Wilson's (2004) view that "individuals have their psychology transformed through social membership" (p. 299). Yet this social membership

does not result in a supra-individual or group cognitive system in which the cognitive properties of interest would belong to the group itself. The phenomenon of interest for the social manifestation thesis is that the occurrence of some sorts of social membership is sufficient to trigger the manifestation of individual psychological dispositions that would otherwise not arise, or even allow the individuals to acquire new and distinct psychological dispositions (ibid.). Simply by participating in the activities of some social collective – or, in accordance with the particulars of we-mode cognition, the *joint action* of some social collective – we find that individual (I-mode) psychology can be radically altered. Whilst the triggering of these psychological dispositions is a social event (in that it necessarily requires the involvement of others), the theory remains “continuous with the individualistic tradition of thinking about psychological states” (ibid., p. 300) because the psychology of any individual within a group “can (and should) still be understood in abstraction from that individual’s social environment” (ibid.).

‘Abstraction’, here, needs to play a subtle role, in that it entails an individual’s detachment from supervenient group cognition, but does not entail detachment in an isolated, observational manner, as the individual has to be appropriately involved in some social collective. It consolidates the idea of socially manifest psychology being an individual (albeit socially enabled) capacity, whilst leaving open the possibility that such psychology differs from disengaged, spectatorial psychology (in virtue of its social manifestation). That is to say, the cognitive content that emerges from socially-manifest/we-mode scenarios is indeed mechanistically individual, but such individual mechanisms must be socially activated, thereby engendering ‘enriched’ content when compared to the isolated mechanisms of non-socially manifest (I-mode) cognition. In this way, other interactive persons are facilitating the emergence of a *novel* set of *individual* psychological propensities and dispositions that simply would not arise without the requisite membership in a social collective. Rather than focusing on groups themselves, this social manifestation thesis/cognitive we-mode approach relies on the causal cognitive connectivity of *individuals* in a manner that allows them to act *together*. We thus find a recurrence of the idea of irreducibly collective, yet individual cognition, except this time we have slightly more conceptual artillery. Although the presence of a social collective and their interactions are essential to the story, they simply provide the subplot to the central narrative regarding individuals. As Wilson (2004) says, it is a “middle ground” that is being sought, in which “socially manifested psychological traits are properties of individuals, but since they occur only in certain group environments, they cannot be understood in purely individualistic terms” (p. 301).

The key insight that our deployment of Wilson’s analysis provides is thus that when cognising in the we-mode, the psychology of a given agent remains a set of *individual* capacities, but these capacities are necessarily enabled by involvement in some social

collective. Such involvement in a social collective can *significantly transform* the mode of individual cognition that occurs when alone, as we see through the transformations that occur in cases of ‘mob mentality’. The crucial enabling role of the social environment (in virtue of membership in a collective) therefore incorporates a *transformative* aspect by inducing psychological capacities that would not be available to an individual on his/her own. On Gallotti & Frith’s behalf (and with the aid of Wilson), the best explanation that can be given for the individual-yet-irreducibly-collective claim is thus that individual we-mode psychological states arise from a socially manifest *transformation* in individual I-mode psychological states. We-mode psychology is non-reductionist with respect to these individual I-mode states, in the sense that we-mode states fundamentally differ (in virtue of being socially enabled) from I-mode states, but the theory itself remains individualistic at heart by resisting ‘group’ explanations of cognition.

Di Paolo et al., on the other hand, endorse an underlying approach to cognition that means they fall into neither the ‘cognitive we-mode/social manifestation thesis’ branch of collective psychology, nor Wilson’s (2005) first branch in which “groups can have or be thought of as having minds” (p. 3). In order to understand why this is, we must briefly look at De Jaegher & Di Paolo’s (2007) definition of *participatory sense-making* as “the coordination of intentional activity in interaction, whereby individual sense-making processes are affected and new domains of social sense-making can be generated that were not available to each individual on her own” (p. 497). Central to this theory, which will be discussed in greater detail in the following chapter, is the idea that the coordination of intentional activity and the generation of new domains of social sense-making create a reciprocally regulated interactive process: “the agents sustain the encounter, and the encounter itself influences the agents and invests them with the role of *interactors*” (ibid., p. 492). Through this reciprocal regulation, the interactive process between individuals is able to acquire an “operationally closed organization” and thereby “enjoy[s] a temporary form of autonomy” (ibid.). In other words, the relational process of interaction itself, *between* individuals, forms a (fleeting) autonomous system and should be at the heart of any investigation into social interaction. Even with this incredibly brief description, we already have a crucial insight for current purposes: it is the *intersubjective* process of relational dynamics *between* individuals, rather than any individual taken on their own, that becomes the focus of attention for analyses of the cognition involved in social interaction. This idea follows directly from the enactive dependence on the dynamic coupling of an organism and its environment (which will be expounded in detail in the next chapter), and such intersubjective cognitive processing can thus only be emphatically non-reductionist. Instead of focusing on enriched individual psychology (as the cognitive we-mode/social manifestation thesis does), or on supervenient group psychology (in which the group itself is said to have a kind of mind that cannot be

reduced to (i.e. supervenes on) individual ones), participatory sense-making concerns itself with the dynamical interplay of neural, bodily and environmental processes within a social system. In taking this stance, it encompasses coordinated collectives of individuals, the individuals themselves, and the interplay between these collective and individual structures. With nothing other than this highly simplistic and superficial discussion of participatory sense-making, it is immediately clear that the theory is not explanatorily limited to the internal mechanisms and cognitive states of single persons. Contrary to we-mode cognition, what matters for participatory sense-making is not just the transformation or enrichment of individual cognition through membership in a social collective, but the notion that the social collective can take on a ‘life of its own’, both feeding from and feeding into the individuals that constitute it (again, this will be properly discussed in the next chapter). Whereas the psychology of individuals cognising in the we-mode is enabled by the social context, participatory sense-makers rely on a social context that is more than enabling, because the social context itself (in the form of a collective) harbours interactive dynamics with their own autonomy. These interactive dynamics draw attention to supra-individual aspects of cognition that *realise* individual psychological processes.

Herein, however, lurks a problem. Whereas we-mode cognition – as a representationalist theory – is operating within traditionally recognised bounds of collective psychology and cognitive science more generally, we have seen in the last paragraph that participatory sense-making relies on the relatively new theoretical framework of enactivism, which is emphatically anti-representationalist. For enactivists, the notion of a ‘first-person plural perspective’, at its core, suggests a failure to acknowledge the autonomy of our interactive dynamics that persist alongside our individual autonomy. But this is not a fine-grained divergence at the level of collective psychology, or even social cognition more broadly; it is instead a divergence caused by the underlying approaches to cognition. Di Paolo et al.’s (2013) criticism of the notion of a ‘first-person plural perspective’ is in fact merely contingent on their fundamental disagreement with the likes of Gallotti & Frith over what cognition is and how it works. Although Gallotti & Frith’s delineation of an “irreducibly collective” psychology that is nonetheless compatible with individualism is, at best, a delicately balanced proposal, one can construe it as coherent within the representationalist framework that they favour – as long as the irreducible collectiveness is not taken to claim that we-mode cognition cannot be reduced to individual processes scaffolded by interactions, but rather that it cannot be reduced to *the kind of* cognitive processing that is at work when individuals cognise without the social-enabling of another person, in a non-enriched, I-mode manner. Conflictingly, the ‘irreducibly collective’ claim simply verges on the nonsensical within the (anti-representationalist) enactive framework of Di Paolo et al., in that it assumes the representation of the world by a pre-given, individual mind. If cognition is the “enactment of

a world and a mind” by an organism that is dynamically coupled to its environment (Varela et al., 1991, p. 9), then the notion of genuinely ‘individual’ cognition is impedimentary to a full picture of cognition, as is the notion of ‘irreducibly collective’ cognition. What require attention instead are the autonomously organised systems across cellular, organismic and supra-individual (social) levels, and how these systems intertwine. Rather than definitive ‘individual’ or ‘group’ minds, enactivism endorses diverse and fluctuating autonomous ‘sense-making’ organisations that align with living and social systems. As long as the jury remains out on which theoretical model(s) is most appropriate for cognitive scientific studies, we-mode cognition is able to avoid the criticism of paradoxicality.

Di Paolo et al.’s (2013) initial worry of incoherence may thus be somewhat misplaced. Their positing of a form of collective psychology that expounds the interplay between individuals, groups, and their relational dynamics, does not in itself (without embarking on a more deep-rooted critique of representationalism) deny the possibility of a cognitive we-mode with a ‘first-person plural perspective’. The cognitive states that emerge in we-mode cognition are indeed individual according to the representationalist paradigm, yet their manifestation is reliant on the transformative effect of some collective participation. As long as the notion of ‘irreducible collectiveness’ is positioned correctly as an individual capacity which has a collective quality only through its novel generation relying on the collaborative involvement of another, then the theory can cling to the right side of coherence. Consequently, Gallotti & Frith (2013a), drawing on Wilson (2004, 2005), can at this early juncture find some vindication in proposing an “irreducibly collective” psychology that is nonetheless compatible with individualism, although the foregoing discussion highlights the difficulty in stabilising such a subtle stance.

## **5. We-mode Cognition in a Tangle**

The coherency of we-mode cognition is unfortunately brought back into question when one looks more closely at its manifestation. In the previous section, I noted that we-mode psychological states could be construed as “novel” in virtue of the induced psychological transformation that takes place when one is appropriately involved in a social collective. But what exactly is it about this social involvement that results in the manifestation of these novel psychological mechanisms and capacities – what exactly ‘triggers’ the cognitive we-mode into life? In particular, when we are told that the manifestation of the cognitive we-mode arises when individuals engage in joint action, should this be taken to mean that one must ‘switch’ into a ‘first-person plural perspective’ in order to satisfactorily engage in joint action, or that engaging in joint action facilitates the adoption of a first-person plural perspective? If one is to adopt a we-mode perspective in order to execute some joint action then “it remains unclear how the necessary coordination is achieved such that group

members adopt we-attitudes” (Di Paolo et al., 2013, p. 303). We are simply left uninformed as to what are the requisite conditions for two disparate individuals to become sufficiently involved with one another so that a first-person plural perspective can be taken. If, on the other hand, it is already being engaged in joint action that facilitates the manifestation of a first-person plural perspective, then it would seem that focusing on the social enrichment of individualistic capacities is a shortfall in the theory of we-mode cognition, and due credit should instead be given to the *interactive* nature of the achievement. As Di Paolo et al. (2013) contend, “[t]he situation is at best unclear” (p. 303). Although Gallotti & Frith (2013b) later try to qualify their position as not claiming that cognising in the we-mode “is enabled by deliberate and voluntary adoption of a *we*-perspective as a precondition of interaction” (p. 304), they give no further explanation of how exactly it is enabled, other than having a reliance on a group environment. Their explanation of we-mode cognition as “a set of psychological propensities and dispositions”, which are brought forth from a latent state once an individual engages in interaction and takes into account the perspective of another, is unsatisfactory in clarifying what it is about social engagement and acting in a group environment that triggers cognising in the we-mode.

This problem is compounded by Gallotti & Frith’s example involving Mr. Blue and Mr. Red. Recall that Mr. Blue’s affordance landscape, of which the relevant aspects are the affordance saliences of the mugs on the table before him, is altered by the arrival of Mr. Red, in virtue of Mr. Red’s presence inducing Mr. Blue to take a we-perspective of the scene (assuming Mr. Red’s arrival is a potentially collaborative one). In slightly more fine-grained terms, we are told that Mr. Red’s presence automatically generates “the co-representations necessary for cognizing in the we-mode” (Gallotti & Frith, 2013a, p. 163 (figure 2)). But what exactly is it about Mr. Red’s arrival that causes the elicitation of these co-representations that are necessary for the cognitive we-mode to take hold? The best answer we have thus far is that Mr. Red’s collaborative presence results in a group environment, and membership in this group environment is sufficient to transform Mr. Blue’s (and Mr. Red’s) psychology and bring his(/their) latent we-mode capacities to life. The question then becomes what is unique about this ‘group environment’? What is it, for instance, about the presence of another *person*, as opposed to a non-person entity, that allows we-mode cognition to be employed? Consider Mr. Blue encountering some sort of grasping tool instead of Mr. Red. This tool would undoubtedly change the affordance saliences of the mugs on his table, and so could its arrival not entail the manifestation of a kind of we-perspective? Without a compelling case for why socially-manifest-*yet-individual* ‘co’-representations must imply ‘with another person’, it could be argued that the tool’s arrival would result in a ‘Mr. Blue + tool’ perspective, which would be mechanistically and functionally akin to the alternate first-person plural perspective of ‘Mr. Blue + Mr. Red’. Both ‘Mr. Blue + tool’ and ‘Mr. Blue + Mr.

Red' seem to result in an enrichment of Mr. Blue's potential for action and his cognitive understanding of the situation "by providing a broader understanding of the options available for action" (Gallotti & Frith, 2013a, p. 164). There is no indication of why the 'group environment' that facilitates co-representation need be a *social* environment (i.e. 'person + person' environment), instead of a hybrid group environment involving cognitively enriched individuals and collaborative worldly entities. This issue is exacerbated by the fact that Mr. Red only need be a *potential* collaborator. That is, Mr. Red may not have even begun to consider engaging in some joint action with Mr. Blue, and yet this has no obvious bearing on Mr. Blue's individual capacity to co-represent his current situation. If it is not the relational dynamics of Mr. Red's interactive presence that matter, then we are simply left none-the-wiser as to why the collaborative presence of another person can enable the conditions that are necessary for we-mode cognising when the presence of a non-social (i.e. non-human) entity cannot.

Yet a further complexity is added by Gallotti & Frith (2013a) puzzlingly endorsing evidence that co-representation can occur even if no joint action takes place, when one is merely observing another who is in one's presence (p. 164; see also Frith (2012)). This would seem to contradict much of what they argue and, at the very least, irrevocably complicates the issue of we-mode processing being 'irreducibly collective', as well as reiterating the problem of why co-representation is seemingly limited to human social environments. Gallotti & Frith explain the evidence as follows:

notice that being in a joint action is sufficient for two agents to co-represent their perspective in the action scene, but the converse does not hold [...]if an action is to count as joint, there is more to the processes that underlie shared mental states than just co-representation (p. 164; Box 2)

So simply being involved in the 'group' that is engaged in some joint action is sufficient for co-representation, but co-representing aspects of the action scene is not sufficient for engaging in joint action. The primary issue here is that we are not told what "more" (in addition to co-representation) is required for joint action, other than an "engage[ment] in the type of reasoning-as-a-we (team) that underlies thinking and enacting things in the we-mode" (ibid.). Without an in-depth explanation of what is meant by 'engagement' – of what is unique about this kind of involvement *with another person* – this elusive "more" could simply be taken to mean joint action itself (Di Paolo et al., 2013, p. 303). Thus, "in circular fashion" (ibid.), the proposal would become uninformative with regards to joint action: co-representation and we-reasoning are required for joint action, and joint action (or, at the very least, membership in a 'group' with the potential for a common goal) is required for we-reasoning.

Moreover, the divergence in sufficiency claims regarding co-representations and joint action seems to suggest a division is being drawn between the processes that lead to we-mode cognition (such as ‘co-representation’ and ‘we-reasoning’) and the actual achievement of we-mode cognition during joint action. In principle, this could perhaps be accepted were it not for the fact that the definition of co-representation that Gallotti & Frith provide us with is of co-representing within ‘an action scene’, in which collective interaction seems probable. Thus, co-representation seems to require (the potential for) interaction. Yet if co-representation is then advocated as possibly occurring observationally, when no interaction takes place, or perhaps does not even seem likely to take place, then we seem to face a contradiction: why should it be considered a process that is tellingly connected to we-mode cognition – which occurs *during joint action* – as opposed to being a straightforward (‘I-mode’) representation of one’s role in some collective group? Put another way, if we-mode processes such as co-representations can emerge from an observational perspective, then their involvement seem at odds with we-mode cognition’s requiring individuals to represent “contributions to something that they are going to pursue together, as a ‘we’” (Gallotti & Frith, 2013a, p. 163). Separating processes that underlie we-mode cognition, such as co-representations, from the actual manifestation of we-mode cognition during joint action seems troublesome for Gallotti & Frith in the sense that it leaves us once again wondering what ‘more’ is required, besides (or implicated within) joint action itself, for the we-mode to take hold.

## **6. Conclusion**

In light of these considerations, it would seem that Di Paolo et al. are ultimately justified in putting the coherency of we-mode cognition in doubt. Central to the problems that Gallotti & Frith must face is exactly how membership in a group environment is sufficient for the manifestation of we-mode cognition. Alongside this, a more exacting mechanistic explanation of such socially enabled thinking – from initial engagement with others, through to we-mode processing (such as co-representations) and ending in transformative, mutually enriched we-mode cognition – is needed. As is hopefully evident, an underlying cause of these problems is a failure to delineate why *human* involvement is necessary for we-mode cognition. Whilst the explanations of the psychological mechanisms involved in we-mode cognition initially seem strong, there is no focus on why (other) humans are integral to the use of such mechanisms. In other words, there is no focus on why we-mode cognition is proposed as a *social* theory.

In spite of this, however, the next chapter will demonstrate that Di Paolo et al.’s alternative to we-mode cognition is also not impervious to criticism.

# **Chapter 4 – Interactive** **Social Cognition:** **Participatory Sense-** **making**



## Interactive Social Cognition: Participatory Sense-Making

### **1. An Introduction to Enactivism and Participatory Sense-making**

In the previous chapter, I considered the theory of we-mode cognition and the difficulty in sustaining its coherence under the conditions that Gallotti & Frith (2013a; 2013b) postulate. As discussed, Di Paolo et al. (2013) are notable critics of the theory and believe that they can provide an alternative explanation as to how interacting individuals come to better understand one another through their theory of participatory sense-making. Throughout sections 2 - 3.1 of this chapter, I will give a broad outline of participatory sense-making and will detail its advantages as a theory of social cognition (in comparison to we-mode cognition). I will then discuss the problems that this theory has to face (section 4).

### **2. The Central Tenets of Enactivism**

Participatory sense-making is the original and dominant enactive account of social cognition, and it has several benefits when compared with we-mode cognition. It aims to “extend[...] the enactive concept of *sense-making* into the social domain” (De Jaegher & Di Paolo, 2007, p. 485), relying on the five core, mutually supportive tenets of enactivism:

1. **autonomy**: a system is autonomous when it obeys rules that are established by its own activity, such that the government of no controlling executive is evident. Autonomy, in the enactive sense, invokes the quality of *self-organisation*, signifying that local rules and nonlinear interactions amongst systemic constituents give rise to (and maintain) a discernible global order (Wheeler, 2011b). If the activity of recursively dependent systemic constituents establishes and maintains a unifying boundary, creating a systemic unity, as well as establishing this unity’s possibilities for interaction with the environment, then a self-organising system is a fully autonomous one (Wheeler, 2011b; Varela, 1979). Exploring this idea a little further, we can say that in virtue of its self-organising constituents, an autonomous system will “actively generate and sustain an identity under precarious conditions” (De Jaegher & Di Paolo, 2007, p. 487). This (self-)identity generation provides the opportunity to elucidate two implicated qualities within enactive autonomy. Firstly, the self-generation of an identity supposes the property of *operational closure*, which is the property that any constituent process in a system is dependent on other systemic processes at the same level of description (ibid.; Varela, 1997; Di Paolo, Rohde & De Jaegher, 2010). This leads to recursive circularity amongst constituent processes, which allows the processes to define an operational unity that will regenerate itself in the space within which the recursively interdependent processes occur. Secondly, the maintenance of an identity in the face of *precariousness* implies that the mutually interdependent processes would discontinue in the absence of

the overall system's unifying organisation (Di Paolo, 2009). The constituent processes thus do not only rely on one another for generating the systemic identity to which they belong, but also for the continuation of their own activity. Without the continuation of the unity that they collectively maintain, systemic constituents will cease their systemic activity. Thus, if a system is operationally closed – generating its own identity – and maintains its operational closure (i.e. maintains itself as a unity) under precarious conditions, then it exhibits the right kind of self-organisation to be considered autonomous.

The operationally closed self-organisation of biological systems is defined as *autopoiesis*<sup>[18]</sup> (Maturana and Varela, 1980, 1987). Autopoiesis is in many ways the seed from which the enactive paradigm has grown, with Maturana & Varela (1980) proclaiming it a necessary and sufficient characteristic of living systems. However, as it is a property that is commonly restricted to the biochemical domain, present purposes require the noting of just two autopoietic qualities: (a). the boundary and identity of a (self-producing and self-maintaining) autopoietic system must be *materially* definable, so that, broadly speaking, “autopoiesis is autonomy plus materiality” (Wheeler, 2011b, p. 152), (b). autopoiesis is sufficient, but not necessary for autonomy, so there can be nonautopoietic autonomous systems (e.g. an insect colony can form an autonomous system with a social and territorial boundary, rather than a material one).

2. emergence describes any novel and distinct macrobehavioural properties that arise from (lower-level) local interactions, such as the seemingly intelligent behaviour of harvester ant colonies that emerges from the interactive behaviour of individual harvester ants. In these ant colonies, there is no central authority that controls the behaviour of the individual ants, and the individual ants themselves possess little (if any) demonstrable intelligence, yet the colony as a whole can build elaborate nests, find the shortest path to the best food source, allocate different tasks to different ants, and defend its territory from threats (Gordon, 2010). In other words, the unintelligent individual ants (inter)act in behavioural patterns that produce a colony which demonstrates intelligent behaviour at its own macro-level. This intelligence is an emergent phenomenon, in that it is a property the colony possesses that arises from the local interactions of individual ants which do not themselves possess such intelligence.

---

<sup>[18]</sup> Autopoiesis is taken to be the paradigmatic example of biological autonomy, in which a “system dynamically produces its own material boundary or membrane” (Thompson, 2007, p. 44). For example, the constituent processes of a living cell recursively depend on each other for their generation and realisation as a network, such that the cell's metabolic reactions produce the molecular components necessary to determine its membrane, and the metabolic network constitutively generated within the membrane defines “the system as a unity in the biochemical domain and determines a domain of possible interactions with the environment” (ibid.).

For a background and introduction to autopoiesis, see Maturana (1975), Varela (1979), Maturana & Varela (1980, 1987) and Thompson (2007).

Within enactive theory, the focus is on *dynamic co-emergence*, in which “a whole not only arises from its parts, but the parts also arise from the whole” (Thompson, 2007, p. 38.), so that “part and whole[...] mutually specify each other” (ibid., p. 60). This kind of emergence is clearly linked in an inextricable manner to the description of autonomy given above. That is, a system that is generated and maintained by the recursively dependent interactions amongst its constituents, and is consequently poised to interact with its environment in such a way that its constituents are themselves maintained, is an *emergent* system as well as an autonomous one. The paradigmatic example of an emergent system is given at the biochemical level of autopoiesis: in a biological cell, interacting molecules give rise to a chemical reaction network and a membrane boundary which define the cell as an operationally closed unity; this unity is then recursively responsible for continuously (self-)generating its molecular constituents through its interactions with its environment. In this way, the ‘parts’ – molecules – and the ‘whole’ – the cell – are co-emerging as entities that persist at specific levels of biochemical description. Importantly, emergence can arise at multiple levels of analysis.

3. embodiment: according to enactivism, cognition “simply cannot *but* be *embodied*” (Di Paolo et al., 2010, p. 42). Embodiment is not intended to convey the notion of the body as an external ‘shell’ which is, at most, some kind of add-on machinery to a central cognitive controller. Rather, “the body is the ultimate source of significance” (ibid.) – it is the animate locus of autonomy, through which meaningful activity is created. This enactive embodiment has two key connotations. Firstly, an embodied organism is *coupled* to its environment through a *sensorimotor cycle*, in which “what one senses depends directly on how one moves, and how one moves depends directly on what one senses” (Thompson, 2007, p. 244; see also section 4 of chapter 1 on the ‘complementarity’ of subjects and their environments). Bodily activity and activity in the surrounding environment are thus mutually specified and inherently informative, relative to the organism. Such sensorimotor coupling entails the conception of perception as a kind of action (O’Regan and Noë, 2001; Noë, 2004). On this view, perception does not involve receiving static ‘inputs’ to be processed, but is an embodied activity that manifests understanding through sensorimotor knowledge of how variances emerge from certain movements. Another way of putting this is that perception is the enaction of sensorimotor contingencies (i.e. the lawful co-dependency of sensory perceptions and motor activity). Embodiment, then, in virtue of sensorimotor coupling with the environment, is a property that endows the body with cognitive competency – most notably, the capacity for perceptual understanding – within the co-determined organism-environment domain.

Intermingled with this sensorimotor insight, embodiment also carries the phenomenological connotation that an organism's body is not merely another physical entity, but is a *lived body* ('*Leib*'; see Husserl (1952/1989)). As such, one's self is not to be dualistically conceived as separate from one's body; rather, one's self is a *bodily* phenomenon – “a nexus of lived meanings” (Merleau-Ponty, 1945/2012, p. 150) through which one's subjectivity is expressed. Embodiment thus encapsulates the idea of oneself as a bodily subject of experience.

Taken together, these two connotations highlight how embodiment, for enactivism, is indispensable as the means by which an autonomous organism establishes a dynamic and recursive connection with the environment, therein defining the environment as a uniquely significant domain for the lived expression of one's subjectivity.

4. experience is the qualitative means by which our activity in the world is consciously intelligible. Rather than being seen as an epiphenomenon or an obscurity, enactivists consider experience to be “intertwined with being alive and enacting a meaningful world” (De Jaegher & Di Paolo, 2007, p.487). Indeed, for any scientific account of the mind to be taken seriously, it is often claimed that phenomenological investigations into lived subjective experience are indispensable (ibid.; Thompson, 2007). Reciprocally, scientific investigations into mind can help to illuminate and naturalise the lived subjectivity that is central to phenomenology. Experience is thus the ‘meeting point’ at which enactivism pitches its tent; it is “a guiding force in a dialogue between phenomenology and science, resulting in an ongoing pragmatic circulation and mutual illumination between the two” (De Jaegher & Di Paolo, 2007, p.487).
5. sense-making is the culmination of the enactive tenets. It is the notion that being an emergent and autonomous organism that self-generates its identity entails the idea that there is always some process towards which the organism is aimed (i.e. the continuation of its autonomous, self-generated identity). The organism protects its continuous self-generation by being its own centre of activity, which correlates to it having a unique living perspective from which the world is normatively laden. This perspective is part of the mutual co-determination of the organism and its uniquely meaningful world, which results in the replacement of the classic cognitivist ‘input-process-output’ model with the idea of the normative world being ‘enacted’ (Varela et al., 1991; Thompson, 2007; De Jaegher & Di Paolo, 2007). Enacting the world means that interactions with the surrounding environment are inherently significant for the organism, and this is presented by De Jaegher & Di Paolo (2007) as the “definitional property of a cognitive system: the creation and appreciation of meaning or *sense-making* in short” (p. 487).

As the capstone of the enactive framework, it is sense-making that is equivalent to the notion of cognition within more traditional frameworks of the mind such as

computationalism and connectionism. By interacting with the world and generating and maintaining a realm of significance that is unique to itself, an organism is not passively receiving and processing information, but playing the central role in the ongoing, situated and autonomous preservation of its purposeful and relational involvement in the world. An organism engages with the significant world by making sense of it during (en)action, which is equivalent to saying that the organism cognises (De Jaegher & Di Paolo, 2007).

These five interrelated concepts define the enactive approach to cognition. Although they are mutually interdependent, it is through the notion of sense-making that the tenets are best brought together: an organism makes sense of the world through being an embodied and adaptively autonomous system that has a unique, normatively-laden experiential perspective of life. According to De Jaegher & Di Paolo (2007), amongst others, this conception of sense-making, which is achieved by enacting one's world, illuminates what it is to be 'cognitive' in a way that other theories fail to achieve. Whilst each tenet mentioned above is worthy of meticulous and in-depth exploration, I have only sketched them here so as to provide a general overview of the framework within which participatory sense-making is put forth.

### **3. Participatory Sense-making**

As should be partly clear from its designated title, 'participatory sense-making' seeks to apply the tenets of enactivism to those instances of social engagement in which agents' activity is dynamically coordinated through their interactions and new 'participatory' spheres of sense-making are created. Such social interactions are claimed to take on an autonomy of their own, and it is the circular dynamics within these autonomous social interactions that participatory sense-making aims to expose and explain.

In order to make clear how *social interactions* can generate their own autonomy, De Jaegher & Di Paolo (2007) provide a precise definition:

Social interaction is the regulated coupling between at least two autonomous agents, where the regulation is aimed at aspects of the coupling itself so that it constitutes an emergent autonomous organization in the domain of relational dynamics, without destroying in the process the autonomy of the agents involved (though the latter's scope can be augmented or reduced) (p. 493)

Using this definition, two distinctive features of social interaction can be highlighted: (i). the relational-dynamic coupling of the involved agents takes on an emergent and operationally closed autonomy of its own, so that the agents' coordinated behaviour generates and sustains the interaction, and the interaction influences and regulates the agents' behaviour, (ii). the individual autonomy of the involved agents persists (in a possibly altered form) alongside the

interaction's autonomy (ibid.). Within this realm of social interaction, individuals can then sense-make (i.e. cognise) in a co-regulatory, participatory manner.

In order to further examine participatory sense-making, it is useful to briefly consider the following three cases which bolster and support the theory:

(a). two individuals walking towards one another down a narrow corridor may accidentally mirror one another's movements, rather than making complementary moves that would allow them to walk past one another. By moving in this manner, the individuals create "a symmetrical mirroring relation [which...], in combination with the spatial constraints of the corridor, increases the likelihood that the next move will also be a mirroring one" (ibid.). This accidental mirroring, which is a kind of coordinated movement, "maintains a property of relational dynamics that forces the individuals to keep facing each other and consequently to remain in interaction" (ibid.). At the same time, the interaction itself sustains the individuals' coordinated movements as they perform a kind of to-and-fro dance in an attempt to navigate their way past one another. Thus, "[c]oordinated sideways movements conserve symmetry and symmetry promotes coordinated sideways movements" (ibid.).

(b). upon arrival at a holiday destination, Janet contentedly breathes in the fresh air through an open window in such a manner that her husband, John, can perceive her appreciation of the situation (Currie, 2007). In performing this act, "Janet is arranging things so that she and John attend to the freshness of the air, in a way that is mutually manifest to both of them.[... S]he is adjusting John's cognitive and affective take on the world: trying to get John to see the world in somewhat the way she is currently seeing it" (ibid.). Her behaviour is an attempt to frame John's interpretation of the portion of the world that is available to them both, so that he can attend to it with the same appreciation she presently has. We thus find that the orienter, Janet, is modulating the cognitive activity of the orientee, John, by opening up a new domain of sense-making that he can turn his attention towards. John's orientation towards this domain does not require an analysis of Janet's behaviour because her clear attracting of his attention towards her intended meaning will affectively generate a coordinated response (De Jaegher & Di Paolo, 2007).

(c). in the game of charades, progressive gestures by a performer, which are readily interpreted by knowledgeable team-mates, eventually lead to the attainment of a correct answer. This process begins with the performer trying to orient her team-mates towards her general intention through established signs, before making increasingly specific and informative gestures until the correct answer is (hopefully) reached. The unfolding of each performance requires both the performer and her team-mates "to adjust their sense-making in a way that converges towards the 'right' gesture and the 'right' interpretation" (De Jaegher & Di Paolo, 2007, p. 501). The interpretation of each gesture is thus co-constructed through

the interactive performance and “evolves through patterns of coordination and breakdowns.” (ibid.)

In example (a), we find that the unplanned interaction ‘takes on’ an autonomy of its own, in that the dynamics of the situation modulate the individuals’ behaviour and their behaviour sustains the interaction. We thus cannot reduce this social situation’s unfolding to the individual level, in that the interaction itself plays a regulatory role; a new, modulatory, relational domain emerges that affects both of the involved agents. Through focusing on this relational domain, enactivism is able to provide a genuinely *interactive* account of social cognition.

In examples (b) and (c), we can see how one individual can orient others towards a new domain of sense-making that they can attend to in a participatory fashion. With charades, in particular, we have an indication of “meanings being generated and transformed during interaction” (ibid.). It is only through playing the game that the performer’s gestures come to take on specific significance for the participants, a significance that is generated in the relational domain of the interaction. This concept can be extrapolated to everyday life, in which “sustained interactions (e.g., in couples, between friends, family members, work-mates) develop their own language and shared perspectives” (ibid.).

What we find across all of these examples is an indication of the everyday prevalence of the kinds of social interactions that can facilitate participatory sense-making. Whether it be the direct orientation of an involved party or an unplanned coordination of movements, we seem perennially on the cusp of falling into a relational domain of participatory sense-making when we encounter others. What’s more, it is evident that paramount amongst the features of interaction is the idea that the relational domain between two (or more) individuals can take on an autonomy of its own. With this interactive autonomy in play, it simply makes no sense to restrict investigations to individual psychology, or even the psychology of groups with supervenient properties. Instead, the interactive process, “as an emergent autonomous domain” (ibid., p. 505), must be given due attention in social cognitive research.

### 3.1. The Advantages of the Enactive Approach

In many ways, the enactive paradigm appears especially suited to social cognition, in that the dynamical and relational qualities on which it relies are expected characteristics of social interactions. That is, even a commonsense understanding of interaction conjures up the notion of a relationship between two or more persons that will rely, in part, on the persons responding sensitively to the various subtleties of one another’s actions (including utterances). It is simply by applying the concepts of (i). interactive autonomy, and (ii). cognition as sense-making, that enactivism builds on this understanding.

This is, in fact, one of the key strengths of the enactive approach to social cognition. Whereas we-mode cognition and a vast array of other theories begin with individual cognition and then add sociality into the mix as a special kind of supplement to individual potential, participatory sense-making takes as its point of departure the unique *relationship* – the interactive process itself – between two individuals. And rather than having to posit novel mechanisms and capacities to explain social cognition, the concept of an agent relating dynamically and meaningfully to the world is already built in to the central tenets of enactivism (see previous section). Participatory sense-making is simply an application of the dynamic and relational tenets of enactivism in the social domain. With the establishment of this coordinated, autonomous interaction process, participatory sense-making is able to take “an inherently social route” (De Jaegher & Di Paolo, 2007, p. 494) to explaining social cognition, rather than treating interactions “as an addendum to a position that departs from what is really still an individualistic perspective” (ibid.).

The establishment of an autonomous interaction process also allows advocates of participatory sense-making to avoid accusations of incoherency, such as the one aimed at we-mode cognition by Di Paolo et al. (2013). This is due to the fact that they are not re-construing an already established cognitive paradigm, attempting to expand or manipulate individual psychology into the collective realm of the social. Enactivists can keep references to individuality to one side as they shift focus to the active relations of an organism with its environment and the coordinated relations between organisms in an interaction. Rather than requiring a (potentially incoherent) marriage of individual and collective perspectives, enactivism, and its social subcategory of participatory sense-making, can simply address relational processes of diverse and vacillating forms. In doing so, an autonomous interaction can become its own subject of attention, avoiding many criticisms regarding individualism and reductionism because individual agents simply persist at a different level of organisation, alongside the interaction itself. As long as participatory sense-making conforms with the five central tenets of enactivism as laid out by De Jaegher & Di Paolo (2007), and stays true to the idea that “cognition is always relational” (Di Paolo, 2009, p. 19), it remains a valid and coherent theoretical viewpoint.

Due in large part to being embedded within enactivism, as a subcategory of the overall theoretical framework, participatory sense-making also finds itself robustly supported in issues pertaining to its manifestation and mechanism. In contrast, we-mode cognition, which can be considered as a kind of off-shoot from the representationalist framework of cognition, struggles to explain how membership in a group environment is sufficient for its manifestation, as well as conveying potential discrepancies between we-mode processes (such as co-representations) and genuine we-mode cognition (see section 4.1 in Chapter 3). De Jaegher & Di Paolo (2007) avoid such confusion by providing a precise definition of

social interaction which, along with the examples outlined in the previous section, seems to leave us well positioned to understand how an interaction process can take on an emergent autonomy of its own in a relational domain. We are also later provided with a concrete definition of ‘engagement’ as “[t]he qualitative aspect of a social interaction as it starts to ‘take over’ and acquires a momentum of its own” (De Jaegher, Di Paolo and Gallagher, 2010, p. 441). This ‘taking over’, or ‘taking on a life of its own’, refers to the co-regulated coupling between individuals that results in the emergence of an autonomous interaction (ibid.). By applying this concept to instances of social cognition, there is purportedly no concern regarding what ‘triggers’ participatory sense-making into life, as there was for we-mode cognition. That is, the co-regulated dynamic coupling of social interactions “can just happen” (Di Paolo et al., 2013, p. 303) as individuals’ activities become mutually coordinated in an unplanned (e.g. example (a) in the previous section) or intentional (e.g. examples (b) and (c) in the previous section) manner. Autonomous interactions thus spontaneously emerge as individuals’ activity becomes coordinated and dynamically coupled, such that the individuals have both generated the interaction process and come to be constrained and parametrically conditioned by it (De Jaegher et al., 2010, p. 441).

Through their exposition of key terms and the natural congruence of participatory sense-making with the ‘standard’ sense-making of the enactive framework, enactivists are therefore able to ground their theory of social cognition with a stability, precision and detail that is perhaps lacking in we-mode cognition (and various other social cognitive theories). Above all, participatory sense-making is able to unproblematically account for plural psychological attitudes being genuinely plural, in that “mutual co-regulation of the interaction develops into mutual recognition and finally into a sense of shared, mutually shaped intentions towards the world” (Di Paolo et al., 2013, p. 304). The sense-making that takes place within a mutually co-regulated interaction is undoubtedly ‘irreducibly collective’ and coherently substantiated by both the central tenets of enactivism and De Jaegher & Di Paolo’s (2007) examples that were outlined in the previous section.

However, participatory sense-making is, of course, not immune to criticisms. In the subsequent section, we will look at problems the enactive approach to social cognition must face; in particular, we will see that the theory encounters difficulties through its failure to satisfactorily defend against ‘bloat’ worries or to satisfactorily acknowledge the significance of (heteronomous) cultural norms.

#### **4. Problems for Participatory Sense-making**

There are, I believe, three issues which participatory sense-making must address if it is to be considered a plausible account of social cognition. The purpose of theories of social cognition is commonly accepted to involve an account of how we come to think with or about

others (in order to understand them), and it would thus be hoped that participatory sense-making can explain how our cognising with or about others in social situations differs, or can be delineated from, ‘non-social’ cognising. However, I believe the following sections will show that participatory sense-making does not satisfactorily achieve this and instead conflates two different senses of sociality.

#### 4.1. Participatory Bloat

The first difficulty that participatory sense-making must face can be approached by looking at the escalation of the relational dynamics that generate a fully autonomous interaction. The issue here is that there seems to be an extensive range of interactions that can take on an autonomy of their own, to the point where even the briefest and weakest of interactions seem to qualify as generating a relational sense-making domain, as do various human-‘object’ interactions. Put another way, this initial problem calls into question the latter part (whilst leaving unexplained the earlier part) of De Jaegher & Di Paolo’s (2007) definition of social interaction as the (co-)regulated coupling between autonomous agents that takes on its own autonomous organisation. As we shall come to see, an organism’s perpetual coupling with the environment presents the possibility of construing all ‘action’ as *interaction*. Building on this construal, it stands to reason that it is the *social* aspect of social interaction and, accordingly, the *participatory* aspect of participatory sense-making that must be explained and delineated, rather than the emergence of an autonomous organisation at a supra-individual level (i.e. it is consideration of *autonomous agents* that should take precedence over consideration of an interaction’s autonomy). Once all behaviour is considered to be interactive, one can no longer rely on an interaction’s autonomy as the explanans of how we come to understand others.

Although, as we shall come to see, enactivists may feel that they can respond to this criticism, it does, at the very least, lead to a ‘bloat’ worry regarding the instances of sense-making that qualify as participatory. The examples outlined in section 3 already provide us with a fairly broad spectrum of autonomous interactions, from the temporally defined and rule-governed activity of charades to the unplanned coordination of activity when two strangers cross paths in a narrow corridor. What we must question is the clarity of the far end of this spectrum, at which action becomes part of an *interaction* that can take on its own autonomy.

##### 4.1.1 Interactions With and Without Momentum

Let’s consider a loving couple, Jack and Jill, engrossed in conversation over dinner in a secluded corner of their favourite restaurant. Every utterance from one of them invokes a response from the other and every subtle movement is affectively attended to as a nuanced

shift in their present dynamics. Their mutually developed history and comfort in one another's presence facilitates a palpable sensitivity to one another's present needs, interests and disposition. They encapsulate an archetypical dyad, performing a conversational dance with the fluent synchronicity so telling of romantic partners. They are, quite literally, 'in their own little world', inhabiting an operationally closed relational sense-making domain that is unique to the two of them and their current circumstances. Such a scenario amply fuels the concept of an autonomous interaction, where neither Jack's nor Jill's actions can be individually confined, but there is instead a co-regulatory connectedness between the two of them that conditions their actions and is generated by them.

There seems, in this sort of situation, to be little mystery regarding the presence of the interaction's autonomy, assuming we have taken the relational doctrine of enactivism for granted. The devoted attention that each individual directs at their partner is highly conducive to the concept of the interaction itself playing such an integral role in their social encounter. However, if we slightly change the situation's circumstances then the presence of an autonomous interaction becomes less distinct. Suppose that instead of Jack and Jill, we find Jackie and Jilly in the restaurant on a blind date. Jackie is nervous. He wants to seem relaxed but is all too aware of his body posture. He wants to be funny and charming, but is over-thinking his words so that the flow of conversation is stilted and staggered. He finds himself second-guessing all of his behaviour. Jilly, too, cannot help but feel nervous and is unable to sit, eat and converse with the comfort and ease that she craves. They are, according to a commonsense understanding of the word, 'interacting' through their polite and awkward dialogue, but there is no immediate and unmistakable indication of the easygoing co-regulated coupling that defines De Jaegher & Di Paolo's (2007) version of social interaction. Nonetheless, if Jilly were to suddenly lower her voice, we can presume that Jackie would lean forwards in order to hear her better. This ingrained social know-how of leaning forwards in order to hear better is in fact De Jaegher et al.'s (2010) definitive example of co-regulated coupling. Similarly, if Jackie were to suddenly raise his hand with his open palm facing outwards, in the manner of a traffic warden signalling 'stop', it is safe to assume that Jilly would cease talking, albeit in a disconcerted or begrudging manner. The problem that emerges from these seemingly instinctive behavioural reactions is that they neatly fit the definition of co-regulated coupling, in that one individual is making a motivated change to the constraints and parametric conditions that influence the two individuals' relation (De Jaegher et al., 2010), and yet they occur in an otherwise incongruent and disjointed 'interaction'. What, then, can we ascertain about the autonomy of such interactions?

I believe the response that many enactivists would favour is that the awkwardness of Jackie and Jilly's date is itself a kind of autonomous interaction. The nervousness of each individual will affect the sense-making of the other, such that their actions become coordinated and

together they generate and sustain a domain of awkward social sense-making. Within this awkward interaction, the fleeting behaviour of Jackie leaning forward in order to hear Jilly better and Jilly stopping talking due to Jackie's outstretched hand would both encapsulate momentary autonomous interactions. These interactions would persist alongside the overall tone of the interaction and the autonomy of Jackie and Jilly as individual agents. Nothing within the description of participatory sense-making prevents such stilted occasions from encapsulating coordinated intentional activity in which new sense-making domains are generated. Indeed, this merely reinforces the enactivists' acceptance of autonomous relational processes taking on diverse and vacillating forms (as discussed in section 3.1), although it is perhaps noteworthy that such restrictive encounters are rarely discussed.

However, the problem is not yet overcome. With the example of Jack and Jill we have a benchmark of a fluid dyadic interaction. With Jackie and Jilly, we have an example of a stilted and awkward coupling that nonetheless seems to encapsulate an autonomous interaction. We can easily entertain numerous coupling variations between these two examples, in which the interaction is neither the enclosed dyad of a romantic couple nor painfully uncomfortable. Thus, archetypically dyadic interactions, awkward interactions and all those in-between seem suitably poised to exemplify autonomous interactions that accommodate participatory sense-making. Even isolated and strikingly brief actions such as one individual lowering their voice or raising their hand can seemingly lead to fleeting, autonomous interactions. All that seems to be necessary in these examples is that the individuals are attending to one another's active presence so that a relational sense-making domain emerges. Is there, then, *any* attentive meeting of individuals – no matter how brief or seemingly insignificant – that does not result in some form of participatory sense-making?

The reason that a negative answer to this question is troublesome for advocates of the enactive form of social cognition is that participatory sense-making is meant to be a *unique* interactive occurrence, in which intentional activity is coordinated in such a way that previously unavailable relational domains of sense-making become available to each interacting subject. These new domains of sense-making will then affect individual processing so that subjects can come to think with, through and about their interacting partner in a *novel way* that was absent prior to the domain's manifestation. As stated earlier, such *participatory* sense-making is naturally tied to the notion of a *social* interaction, in which (referring to the earlier definition from section 3) co-regulated coupling leads to "an emergent autonomous interaction in the domain of relational dynamics" (De Jaegher & Di Paolo, 2007, p. 493). However, it is at no point suggested that all social encounters – in which there is some qualitative element of engagement with another person – are instances of participatory sense-making; instead, the theory is intended to focus on those encounters

in which intentional activity becomes sufficiently coordinated and the relational dynamics between individuals take on their own autonomous organisation. If the theory fails to pick out these encounters then, as I shall make clear in the subsequent discussion, it will either be forced to incorporate ostensibly non-social interactions into its theoretical range, or an arbitrary line will be drawn between participatory and non-participatory instances of sense-making.

Recall that the demarcation we are provided with for the occurrence of participatory sense-making is that it “can just happen” when the interaction gathers its own momentum and “takes on” its own autonomous ‘life’ (Di Paolo et al., 2010). In section 3.1, this demarcation was put forward in a positive light as the notion of an implicit manifestation of relational sense-making withstood the need for any involved individual to make a deliberate effort to engage with the conditions of an autonomous interaction. It also permits a broad view of the scope and depth of our social interactions, rather than being restricted to certain instances of ‘joint action’ or ‘collective intentionality’, as is the wont of many other theories of social cognition. Unfortunately, the demarcation now appears unsatisfactory in that interactions do not seem to need to ‘gather momentum’, but can arise in many fleeting, feeble and unexceptional encounters. Indeed, in chapter 2 we considered how even the mere presence of another can transform a given agent’s phenomenology to an *I-before-another* structure and one could argue that this is itself a kind of minimal autonomous interaction. Goffman (1956, 1961, 1964), for instance, takes this route in suggesting that basic co-presence generates “mutual monitoring possibilities” (Goffman, 1964, p. 134), which result in interaction-specific normativity (i.e. what ought to be done in this other’s presence) and unique experience (i.e. as being ‘before-this-other’). If there is even the possibility of ‘bare encounters’ being conceived of as interactions with their own novel organisation then it becomes increasingly difficult, if it is in fact possible, to parse participatory sense-making from social encounters more broadly. And if all social encounters are instances of participatory sense-making, then the idea of an interaction gathering momentum and taking on its own autonomous life becomes entirely unnecessary.

This problem can be separated and addressed in one of two ways, but neither method is fully satisfactory. On the one hand, one could seek a more definitive ‘switching point’ for exactly how an interaction absorbs individuals into its autonomy. This can be Interpretation A. For Interpretation A, once one provides an account of exactly how and when an interaction ‘takes on a life of its own’, there is no longer any dissatisfaction with the demarcation of participatory sense-making. An empirically observable distinction between those social encounters that involve participatory sense-making and those that do not would allow enactivists to avoid any accusations of participatory interactions occurring as if “by magic” (Gallotti & Frith, 2013b, p. 304). However, positing a definitive ‘switching point’ at

which sense-making becomes a participatory event is a significant challenge, in that the notion of interactions having more or less ‘momentum’ (i.e. being stronger or weaker) is intuitively in keeping with how we experience our encounters with others. There is no reason to suspect that there is in fact a ‘switching point’, rather than a gradient of change to sense-making processes in accordance with the strength of an interaction and an individual’s specific involvement.

The alternative resolution is to take it as given that participatory sense-making is an automatic form of cognising that occurs when one is part of a co-regulative autonomous interaction with appropriate momentum. On this interpretation, which we can call Interpretation B (and which most enactivists will favour), the inexact notion of an interaction ‘taking on’ an autonomous life of its own is simply in keeping with the great variety of interactions that can seemingly exhibit their own autonomy; that is, there are bound to be numerous ‘boundary cases’ of participatory sense-making when one considers the diverse, vacillating and often-fleeting ways in which individuals interact with one another. Thus, the demarcation of participatory sense-making may not be robustly precise, but then we do not live according to precise laws – a certain natural vagueness is par for the course.

The trouble here is that this second interpretation is not really a resolution of the issue; it leads us back to the idea that all social encounters can involve fleeting autonomous interactions of various forms, so that it becomes unclear why participatory sense-making is a specific explanans for social cognition, aside from what would be a trivial claim that it involves an encounter of at least two autonomous agents. For some, this simply isn’t a problem: one simply accepts that there will be some clear-cut cases of participatory sense-making and some cases that are clearly not participatory sense-making, with a gradient of cases between these. Yet following this line of thinking is worrisome for two related reasons. Firstly, it renders participatory sense-making a purely descriptive theory of enactivism in the ‘social’ realm, rather than elucidating any novel insights (in terms of cognitive mechanisms, processes or content) into “what the interaction process does for social cognition” (De Jaegher & Di Paolo, 2007, p. 486). If De Jaegher & Di Paolo wish to “reframe the problem of social cognition as that of how meaning is generated and transformed in the interplay between the unfolding interaction process and the individuals engaged in it” (ibid.), then we need to be able to ‘frame’ the instances in which the ‘interplay’ in question is *social*. Otherwise, we’re merely considering an interplay that occurs in all forms of cognition and what De Jaegher & Di Paolo are actually doing is simply suggesting that we use enactivism as our underlying theory of cognition. Nothing about their proposal would be unique to *social* cognition. The second issue, which echoes the first, is that accepting a ‘gradient view’ of instances of participatory sense-making exposes a misstep in De Jaegher & Di Paolo’s

elucidation of their theory: whilst their aim (stated in the quote above) is to reframe social cognition according to the interplay between interaction processes and individuals, their explanatory focus is predominantly on the interaction process itself. And if the interaction process emerges in all forms of fleeting encounters, then around what exactly is a new 'frame' being drawn? The focus seems to be on the wrong explanatory feature; that is, it is not the autonomy of interactive processes that should be of primary interest, but the influence of such autonomy on individuals. The 'frame' must surely fall first and foremost around what is unique about the *individuals* who are interacting with one another, and if participatory sense-making cannot do this then it has no proper frame to work within.

Put another way, the purpose of theories of social cognition is, generally speaking, to define and explain the processes that occur when agents understand – often by thinking *with* – other agents, and most theories suggest that there is some change in our cognition in these social situations when compared to non-social situations. If, however, the enactive theory cannot pick out those interactions with sufficient 'momentum' to be classed as participatory sense-making – and instead *all* social encounters seem appropriate for participatory sense-making to take hold – then the theoretical uniqueness seems to simply lie in the fact that another autonomous agent *is there*. Such an outcome seems at odds with a focus on the creation of a meaningful relational domain between individuals; mere social presence is presumably not what De Jaegher & Di Paolo are aiming for.

#### 4.1.2. Sense-making: participatory or not?

Taking the response to Interpretation B (above) a step further, we find that the 'bloat worry' for participatory sense-making is exacerbated.

All sense-making is, by definition (see point 5 in section 2), a relational process. If we lose faith in the idea that the co-regulated coupling of an interaction must 'gather momentum' in order to 'take over' with its own autonomy (and can instead instantaneously come into play when another autonomous agent is encountered), then it is easy to also lose faith in the idea that participatory sense-making is different from any other form of sense-making in the manner that De Jaegher, Di Paolo and company propose. In order to illustrate this, let's briefly consider some loose analogies:

- the coordination of intentional activity between two individuals attempting to walk past one another in a narrow corridor (example (a) in section 3) could easily be achieved between a human and an adaptively mobile robot (which is programmed to continuously navigate its way around obstacles).
- similarly, the participatory sense-making of a group playing charades (example (c) in section 3) can be likened to a person attempting to ask a troublesome PA app on their

Smartphone a question, e.g. the question (i.e. attempted answer) is progressively modified via various wordings and tones until the 'correct' pronunciation (i.e. answer) is reached.

- a person leaning forward and straining to hear in response to another's lowered voice (as in the Jackie and Jilly example) is akin to a person leaning forward and straining to hear in response to unexpected sections of lowered volume within songs (such as the whispered sections of Bjork's 'It's oh so quiet' or the crescendo sections of Supergrass's 'Pumping on your Stereo').

- a person halting a speech in response to a raised hand (as in the Jackie and Jilly example) could arise courtesy of an electronically manifest red hand from an interactive games console that a person is talking to.

On the face of it, none of these analogous examples should be expected to count as occurrences of participatory sense-making. This is because the couplings between the systemic variables of the environment (in the form of a robot, a smartphone's PA app, a pop song or an interactive games console) and the agent's relational parameters are not *co-regulated* couplings. There is no interaction *between individuals* that leads to mutual participation in a relational sense-making domain. And yet there are couplings and there are interactions, they are simply of a human-'object' form rather than a human-human form. These interactions, it could be argued, do indeed have their own autonomy in that their constituent systemic processes will depend on other constituent systemic processes, and this global mutual dependence will sustain an identity. So, for example, in the first analogy, the side-to-side movements of the robot depend on the side-to-side movements of the human and vice-versa. Both robot and human are acting in a relational domain that they have generated and are sustaining *together*. In this way, there is a coordination of activity and a generation of sense-making possibilities that were previously absent. All that is lacking in this analogy when compared with De Jaegher & Di Paolo's (2007) original example is the absence of another autonomous person (and the presence of a programmed robot instead). Therefore, once we have conceded that the point at which a co-regulative autonomous interaction 'takes over' is, at best, a murky division, the real difference between the above examples and bona fide cases of participatory sense-making seems to be the involvement of another autonomous individual.

With all sense-making being relational – that is, involving interaction – it now seems that, for De Jaegher & Di Paolo (2007), what is initially crucial in generating and sustaining “new domains of social sense-making” (p. 497) is mere social presence. Yet even this can be brought into question once one considers that novel domains of sense-making can be generated and maintained in human-object interactions, as well as human-human interactions.

A possible response for the enactivist would be to explain exactly what it is about the involvement of another *autonomous* agent in interaction that differs from the involvement of non-autonomous entities. Such a response would involve a closer scrutiny of the concept of ‘autonomy’ that is used by enactivists, particularly how the autonomy of a sense-making individual both differs and relates to the autonomy that interactions themselves can display. This is the approach that I will take in chapter 6, demonstrating that it is not just the coordination of activity and regulated coupling that are essential to relational dynamics, but the complexity and mutual openness that is present in individuals’ autonomy. It is the manner of coordination and coupling – courtesy of the impact that one individual’s autonomy can have on another – that matters more than the emergence of relational dynamics with their own momentum. My claim will be that a mutual recognition and openness towards others’ autonomy, which I will call our *sense of otherness* (see chapter 6), leads to a more satisfactory demarcation of social cognition. For De Jaegher & Di Paolo, the lack of explanation regarding the significance of agents’ autonomy (and their capacity for sense-making) for the escalation and maintenance of an interaction (with its own autonomy) renders participatory sense-making too accommodating. Indeed, their presentation of the theory seems to succumb to bloat in such a way that the truly interesting and illuminating aspects of it become obscured. That is, participatory sense-making is a theory for explaining our capacity for “understanding others [...and] understanding with others” (De Jaegher et al., 2010, p. 442), yet in the absence of an explanation of how we come to sense and appreciate the autonomy of another agent, it becomes difficult to bind the theory to genuinely *social* understanding.

#### 4.1.3. Pervasive Participatory Sense-making

Another way of viewing the ‘bloat worry’ is by further expounding the idea that nearly all sense-making, whether others are directly involved or not, is ‘participatory’. A recent paper by Popova (2014) will help to exemplify this. Popova argues that literary narrative understanding depends on participatory sense-making. Her claim is that literary narratives are not “static or inert cultural artifacts; they are expressions of intersubjective meaningful action and participatory sense-making between tellers (narrators) and readers” (p. 8). This conclusion relies on the premise that the conveyance of linguistic meaning is an active relational process, rather than a back-and-forth delivery and retrieval of fixed and inflexible expressions. And this active relational processing of language is not limited to speech or interpersonal gesturing; for Popova, written narratives are enacted by a reader, with the written text being treated “as a conversational participant” (ibid.) in virtue of its contribution to the *interactive* interpretation that it undergoes. That is to say, the ‘reality’ of the narrative is *co-created* in the relational domain that is generated by the interaction between the

narrator's (written) language and the reader's co-regulatory interpretation of linguistic meaning. Far from being an unchanging delivery of meaning for all readers, a narrative and each individual reader will, together, generate the story's experiential world (ibid.). In this way, the narrator and the reader dynamically interact – sense-making in a participatory manner – to construct a narrative meaning. If one were compelled to accept Popova's argument, then the temptation is to extend the notion of written narrative as a 'participant' in interaction to various other uses of narratives and language. For example, listening to music (certainly lyrical music and arguably instrumental compositions too) would seem to be a participatory act in which the interpretable music and interpreting listener generate and sustain an interaction with its own autonomy (that can affect the individual's sense-making processes). This would seem to be nicely evinced in the case of live musical performances, in which performers and audience recursively adjust their playing and responses, but it is also true of recorded music if we follow Popova's reasoning and take the music's meaning to be non-fixed. Similarly, a vast array of other means of storytelling, such as films, television, works of art, theatre and computer games can be considered interactive achievements so long as the 'narrator' is an active participant in the generation of a unique interpretation of the narrative for each 'reader' (i.e. receiver of the media). This active-ness emerges from consideration of narratives' language, structure and delivery not as fixed information, but as inherently "socially recursive" and relying "on conscious modification and regulation between intentional agents (real or imaginary)" (ibid., p. 10). With such an idea in tow, a plethora of activities that are commonly considered to be 'individual' become participatory accomplishments. Indeed, one could go so far as to argue that *any* transfer and interpretation of linguistic expression, be it written, vocal or bodily, is in fact a dynamic *interactive* process.

Yet again, this all may not amount to much of a concern for many enactivists. As long as the central tenets of enactivism remain intact, the various interactive dynamics that emerge and affect our sense-making should be embraced. The problem is, however, that the further we move along this path, the more participatory sense-making begins to look like the only way to cognise, rather than being a unique consequence of the enactive framework being applied to the social realm. In order to avoid this theoretical subsumption, the *social* aspect of the enactive perspective on social cognition must be clarified. Yet again, we are left facing the idea that there needs to be an explanation of why the escalation of relational dynamics and the autonomous interaction that individuals (co-)create are *social* qualities requiring an organisation that spans individuals and reconfigures their sense-making in a novel way. An interaction 'taking on a life of its own' is insufficient in responding to this; so, too, is mere

social presence, unless further explanation is provided as to why such presence differentiates the domain of the social from other forms of sense-making.<sup>[19]</sup>

Ultimately, the demarcation of participatory sense-making turns on how individuals are poised to generate and sustain an interaction with its own regulating and regulated organisation, and I believe it is through an ontological analysis of human agents – and how human agents encounter one another’s autonomy – that headway can be made (see chapters 5-7).

#### 4.2. The Relationship between Individuals and their Interactions

Whereas the bloat worry focuses primarily on the initial escalation and scope of the interactive dynamics that sustain participatory sense-making, the second major concern addresses the nature of interaction itself. As we have already seen, the driving force behind the enactive approach to social cognition is the idea that an interaction takes on its own autonomy so that participatory sense-making emerges “between individuals and not in their heads” (De Jaegher et al., 2010, p. 446). The idea of a social collective – that is, a collective of individuals that can together generate and sustain an interaction – is thus essential to the formation of a relational level of autonomy that persists alongside individual autonomy. What is not properly considered, however, is the balance within an autonomous interaction and the relationship it has with the involved individuals’ autonomy.

Consider, for example, a mother trying to teach her young child a new word. Utterances and gestures from the mother will prompt responses from the infant which, in turn, will prompt further utterances and gestures from the mother (and so on). There is thus a clear coordination of intentional activity in that mother and child co-regulate one another’s actions. They are gripped by the autonomy of their interaction, in that they are both regulating it and being regulated by it. This seems to be a good example of a robust dyadic interaction that organisationally spans its participants. As with the earlier example of Jack and Jill, mother and child operate ‘in their own little world’, inhabiting a relational sense-making domain that is unique to the two of them and their current circumstances. Part of the reason that relationships such as these seem to naturally fit with the notion of autonomous interactions is that both participants seem to equally contribute to the generation and maintenance of the interaction’s co-regulatory dynamics. However, in many (if not most) circumstances, relationships tend to take on a much more one-sided character, such as

---

<sup>[19]</sup> Di Paolo et al. (2010) do in fact provide an alternative description of how an interaction ‘gathers momentum’ and ‘takes on its own life’ through discussion of “rhythm capacity” (p. 70). Rhythm capacity is a social skill entailing “*the ability to coordinate through the interaction with another person*” (ibid.); that is, the ability to co-create and sustain a specific ‘rhythm’ for an interaction. Rhythm capacity is not an individual ability, but something that participants in an interaction bring about and manipulate through both individual and collective processes (ibid.). Although this is a helpful concept, participatory sense-making still lacks a definitive justification of why it is a truly *social* phenomena. What is needed is an explanation of why the interactions of participatory sense-making are distinct from other (inherently interactive) instances of sense-making, without implying an arbitrary human/non-human division.

‘authority-subjection’ interactions that may occur between teachers and students, doctors and patients, employers and employees, or sports captains and team players. De Jaegher et al. (2010) themselves raise the question of how these “asymmetric interactions[...] influence each participant’s possibilities for understanding and acting?” (p. 446 (Box 2)). A more pressing question, however, is how the most extreme examples of these asymmetric, authority-subjection relationships fit with the overall picture of participatory sense-making.

For instance, let’s imagine a despicable despot, Emperor Evil, with a personal slave. Emperor Evil treats his slave with such inhumanity that he does not even name the poor subject, he is simply referred to as Slave. Indeed, from Emperor Evil’s perspective, Slave is simply an offloading device for his daily activities; he is an instrument for the execution of Emperor Evil’s whims.<sup>[20]</sup> It is even decreed by Emperor Evil that Slave must bow his head and kneel down whenever he is in Emperor Evil’s presence, unless he is instructed to act otherwise. From these basic and superficial facts about Emperor Evil and Slave’s relationship, we can entertain numerous scenarios in which they interact. For example, Emperor Evil may regularly command Slave to run him a bath, serve his dinner, or wash his clothes. There could even be occasions in which Slave is asked to have a more constructive input to activities, such as reading Emperor Evil a book, helping him with a crossword, or introducing other arrivals at the palace to him. In spite of this, no matter what Slave’s involvement, Emperor Evil – being a callous, spoilt and apathetic tyrant – views Slave merely as a means to fulfilling his desires and interests, rather than as an individual subject with any desires or interests himself. Their communicatory experiences seem to depend exclusively on Emperor Evil’s moods, beliefs and desires, and their interactions thus appear entirely one-sided: Slave simply kneels, bows his head and does as Emperor Evil commands. Yet these are still genuine *social interactions* in a straightforward sense: there is a coordination of agents’ intentional activity through “the nonaccidental correlation” (Di Paolo et al., 2010. p. 62) of their behaviour, and one can safely presume that this should generate a relational domain of activity with its own autonomy.

There is thus something slightly strange going on in the example of Emperor Evil and his Slave. We are seemingly faced with social interactions, but qualifying them as *participatory* sense-making seems to turn a blind eye to their uncollaborative and authoritarian nature. Indeed, it would not seem unreasonable to claim that the interactions are in fact merely self-determined actions for Emperor Evil, akin to him offloading chores to domestic appliances. So one perspective here would be that the couplings involved in Emperor Evil and Slave’s interactions lack any *co-regulation*; that is, the interactive dynamics are exclusively regulated

---

<sup>[20]</sup> As fanciful as this example sounds, there is actually considerable evidence of slaves being employed in Ancient Greece as ‘offloading devices’ for their masters. For example, elderly politicians would often employ a ‘memory slave’ (or ‘nomenclator’) to remember all of the names of their political connections. In this way, the slave acted as an offloading device for the politician’s failing memory (Short, forthcoming).

by Emperor Evil. According to this stance, enactivists could dismiss the interactions as cases of participatory sense-making because they are not in fact genuine *social* interactions – they are conceptually comparable to an individual ‘interacting’ with appliances such as an oven or washing machine.<sup>[21]</sup> Each act is simply the completion of an individually willed task, in which Emperor Evil is using the world around him for assistance, whether it be Slave, an appliance or some other present entity.<sup>[22]</sup> From another (preferable) perspective, however, one could argue that Slave’s involvement – even if it is purely instrumental – is still having a regulatory effect on the interaction. For example, no matter how much Slave’s humanness is ignored, he is still a human and will thus be prone to variable behaviour, including errors (as well as belonging to the basic ontological connectedness of humans that was outlined in chapter 2 and will come to play a further role in chapter 6). So every time Slave completes some mundane task – say, bringing Emperor Evil his dinner – it will vary (even if only slightly) in the manner of execution. Such variance will have minimal impact on the interaction, but it will have *some* impact. Slave’s inherent human foibles and personal traits will result in minute inconsistencies from interaction to interaction, which will affect the parameters of the nonaccidentally correlated behaviours of himself and Emperor Evil.

The difficulty is that these two perspectives, both of which seem to be consistent with enactivism, cannot both be right. According to the first view, these heavily one-sided interactions are not truly *co-regulatory* and so are not truly *social*; according to the second view, the interactions do involve (minimal) *co-regulation* and are genuinely *social*, but are unusual in that they are subject to one participant’s absolute authority at the cost of the other’s submission. It is likely that most enactivists would favour this second view, with it being more in keeping with their nondiscriminatory acceptance of a broad variety of autonomous interactions. Once more, however, this leaves us facing an unrestrictive delineation of participatory sense-making and autonomous *interactions*. Whereas before we encountered this problem through the ambiguous notion of an interaction gathering its own momentum, we now find that even within interactions the hold of the relational domain’s

---

<sup>[21]</sup> De Jaegher, Peräkylä and Stevanovic (2016) would endorse this position. They claim that within an interaction, “if one of the participants completely dominates the interaction, we are not dealing with a social *interaction* (it would be like interacting with an object, not with another subject)” (p. 6). They go on to give the following example: “Imagine a couple dance: one cannot lead unless the other assumes the role of follower, and if one participant does not contribute to the moves, it would be like carrying a doll across the dance floor” (*ibid.*). In such a scenario, De Jaegher et al. suggest that this is *not* a genuine example of social interaction and, as social interactions are the playing field of participatory sense-making (De Jaegher & Di Paolo, 2007), we can assume that it is also not an example of participatory sense-making. Yet this unearths two issues that have already been considered. Firstly, it once again gives us an unclear notion of when participatory sense-making occurs; for instance, how little movement does one require from a dance partner for the dance to become something that is achieved *together*? Secondly, De Jaegher et al.’s (2015) example seems to suggest that participatory sense-making requires purposeful collaboration. However, as the case of Jackie and Jilly exemplifies in section 4.1.1, coordination and *co-regulation* can be achieved in tense and awkward interactions as well as harmonious ones. Indeed, I see no reason why even outright conflict cannot result in the dynamic coordination of agential intentional activity such that new participatory domains of sense-making are generated.

<sup>[22]</sup> From the discussions of chapter 2 and those that will follow in chapter 6, it should hopefully be clear that I believe there to be a basic phenomenological difficulty in supposing that one person can treat another as an object (e.g. in the case of Emperor Evil and Slave). The purpose of hypothetically considering Slave’s status as an ‘object’ is to highlight the confusion surrounding when an interaction has sufficiently ‘gripped’ its participants, and so when it is truly emerging *between* the involved individuals.

emergent autonomy can be tenuous. That is, in highly asymmetric interactions, the co-regulation of relational dynamics is so one-sided that it is hard to see a collective influence on the individual sense-making processes. Both the authoritative agent and the submissive agent are *participating* in an interaction: there is still coordination and (minimal) co-regulation of activity, but the submissive agent's role is such that his contribution to the regulation of interactive dynamics is almost negligible.

Once more, the theoretical delineation of participatory sense-making seems to be caught at a crossroads. If the Emperor Evil-Slave example is *not* a case of participatory sense-making, then we counterintuitively have non-participatory interactions between autonomous agents, which is all the more strange when we consider that we can seemingly have participatory interactions between autonomous agents and non-autonomous entities (e.g. the examples involving robots or narratives from the previous section). If it is a case of participatory sense-making, then the 'bloat worry' for instances in which sense-making is a participatory event emerges yet again. With participatory sense-making being proposed to describe how individual sense-making processes can be affected in (and by) an interaction, it is not enough to simply state that the social presence of sense-making agents creates interactive dynamics in a novel relational domain; there must also be an explanation of *why and how* such presence creates a novel relational domain, along with how these relational domains vary across certain types of relationships and how these domains differ from those that are created in non-social circumstances (if, indeed, they are to differ at all).

#### 4.3. Heteronomy and Autonomy – the Social Context of Interactions

There is a final issue that participatory sense-making must address which concerns the social context of autonomous interactions. This issue builds on a criticism that was initially raised by Steiner & Stewart (2009). Steiner & Stewart begin with the premise that there are two different construals of social cognition:

(a). firstly, systems are to count as genuinely 'social' when they "consist of interactions between autonomous individuals" (ibid., p. 527) that encapsulate irreducibly collective autonomous domains. These collective autonomous domains are "based on the co-regulation of mutual coupling" (p. 543) and are most apparent in dyadic interactions. It is this construal that Steiner & Stewart align with participatory sense-making.

(b). the second construal is that social cognition should be defined by its embeddedness within pre-existing normative structures. Importantly, "these social structures are not only a set of constraints, but actually constitute the possibility of enacting worlds that would just not exist without them" (ibid., p. 528). Thus an engagement of persons is not truly social in virtue of involving an interaction between autonomous agents, but because it occurs within

the contextual setting of normative social structures which imbue the interaction with significance (ibid.).

For Steiner & Stewart, it is (b) that is the viable definition of social cognition (and, indeed, as Torrance & Froese (2011) point out, of the word ‘social’ more generally). Their opinion is that the view of social cognition arising from interactions, as is purportedly the case with participatory sense-making, is insufficient in explaining how our behaviour, even within interactions, is produced in accordance with social norms that are “*to a large extent independent of the agent*” (ibid., p. 529). This reliance on largely agent-independent norms amounts to every individual possessing a *heteronomous* attitude (or state); that is, being systemically governed by organisational structures that are external to the agent. Such heteronomy is a crucial feature of human agency and is continuously demonstrated by agents’ meaningfully executed and meaningfully interpreted actions (ibid.). Heteronomy is in fact the means by which social significance arises:

Becoming socialised is achieved by becoming heteronomous: it involves knowing that the behaviours one produces have to be performed in a certain way, and acting accordingly. Abiding by norms is a relational property of agents: it depends on the existence of these norms independently of the agent (this existence consists in their following and practical acknowledgement by a community of agents), and on the fact that the performances of the agent are recognised by other agents as being sensitive (and not randomly conforming) to these norms. (Steiner & Stewart, 2009, p. 530)

Social cognition – and, indeed, any social performance whatsoever – is thus dependent on heteronomy. Such a claim is not meant to imply that agents are bound by the heteronomous influence of social norms, but that compliance is required for behaviour to be understood as “socially meaningful and appropriate action” (ibid.). Contrary to participatory sense-making, therefore, it is not interaction itself that imports the conceptual quality of sociality. It is instead the active constitution of agents by social normative structures – a constitution that is pervasive, rendering all human cognition ‘social’.<sup>[23]</sup>

Looking at this in a little more detail, we can say that for Steiner & Stewart (2009) the idea that all human agents are “actively embedded in a shared normative order, independent of individual minds” (p. 528) highlights a serious weakness in the participatory sense-making approach; namely, that participatory sense-making only focuses on the dynamic cognitive processes *within* an interaction. In doing so, participatory sense-making may well be appropriate for analysing the dynamic processes that allow an interaction to gather

---

<sup>[23]</sup> The idea that human agents are constituted by pervasive social normative structures will be the topic of the next chapter. However, it is also worth noting that such a view stokes a tension that is implicit within the enactive paradigm. This tension is that, on the one hand, we are clearly all subjective agents, with our own autonomy and associated experiences; on the other hand, we are always in a social world, engaged if not directly with others then with social norms. Where, then, does the boundary of selfhood lie: should it be individualistically confined to embodied subjectivity or allowed to meander across the various social relations and norms that we form and engage with? And if the latter is preferred, how are such social relations navigated and how is a discernible self maintained? I will turn my attention to this in chapter 7.

momentum and be maintained, but it presupposes the contextual constitution of overarching social normative structures. The accusation is thus that participatory sense-making fails to take into account that the manifestation of interactions relies on a pre-existing social domain. What's more, if reliance on this pre-existing social domain is given its proper place, the (intra-)interactive processes (such as joint attention, action coordination and empathetic engagement) that generate and maintain the interaction, thereby affecting individual sense-making processes, are always pre-conditionally sensitive to the social normative structures that scaffold the interaction (ibid., p. 544).

According to this account of pre-existing sociality, the activity coordination and co-regulation that drive interactions are in fact constituted by social normative structures. Any regulating movements or utterances within an interaction are meaningful and meaningfully interpretable in virtue of agreeing with or violating social norms. A 'bare' interaction involving two individuals in the absence of pre-existing social norms simply could not generate the co-regulating behaviour that sustains an interaction:

the social is constitutive of the activity in question; one cannot say what an individual is doing (or intends to do) without referring to the social institution that defines the conditions under which this or that behaviour has a certain meaning. (Stewart & Steiner, 2009, p. 540)

Without addressing this heteronomous constitution and the modulation of interactive (and individual) conditions by macro-social structures, Steiner & Stewart believe that participatory sense-making is "*paradigmatically* insufficient" (ibid., p. 543) in capturing the truly social aspect of interactive dynamics. This criticism is not intended to be at the expense of interactions and their value to social studies; it is simply intended to emphasise the dependence of co-constructed interactions on pre-existing social norms.

In response to Steiner & Stewart's claim, Torrance & Froese (2011) defend participatory sense-making by expanding its theoretical reach. Although they accept that ignoring the influence of social normative structures is a flaw in the participatory sense-making approach, Torrance & Froese argue that the (pre-)existence of such structures is highly compatible with participatory sense-making. Reconsidering the 'corridor' scene (example (a) in section 3), Torrance & Froese point out that the involved participants are heteronomous with respect to the interactive dynamics in the sense that they are 'subject to' the interaction's autonomy. From an individual perspective, the participants are generating and maintaining the to-and-fro movements of the interaction; from a supra-individual perspective, the participants are performing to-and-fro movements because of the regulating influence of the interaction. Depending on the perspective that is taken, the interactive dynamics can thus be referred to as both a consequence and cause of behaviour. Similarly, the interaction can be considered as both a heteronomous structure, when considering the participants as individuals, and as

autonomous, when considering the co-created supra-individual structure itself. This perspective-taking can in fact be applied to the pre-existing social normative domain: although individuals and their micro-sociological interactions are heteronomous with respect to the domain's societal organisations, the organisations can themselves be considered autonomous in the sense that they are classed as existing as largely independent of agents. As Torrance & Froese (2011) explain, whether the supra-individual structure is a dyadic interaction or a macro-social organisation, the involved agents are heteronomous with respect to it because of the constitutive role it plays for their activity, whilst the structure itself has autonomy through constraining and modulating the individuals and "having its own 'life'" (p. 43). Torrance & Froese's conclusion is therefore that Steiner & Stewart do not provide a competing account of sociality, but one that can be united with participatory sense-making by altering one's perspective. Once autonomy and heteronomy are exposed as "point-of-view relative terms" (ibid., p. 43), it seems that the two accounts can "be brought together in order to have a properly filled-out picture of social inter-(en)action" (ibid., p. 44).

Interestingly, Torrance & Froese are able to take their reasoning one step further. For instance, if we consider how macro-social normative structures can be said to 'have their own life', then it seems that Stewart & Steiner's account of sociality may in fact be best conceived as an extension of participatory sense-making. Recall once more that the dyadic interactions which participatory sense-making focuses on are generated and maintained by individuals, and simultaneously affect the cognition of the individuals, such that the individual and supra-individual dynamics form an operationally closed, interactive system (De Jaegher & Di Paolo, 2007). There is an entirely analogous dynamical relationship at the macro-social level, in which large social normative structures are generated and maintained by social collectives, and then play a constitutive and constraining role in individual cognition. Just as two individuals play a *participatory* role in generating a dyadic relational domain that they enact together, society-wide collectives of individuals play a *participatory* role in generating society-wide relational domains that are enacted across a society. Although Steiner & Stewart are correct in describing these social normative structures as pre-existing in a general sense (i.e. relative to an individual), they fail to emphasise the fact that their continued existence is dependent on their being continually enacted at an individual and 'micro-interactive' level (such as in dyadic interactions). 'Pre-existing' should thus not be extrapolated to propose the isolated existence of these large normative structures in a completely independent social realm; the structures are instead threaded through the daily living of societies and the individuals that they are composed of, being both a modulating medium and emergent consequence of individual behaviour.

At first glance, then, it may seem that Steiner & Stewart's 'criticism' of participatory sense-making only serves to extend and bolster the enactive approach to social cognition. However, a closer examination reveals that it substantiates the troublesome issues that were explicated in points (b) and (a). On the one hand, the fact that individual cognition (whether an individual engaged in an interaction or not) is partly constituted by social normative structures reiterates the difficulty in satisfactorily delineating participatory sense-making in the form that it is presented by De Jaegher, Di Paolo and colleagues. That is, with Steiner & Stewart's account of sociality incorporated in the enactive picture, it seems that as far as humans are concerned *all* sense-making is in fact participatory, in virtue of the fact that individual cognitive processes are regulated by the social normative structures that are maintained by individuals who *collectively participate*, as organisational constituents, in the maintenance of the structures. There are no isolated circumstances, away from a social context, in which an individual can sense-make in a non-participatory manner. Even biological norms are socially codified from a human perspective.<sup>[24]</sup> On the other hand, there does seem to be something of value in face-to-face interactive dynamics and their influence on individual cognition. Unfortunately, the current theoretical picture of participatory sense-making is unable to pick out and delineate these purportedly basic cases of social cognition from amongst human-object, asymmetrical and society-wide interactive behaviours.

## 5. Conclusion

Taking the issues from sections 4.1–4.3 together, it should be clear that although participatory sense-making delivers much of value to social cognitive studies, there is still work to be done in explicating how we think about, with and through others. The issues can be broadly compiled into a double-edged problem:

(i). if participatory sense-making is the extension of enactive tenets into the social domain so as to provide an enactive account of social cognition, then it must be conceded that it fails to differentiate social cognition from the cognition of (allegedly) 'non-social' events in which a

---

<sup>[24]</sup> To exemplify the social codification of biological norms, consider internal bodily processes, which are often viewed as being exclusively biological. For instance, hormone levels, or the internal regulatory processes of internal systems (e.g. respiratory, cardiovascular, or renal systems) seem peripheral to sociological or social-cognitive research. However, it is easy to highlight the bodily impact of a social phenomenon such as stress (Brotman, Golden and Wittstein, 2007). At a 'biological' level, stress is tied to the release of hydrocortisone and Engert, Plessow, Miller, Kirschbaum and Singer (2014) have recently shown that our hydrocortisone levels can be increased by merely observing others undergoing a stress test, with the increase being more pronounced if the observed individual is a loved one. This would suggest that, at least in some cases, internal biological systems are modulated by the social world. A further example would be psychophysiological interdependence amongst long-term partners (Bourassa, Memel, Woolverton and Sbarra, 2015), which continues after one partner has died (Bourassa, Knowles, Sbarra and O'Connor, 2015), with the suggestion being that we can engage in relational patterns which affect our physical and psychological well-being even when no other is present (*ibid.*; De Jaegher et al., 2015).

For further studies regarding the social codification of biology, one could point to subjects unconscious imitation of one another in interaction (Chartrand and Bargh, 1999), or the widespread modulation of the endogenous opioid system through social activities such as music-making (Tarr, Launay and Dunbar, 2014) and synchronised sports (Sullivan and Rickers, 2013).

One may claim that all I am doing here is painting a causal picture of the world influencing the body and this, in turn, influencing an agent's behaviour in the world. Yet while there is of course a place for considering bio-physical systems (e.g. organ and tissue systems) as largely isolated, when one's concern is a human agent considered holistically – that is, as an organism – then it is narrow-sighted to ignore the amalgam of bodily and social processes (see chapter 7).

relational domain of interactive dynamics is present.<sup>[25]</sup> Sense-making involving non-autonomous entities, asymmetric interactions and any behaviour within a collectively maintained social context (which is *all* meaningful behaviour) all seemingly have an important participatory aspect. Participatory sense-making, then, seems to be the ‘standard’ of sense-making and it becomes unclear why its application to the social realm (or, more accurately, the realm of ‘face-to-face’ social interactions) differs from any other form of application.

(ii). if the theory wants to claim that there is something unique about our social cognition, and that individual sense-making processes are affected *in novel ways* during direct interactions between autonomous agents, then there must be some account of what makes these social situations unique. That is, the enactive approach to social cognition needs a robust explanation of how sense-making agents are imbued with interactive capacities that allow them to generate relational dynamics of a novel form.

In the subsequent chapter I tackle these two interrelated problems. I will firstly argue that cognition is indeed always a ‘participatory’ phenomenon in the manner that De Jaegher & Di Paolo propose for direct social interactions, by which I mean that we can consider all cognition to be ‘social’ in a specific way. I will then argue that what separates ‘social cognition’ as a discernible field of study are the *fundamental intersubjective qualities* (FIQs) of human nature and phenomenological *immersion* into another’s activity.

---

<sup>[25]</sup> For Gallagher (2009, 2010), the reason why participatory sense-making fails to separate occurrences of social cognition from ‘non-social’ cognition is that participatory sense-making does not address the question that theories of social cognition should be responding to. That is to say, participatory sense-making addresses the question: “How do we, together, in a social process, constitute the meaning of the world?” (Gallagher, 2010, p. 113). Whereas the question that Gallagher believes theories of social cognition should address is: “How do we understand another person?” (ibid.). Clearly these two questions are related, yet there is also a notable asynchrony between them. The issues that I have raised throughout this chapter could be used to substantiate Gallagher’s idea that participatory sense-making is addressing a different topic – namely, it is addressing meaning constitution in interaction – to the ‘standard’ topic of social cognition.



# **Chapter 5 – The Pervasive Social Constitution of Human Existence**



## The Pervasive Social Constitution of Human Existence

### 1. Introduction

#### 1.1. A Recap

Throughout the two preceding chapters, I have described and criticised two recent ‘interactive’ theories of social cognition; namely, *we-mode cognition* and *participatory sense-making*. Whilst I have argued that participatory sense-making is the stronger of the two, I made clear in the last chapter that it is still flawed. I outlined three challenges – the problem of participatory bloat (section 4.1), the relationship between individuals and their interactions (section 4.2), and the issue of socially contextualising interactions (section 4.3) – which I believe participatory sense-making must respond to if it is to be considered a satisfactory theory of social cognition. Implicated in these three challenges is a more broad and perhaps more pressing problem regarding how we conceive of cognition (or ‘sense-making’) generally within enactive theory (although, as we shall see in section 4 of this chapter, the problem is not in fact limited to enactivism). This problem can be summarised as follows: enactivism produces a confrontation between the idea that all cognition is ‘social’, and the idea that there is something unique about our cognising within social interactions. The uniqueness of cognition within social interactions is allegedly due to the generation of a regulatory relational domain in interaction, the autonomy of which affects individual cognition so that sense-making becomes a collective process (De Jaegher & Di Paolo, 2007); yet, as I showed in the previous chapter (sections 4.1.3 and 4.3 in particular) and as I will elucidate further in this chapter, *all* forms of cognition seemingly generate a regulatory relational domain that affects individual cognition. The problem is thus one of parsing truly social cognition, in which we come to understand others by thinking *with* or *about* them, from all other forms of cognition. It is this problem that I will tackle over the next two chapters.

#### 1.2. Clarifying the Problem

The aforementioned challenges to the robustness of participatory sense-making have left it (and enactivism more generally) theoretically open to two different interpretations:

(I). the first interpretation is that *all* cognition can be considered ‘social’, if ‘social’ is equated with the notion of ‘participatory’ in participatory sense-making.<sup>[26]</sup> This is because all cognition seems to involve the generation of a regulatory relational domain that affects individual sense-making processes, whether this domain is generated through interactions between persons, or between an individual and some aspect of the material world, or between an individual and the social normative structures that condition our behaviour

---

<sup>[26]</sup> I am taking for granted the equivalence of sense-making with cognition (see Thompson (2007) and section 2 in the previous chapter).

(there will be more on this shortly, in section 2, although the idea is already present in the previous chapter). In other words, *social* cognition (or participatory sense-making) is all there is – it is our *only* form of cognising (or sense-making) whether with others or alone – and its application to interactions between people is interesting purely in virtue of the frequency and familiarity of these circumstances. On this interpretation, the claim is that there is nothing uniquely novel, in terms of mechanisms and processes, about cognising in the social realm; such instances are simply captured by the application of our ‘blueprint’ cognition to situations involving other persons.

(II). the second interpretation rejects the thought that *all* forms of cognition can be considered ‘social’ (or ‘participatory’) and instead states that there *is* something uniquely novel about cognising in the ‘social domain’, namely, that the cognition of social situations involves the *co*-regulated coupling of *autonomous entities* (in which ‘entities’ are typically, but, as we shall see shortly, not always, human agents). According to this view, the social cognition that is captured by participatory sense-making is distinct from non-social cognition because the interaction of participants constitutes an autonomous self-sustaining organisation in the realm of relational dynamics that is both generated by and regulates entities’ activities (De Jaegher & Di Paolo, 2007; De Jaegher et al., 2010). The distinctness of social cognition, according to this view, is that there is more than one autonomous organisation engaged in a back-and-forth interplay that manifests an interaction which ‘takes over’ with its own autonomous momentum, so as to engender a collective character to the generation of meaning. The idea is that our everyday, non-social cognition does *not* involve relational dynamics with the capacity to generate an autonomous interaction of the aforementioned kind; thus, the *co*-involvement of more than one autonomous entity leads to the generation of a novel and distinct kind of cognising/sense-making.

The divergence of these interpretations can be demonstrated by looking at ‘differences in kind’. For (I), there is no difference in kind between the relational dynamics that are generated in allegedly ‘social’ situations and the relational dynamics that are generated in non-social situations. Both simply involve the generation of cognitive domains of equivalent kinds in virtue of the relation between a given agent and some aspect of the agent’s world (if there is to be a difference, it is one of ‘degree’ rather than kind). For (II), there *is* a difference in kind between social and non-social situations, in that the relational dynamics that are generated between autonomous entities are novel in virtue of being *co*-regulated and thereby constituting an autonomous organisation that regulates the entities’ activities.

The problem arises once one realises that both (I) and (II) are fundamentally flawed. With regards to (I), the difficulty in claiming that all cognition is a matter of variations across dynamical uniformity is that there seems to be a commonsense consensus that there is

something significantly different, beyond mere mental contents, in my thinking with or about another present person, when compared to my thinking of, say, a teacup. The trouble that can be levelled at this interpretation is thus that without further development, it posits no methodological criteria for parsing our cognising with/about other persons (in virtue of which we come to *understand* others) from the cognition of an isolated individual, which is a radical and controversial stance to hold. With regards to (II), the difficulty (as shown in the last chapter) is that almost *any* interaction – those between a human and non-human autonomous entities, heavily asymmetric interactions between humans, and any interaction of a human with a collectively maintained ‘autonomous’ social structure (which I will shortly argue incorporates *all* meaningful behaviour) – can seemingly display an important ‘participatory’ aspect. That is, all of these interactions and not just those involving only humans seem to be able to generate a realm of relational dynamics that regulates individuals’ cognitive behaviour. The criterion of *co*-regulation leading to a novel form of *social* cognition (i.e. participatory sense-making) is thus unsatisfactory, unless it can be shown why the ‘co-’ prefix is itself unique (i.e. why the involvement of (human) agents creates relational dynamics that differ significantly, and not just in complexity, from all other relational dynamics). In other words, this interpretation fails to respond to the criticisms of the last chapter, and satisfactory methodological criteria for parsing our cognising with/about other persons from our cognising from a seemingly individual perspective are still lacking.

### 1.3. The Plan to Tackle the Problem

When two theoretical interpretations conflict with one another, as with (I) and (II), the normal plan of action is to argue in favour of one interpretation and against the other. I am going to shun this plan of action in favour of arguing that, with minor adjustments, both (I) and (II) encapsulate justifiable claims.

The first claim (this chapter) will build on response (I) by justifying the idea that all sense-making/cognition *can* be conceived of as being social in a specific way. The second claim (next chapter), drawing on response (II), is that there *is* something unique about our understanding others through cognising with and about them. To this end, I will put forward criteria for delineating a new framework for ‘social cognition’, with social cognition being taken to encapsulate the transformation of human cognition that unfolds when we are immersed in interactions with co-present others. The two claims, developed from the nuclei of (I) and (II), thus draw out and delineate two forms of ‘social’ – one which encapsulates everything we think or do as humans (in virtue of the fact that human existence, in the philosophical sense, requires (*en*)socialisation in self-other generated norms) and one which captures the results of the novel relational dynamics within engagements between co-present agents.

An underlying motivation for this plan is that I believe the term ‘social’ is widely taken for granted and misused in a wide array of disciplines, but particularly in those associated with cognitive science. In order to rectify this, I will explain in this chapter that everything we think, feel, or do is ‘social’. I will firstly show that all human actions are inherently social (section 2), then that all thoughts are social (section 3), whether from an enactivist or functionalist perspective (section 4), before ending with some remarks on the enmeshed nature of societal and individual existence (section 5). In the subsequent chapter, I will address the idea that situations in which we directly think *with* or *about* co-present others seem to be worthy of delineation as a unique kind of narrowly defined ‘social cognitive’ accomplishment.

## **2. Being Social**

### **2.1. Social Structures and Heteronomy**

In order to argue for the idea that all human thoughts and behaviour are social, I will begin by briefly readdressing the work of Steiner & Stewart (2009) that was encountered in the last chapter and looking at how it resonates with the work of others (e.g. Dreyfus, 1991; Giddens, 1976; Haugeland, 1998; Heidegger, 1927/1962). Included in the motivation for Steiner & Stewart’s work, as it is for mine, is dissatisfaction with the definition of ‘social’ within the discipline of social cognition. Thus, they point out that phenomena as inconsonant and wide-reaching as ant colonies, wolf packs, ape communities, domesticated pets and their owners, traffic jams and empathy between humans are all seemingly ‘social’ from different theoretical or empirical viewpoints (p. 528). More than this, for Steiner & Stewart, any theory that takes the starting point of social cognition (or ‘social’ activity in any sense) to be ‘micro-level’ interactions between individuals will obscure “the root of sociality” and run the risk of reducing all cognitive events to an individual level (ibid., p. 534). In an attempt to eliminate such confusion, Steiner & Stewart claim “that a truly social domain is always defined by a set of structural norms [...that] actually *constitute* the possibility of enacting worlds” (ibid.). Rather than focusing on some sort of qualitatively discernible interaction being “the lowest common denominator [that is] necessary and sufficient for calling a phenomenon ‘social’” (ibid., p. 528), their claim focuses on embeddedness within pre-existing normative structures. These pre-existing structures – that are present in the form of the institutions and social norms that pervade and condition our behaviour – are responsible for our ability to *socially* interact in the first place; they are the bearers of our sociality and facilitate any interactive formation of an intra-individual autonomous organisation. Due to being the bearers of our sociality, the actualisation of these normative structures is how Steiner & Stewart (2009) believe ‘social’ should be defined, rather than its definition relying on the production of intra-individual domains during interactions that manifest novel

cognitive processes in individuals (p. 532). In turn, these social normative structures are said to *constitute* and structure our cognitive behaviour, such that they are able to claim that “[h]uman beings are socialised through and through” and “human cognition cannot but be social cognition” (ibid., p. 527). Whilst I will outline my sympathies with this view throughout this section, we will come to see throughout this chapter and the subsequent one that Steiner & Stewart’s definition of ‘social’ ultimately obscures a key aspect of human social life.

Central to Steiner & Stewart’s position is the idea that we are *heteronomous*: we implicitly abide by norms, which are largely external to ourselves, that elevate our behaviour from mere physical movements to actions with attached responsibilities (ibid., p. 529). Because of the implicit adherence to heteronomous – and thus communally available – norms, actions can be interpreted by others (relative to the norms) and so become meaningful to them. Heteronomy of this kind does not just normatively condition our behaviour in the form of constraining conditions, but is also the resource for encountering new action possibilities, in that conformist or iconoclastic behaviours alike depend on the social structuring of the involved parties. Both progression and restriction of individuals and societies will unfold within normative structures of the same social form, in terms of their historical generation and present/future persistence. Individual behaviour is permeated by the heteronomous structures of institutional and societal norms. Heteronomy is thus seen as an extension of “a conception of autonomy that is dependent on various figures of otherness” (Kaplan, 2003, p. 113), in virtue of the world appearing to us as possibilities or conditions that are enveloped in an immersing network of self-and-other-generated normative structures. The divergence from autonomy lies in the fact that heteronomous structures are *self-and-other* (or ‘intersubjective’) phenomena: individual subjects play a role in the generation and maintenance of heteronomous structures, but heteronomous persistence relies on the relational contributions of others, rather than any one individual. Essential to Steiner & Stewart (2009) is the idea that heteronomy is not opposed to autonomy, but enhances it. It is a requisite of heteronomy as much as it is of enactive autonomy that an organism is not physically isolated, but is embodied and embedded within an environment with which it interacts, such that heteronomous structures are not external determinants, but necessary resources that allow autonomous organisms to enact aspects of the world that would otherwise be unavailable (ibid.). Being heteronomous is therefore tied to our very existence as autonomous humans and allows us to continuously enact a social life at all times, rather than social enactment requiring some direct, face-to-face engagement with another (or others).

There are myriad ways of exemplifying the enaction of these social normative structures that pervade human life. Steiner & Stewart (2009) turn to our use of language, describing

how “[e]ach of us can only actually engage in the activity of speaking because there is “always already” the normative social structure of a shared “language”” (p. 533). Quite simply, an individual’s spoken communication relies on engagement with the pre-existing body of language that is immersed in the individual’s culture. For Steiner & Stewart, it is not the conveyance of language between individuals that should have hegemony over the notion of ‘social’ within linguistic communication, but the presence of language as a normative structure that can be ‘tapped into’ by norm-abiding agents. Language is thus not a ‘social tool’ purely in virtue of allowing us to communicate with one another, but also in the sense that it is generated and maintained by agential communities, in a manner that is specific to them and largely independent of any isolated individual. Luhmann (2002/2012) makes a similar claim to this, pointing out that, under normal circumstances, we presuppose the social functioning – the comprehensibility – of language when we talk (p. 196). We do not explicitly actualise all of the conditions and rules that are necessary for the successful use of language; rather, we implicitly engage with them in virtue of their communal maintenance.

Analogous to this use of language, but less obvious in its manifestation, Dreyfus (1991) describes how there are normative structures conditioning how far individuals stand apart from one another during conversations. Just as the ability to verbally communicate in an appropriate manner depends on engagement with normative structures of language, so the ability to stand an appropriate distance from another depends on engagement with normative structures of distance-standing practices. In North Africa, for example, distance-standing practices result in individuals generally standing closer together than people do across Scandinavia (*ibid.*, p. 18). This is clearly not something that is deliberated about under normal conditions; rather, it is due to implicit conforming to norms that modulate the appropriate distance to stand from someone. And these norms are specifically moulded across every circumstance, being relevant to culture and to the persons involved (so that, for instance, it is in general normatively appropriate to stand closer to a lover than to a stranger).

To further consolidate the implicit adherence and ubiquity of these social norms, consider that in most Western countries, a greeting between friends will perhaps involve a hug, handshake or kiss, as this is what is expected due to the cultural embeddedness in normative structures of greeting practices. In certain Arab social circles, on the other hand, an amiable greeting may involve one individual sniffing another (Wheeler, 2005). Practices can thus vary vastly both within and across cultures, but the uniting factor is that they are always relative to some enveloping normative structure. Even when alone, behaviour is regulated by these norms. When I eat a sandwich by myself in my kitchen, the manner in which I eat is still regulated by the norms of the culture (relevant to eating) that are specific to me. I may eat more messily and noisily than if I were in a crowded room, but this can simply be

attributed to it being normatively conventional to eat more messily and noisily – or simply in a more uninhibited manner, whatever this may involve – when alone.

Of course, it is not the case that we always blindly follow these norms, nor that they are always rigidly entrenched (and therefore ostensibly ‘breakable’) like the rules of a legal system. It is rather that under normal conditions we follow these norms without realising it as they are simply *what one does* in different situations. There is always the option of behaving in some unconventional manner and not doing *what one does*, but, as we shall see shortly, this too will be enmeshed within pervasive social structures of normativity.

## 2.2. Socially Conditioned or Socially Constituted?

Thus far, a critic could be forgiven for arguing that Steiner & Stewart’s (2009) claim regarding our heteronomous immersion in social normative structures is merely revisiting well-trodden ground. The idea that our cognitive processing and behaviour is meaningfully embedded in causally influential social contexts is far from radical and is something that very few cognitive scientists would deny. However, Steiner & Stewart do not merely propose that normative structures provide a regulatory context for human behaviour; they instead make the stronger claim that the social normative structures are *constitutive* of human behaviour. This idea rests on our active involvement and *reciprocal interaction* with norms, such that we are the ‘parts’ that generate and sustain the norms and the norms are then ‘parts’ within our cognitive processing. As this idea is vitally important to much of the subsequent discussion, I will take a closer look at it in this section.

Although I have followed Steiner & Stewart in describing the social normative structures that pervade our cognitive world as “pre-existing” and “largely independent”, these terms are only applicable from a third-party perspective. From the perspective of the normative structures themselves, it would be more appropriate to describe the structures as ‘co-existing’ and ‘co-dependent’ with individuals, as they are generated, sustained and continually altered by individual activity. It is for this reason that the normative structures do not merely condition, or contextualise, our behaviour in terms of having a unidirectional causal influence on our cognitive processes. If it were the case that we were causally connected to these normative structures in a unidirectional fashion, then our behaving in accordance with (or against) norms would have no reciprocal impact on the norms. However, this is not what happens. Torrance and Froese (2011) exemplify this through the behaviour of Rosa Parks who, in 1955, refused to give up her seat on a racially segregated bus to a white man. In doing so, she defiantly acted against the conventional norm of her society, yet this act was not independent of the norm she was expected to follow; instead, it was a kind of flexible questioning of the norm that was poised to either gather culture-wide momentum and lead to the norm changing, or be culturally ignored and thus reinforce the

norm's present state. In Rosa Parks's case, her brave behaviour reverberated throughout her society in such a way that the repercussions were nationwide and still persist in many societies to this day. That is to say, she successfully modulated part of *what one does* when sitting on a bus. The same can be said of any act, although few have such strongly transformative and far-reaching consequences. Indeed, the relationship between individual behaviour and society-wide structures is rarely as direct as in Parks's case. There is a complex non-hierarchical entanglement of individual/group behaviour and contingent normative 'levels' amongst dyads, cliques, work, groups, friendship groups, crowds, institutions, neighbourhoods and cities, right through to what could be conceived of as a global society. Consider how the popularity of music genres changes not just from culture to culture and generation to generation, but from clique to clique within the same generation and culture, with teenagers listening to a certain genre because it is 'what is cool' at that time and place. The continual change in music popularity amongst teenagers is due to individuals and social groups interacting with the norms of what is considered to be 'good music' within various communities. The same rhetoric could be applied to fashion, or films, or social media. 'Trends' continually change because individual-level engagement with them (either pro- or anti-) alters their organisation across various group 'levels', which, in turn, alters how individual-level intentions and actions are formed and executed.<sup>[27]</sup>

To enrich the above point regarding individual interaction with norms, let us consider both the interactive modulation of norms and the norm-immersion of behaviour in a slightly more technical fashion. Take the distance-standing practice mentioned earlier, which is generally followed by members of a society under normal circumstances, to be norm  $X$ . Let us then imagine that an individual who has great familiarity with this society intends to ask a stranger for directions while she is out for a walk. We can refer to the generic act of asking a stranger for directions while out for a walk as  $Z$  (such that  $Z$  is to encapsulate the set of possible variations,  $(Z_n)$ , of how one can ask a stranger for directions whilst out for a walk). In order for the individual to execute  $Z$  in the manner of *what one does* in the society, she will need to behave in accordance with  $X$ . As she does this, she is not merely satisfying her desire for directions, but is also reinforcing norm  $X$  (for herself, the stranger of whom she is asking directions and any spectators of  $Z$ ).  $X$  will then go on to impregnate the behaviour of other society members wishing to execute  $Z$ , perhaps having novel consequences for various dyads, cliques and institutions. Importantly, if the individual were to execute  $Z$  in a manner that did not accord with norm  $X$ , then she would be potentially altering  $X$ , which, in turn, would have potential repercussions for the execution of  $Z$  (within the social groups to which  $Z$  pertains). Norm  $X$  is thus modulated by both rebellious (i.e. potentially transformative)

---

<sup>[27]</sup> There is a vast amount of literature on cultural transmission, conformity, cooperative behaviour, self-organisation in groups, and cultural evolution that is relevant to this notion of a dynamic web of interactions amongst individuals and various societal groups and subgroups. See Bloch, 2005; Smith & Kirby, 2008; Kirby, Cornish & Smith, 2008; Coultas, 2004.

and conformist (i.e. reinforcing) behaviour. Of course, however, the individual in question would not be behaving independently of norms even if she doesn't conform to *X*. There would still be norms involved in how loudly directions were asked for, the tone that was used, the words that were used, who was asked, how the inquirer positioned her body, how she gesticulated and used facial gestures, how responses were given, and so on. There is simply no possibility of *Z* being executed independently of social norms, howsoever it is performed. And howsoever *Z* is executed will have a modulating affect on the norms that are engaged with (most likely including *X*). Under normal circumstance, we can thus say that *Z*'s execution is partly constituted by *X* (and other norms) and *X* is partly constituted by *Z*, in that the persistence of a norm depends on continued conformism within a group. If group members alter how they behave with reference to these norms, intentionally or not, then the norms themselves change, which may further alter the behaviour of group members, which may further alter the norms (and so on). It is in this manner that the social norms and individual actions can be said to be *(co-)constitutive*, in that persistence in one ensures persistence in another, and changes in one can lead to changes in another. There is a co-dependent circularity between them.<sup>[28]</sup> Giddens (1976) refers to this relationship as the "duality of structure" of the social world, such that "[s]ocial structures are both constituted by human agency, and yet at the same time are the very *medium* of this constitution" (p. 121). Thus, human actions cannot exist independently of certain norms, and each norm relies on certain actions for survival.

### 3. **Thinking Social**

#### 3.1. Behavioural Constitution and Cognitive Constitution

In the previous section I have outlined how social norms and behaviour can be co-constitutive. How we walk, talk, laugh, learn, eat, drink, sleep, fight and love will always be dependent on some social norm, which will reciprocally depend on how we live. Even when acting alone, our actions will carry traces of our current culture and upbringing, thereby

---

<sup>[28]</sup> One could claim that there is a kind of *continuous reciprocal causation* (CRC; Clark, 1997) going on between individuals and social norms. CRC is "the presence of continuous mutually modulatory influences" (ibid., p. 163) within a system, such that each part of the system simultaneously modulates other parts through feedback loops (where componential outputs become inputs for other components). This is thus a special kind of 'causation', in which the separation of systemic parts into insulated components obscures any useful insight into the behaviour of the system as a whole. For some, such as Palermos (2014), it is this kind of CRC that encapsulates what is meant by *constitution*. Thus, Palermos (2014) argues that when "systems mutually interact on the basis of feedback loops, there is an ongoing *causal amalgam* between the two units that disallows their decomposition into two separate systems on the basis of distinct *inputs* and *outputs*" (p.31). He goes on: "since we cannot disentangle the behavior of the two components in terms of distinct inputs and outputs from the one to the other, we must accept they *constitute* an overall system comprising of both of them" (ibid., my italics). It is through analogous reasoning that I favour the notion of *constitution* for describing the connection between societal individuals and social norms, as the norms are comprised of individual activity and individual activity is comprised, in part, by the norms. Any description that causally disentangles the two entities (i.e. social normative structures and individuals) would overlook their enmeshed nature and would instead have to treat them as insulated components, which would ultimately lead to a wrongly reductionist picture of the self and cognition.

Having said this, however, I do not wish to digress into a 'causal vs. constitutive' debate (see Kirchoff (2015) for a recent discussion). The important factor to consider, whether one is inclined to favour a 'causal' or 'constitutive' description, is that societal individuals and their norms are part of the same organisation (i.e. a culture, or society). Social norms and persons are co-dependent and one cannot exist without the other.

involving the social normative structures that help define our current culture and cultural upbringing. As human behaviour is always constituted by social normative structures in this fashion, there seems to be considerable worth to Steiner & Stewart's (2009) claim that we "are socialised through and through" (p. 527). Rather than restricting this claim to behaviour, the same idea carries through to our cognitive processing. If we return to the execution of action *Z* (section 2.2), there is also necessarily an accompanying cognitive state, *Y*, which will be implicitly laden with numerous social norms (including *X*). These norms are constitutive of our thoughts just as they are for observable actions, as these thoughts will also be fecund with the social norms of present culture and cultural upbringing. By conceiving of cognitive processing as comprising both thoughts and dispositions to behave, Haugeland (1998) explains that "when community members behave normally, how they behave is in general directly accountable to what's normal in their community; their *dispositions* have been shaped and inculcated according to[...] norms" (p. 150; my italics). Circularly enclosed in this description is the notion that norms are defined as "community-wide classes of similar dispositions that coalesce under the force of conformism" (ibid., p. 149). Importantly, these dispositions are not internalised and thus manifest by *inner* cognitive states, but obtain through "the instituted relations among public paraphernalia" (ibid., p. 154) – that is, the dispositions obtain through the communal norms that are present in societal mores and infrastructure and are enacted by collectives of individuals. The dispositions and thoughts involved therein are thus heteronomously constituted as per the behavioural constitution described in the previous section. This provides us with a repetition of the reciprocal dependence of social norms and individual actions, this time incorporating accompanying dispositions and thoughts. *Any* disposition and any consequent behaviour, along with accompanying thoughts, will be constituted by social normative structures.

Within the phenomenological tradition, this is a long-standing idea. Merleau-Ponty (1945/2012) and Schütz (1932/1967), for example, spoke of "*anonymity*" when defining the social normative background that is shared by humans and permeates our cognition, whilst Heidegger (1927/1962) referred to the same concept (broadly construed) as "*das Man*". Although for these phenomenologists there is a reciprocal dependence of collectives of individuals and social norms, there is not always an advocacy of the reciprocal manipulability of individual actions and social norms. Heidegger (1927/1962), for instance, thought that it was generally the revolutionary behaviour of authentically acting 'heroes' that could manipulate social norms. Nonetheless, the notion of a normative background of *anonymity*, constituted by 'the they' (*das Man*), is useful in highlighting the presence of a socially maintained normative background that we tend to unreflectively engage with. Indeed, an appropriation of Wheeler's (2011a) translation of *das Man* as the "they-self" is useful for capturing my present description of the communal and personal aspects of

cognition-constituting social norms. That is, the background of social norms is both inherently dependent on others (the ‘they’) whilst being personally manipulable and itself dependent on community-wide assimilations of individual-level (the ‘self’) behaviours. This background of social norms can be taken to inform “everything we expressly say, explicitly think and deliberately do” (Hutto, 2012, p. 37), such that thoughts and actions are always and inevitably socially normative thoughts and actions. By collating this phenomenological support alongside Steiner & Stewart’s view of heteronomy and my subsequent discussion of constitutive social norms, there is considerable momentum behind the notion of omnipresent normative ‘social’ existence in human life.

In the next section, I will apply this notion of ‘social’ to the theories of enactivism and functionalism.

### 3.2. Social Structures and Theories of Cognition

The ‘social’ that has been outlined in the previous section as pervasive of all human thoughts and actions is notable in that it means that whether we are thinking and acting with others, about others, or with/about some inanimate object, we are still thinking and acting in a ‘social’ way. Thus, this stance carries the claim that all cognition is social cognition. I will now briefly recount how this claim accords with an enactive account of cognition, such that all sense-making is *participatory* sense-making and interactively generated social normative structures are themselves ‘autonomous’ in the sense defined by enactivists.

Recall that sense-making is a participatory act if activity is coordinated in interaction in such a way that individual sense-making processes are affected and new domains of social sense-making, that were otherwise unavailable to the involved individuals, are generated (De Jaegher & Di Paolo, 2007, p. 497). In section 4.3 of the previous chapter, it was shown that this newly generated kind of sense-making domain can also be achieved in interactions with social normative structures. Just as the relational dynamics of an interaction – which are generated and maintained by individuals within the interaction – can take on an autonomy of their own and thereby affect individual sense-making processes, so too can a large-scale social normative structure – which is generated and maintained by interacting individuals across a community – have its own autonomy that affects individual sense-making processes. From an enactivist perspective, differentiating the interactions between individual agents that typically exemplify participatory sense-making from the interactions between an agent and a communal normative structure should come down to a matter of scale: small-scale, locally heteronomous interactions and large-scale, societally heteronomous interactions (De Jaegher, 2013). So as one individual interacts with another, relational dynamics can be created that affect the individuals’ sense-making, and as an individual interacts with a social normative structure, relational dynamics can be created that affect the

individual's sense-making. Both cases involve the reciprocal modulation of the immersed entities (two individuals, or an individual and a social normative structure). The key point for present purposes is that there is a direct analogy between two individuals playing a *participatory* role in generating a dyadic relational domain that they enact together, and society-wide collectives of individuals playing *participatory* roles in generating society-wide relational domains that are enacted across a society.

Returning to the behaviour of Rosa Parks in 1955, we can illustrate this *participation* with social norms. As Parks refused to give up her seat on the bus, she was not engaging with a static system of norms, but one which was open to questioning and rejection by her situated, affective and relational behaviour. This is the case with all normative structures, whether formalised like the legal system or implicit like distance-standing practices. If normative systems were completely static and closed to interpretation, there would be no potential for fluid cultural variation or progression. For Parks, her behaviour opened up a relational domain involving the norms associated with bus-sitting etiquette in Montgomery, Alabama (with these norms being constituted by the relevant members of Montgomery society). As the norms themselves were interpretable – and therefore manipulable – both Parks (and the immediate Montgomery community) and the norms were reciprocally open to one another's modulation. That is to say, both Parks and the societally constituted norms were *participating* (in the sense of co-regulation between autonomous organisations) in the relational domain that had been generated through interaction, in such a way that Parks's sense-making could be affected and the norms themselves could be transformed. There was no unidirectional transmission of information and no linear interaction amongst insulated components. Instead, the event was co-constituted by the nonlinear relationship between entities that were reciprocally open to modulation. As De Jaegher (2013) says, “[w]e do not just follow rules, but create them, engage with them, and interpret them in a contextual manner” (p. 22), with this flexible engagement leading to the changed or reinforced ‘rules’ then modulating future individual behaviour.

Interestingly, De Jaegher (2013) claims that only an enactive framework can satisfactorily capture this idea of a dynamic interaction between an individual and a social normative structure. However, in the following interlude I will show how it may be possible for a functionalist to make a similar claim. The purpose of this brief interlude is to demonstrate that the social normative framework of all human behaviour and cognition that I am expounding has applicability beyond the confines of enactivism.

#### **4. A Functionalist Response to De Jaegher**

##### **4.1 Functionalism and Dynamic Interactions**

Functionalism is the theory that cognitive states should be characterised by how they function within the system of which they are a part; that is, *what they do* and how they causally relate to other states, sensory inputs and behavioural outputs (Block, 1996). In recent literature, it has been suggested that the hypothesis of extended cognition – the view that the vehicles of cognition can be constitutively distributed across brain, body and world (Clark & Chalmers, 1998) – is logically entailed by functionalism (Clark, 2010; Sprevak, 2009; Wheeler, 2010a, 2010b). This has resulted in the concept of *extended functionalism*. Whilst the combination of functionalism and extended cognition is not unanimously accepted (see Shapiro (2008), Drayson (2010) and Walter (2010) for doubts and discussion), De Jaegher (2013) falls into the category of those who consider the combination theoretically unproblematic, at least in some circumstances (though not all, as we shall soon see). It is this extended form of functionalism that is primarily attacked by De Jaegher (2013) during her claim that only enactivism can satisfactorily account for the reciprocal dynamic interaction between individuals and social normative structures. As I intend to oppose De Jaegher's (2013) stance in this section, I will also presently assume the possibility of extended cognition grounded in functionalism.

De Jaegher (2013) takes specific aim at the hypothesis of the *socially extended mind* (Gallagher and Crisafi, 2009; Crisafi and Gallagher, 2010; Gallagher, 2013a), which starts out from an ostensibly functionalist perspective and argues that our minds can be “socially extended” such that they are “constituted[...] in ways that involve institutional structures, norms, and practices” (Gallagher, 2013a, p. 4).<sup>[29]</sup> This is exemplified through an individual, Alexis, who is given a set of facts and a body of evidence and then asked to judge the legitimacy of a claim (ibid.). In one possible scenario, Alexis is provided with answers to her task by experts, along with a set of rules to assist in reaching a decision. Gallagher explains that these rules reduce the cognitive effort that takes place internally in both the experts' and Alexis's heads; the rules provide a way to extend cognitive reach (ibid., p. 8). These rules and their expert interpretations, both of which are supplied to Alexis, form a large and complex system that is institutionally maintained and provides “the tracks” along which Alexis's cognitive processes must run if her judgement is to be accepted as ‘legitimate’ (ibid., p. 8). In making her judgement, Alexis is thus interacting with the institutional rules and expert opinions such that otherwise (potentially) impossible cognitive processes emerge. Her cognition is said to be ‘socially extended’ by her interaction with these institutional rules and opinions, as the rules and opinions realise – in a *constitutive* manner – new cognitive processes that drive her reaching a judgement.

---

<sup>[29]</sup> Despite starting out from the functionalist perspective of the original extended mind hypothesis, the socially extended mind is not strictly a functionalist account of mind; rather, it attempts to unite aspects of enactive and extended perspectives on cognition (Gallagher, 2013a). However, my argument that De Jaegher (2013) is wrong in so readily dismissing functionalism's ability to account for the dynamic interaction between individuals and social normative structures focuses on the non-enactive basis of the socially extended mind, which is the aspect of the socially extended mind theory that De Jaegher denounces.

De Jaegher's (2013) criticisms of the socially extended mind can be separated into three broad and interrelated claims:

(a). in virtue of its functionalist basis, De Jaegher claims that the socially extended mind is focused on "information-processing, regulated reason and affect, task-orientation, and individuality" (ibid., p. 21). As far as De Jaegher is concerned, being conceptually targeted in this way renders the socially extended mind appropriate for interactions with "patriarchal, rule-based institutions" (ibid., p. 19), with which we have "rigid, hierarchical, gendered relations that, once installed, cannot easily be negotiated" (p. 20), but inappropriate for capturing the more fluid and messy interactions which typify normal engagements with institutions and social groups.

(b). for De Jaegher, the informational processing rigidity of functionalism comes at the expense of "free and spontaneous interpersonal connection, emotion and affect, the body,[...]and, generally, dynamic change" (ibid., p. 21). That is, the socially extended mind's exploitation of functionalism reduces it to a non-embodied, non-affective and ultimately individualistic theory that is concerned with the stilted "functioning of ready-made, rigid systems defined by their place in a functional economy" (ibid.).

(c). due to the alleged inadequacy of its functionalist basis, De Jaegher also accuses the theory of the socially extended mind of being "incomplete" in its capability to account for interactions between individuals and institutions. Furthermore, the incompatibility of the extended mind and enactivism is claimed to impede any attempt to complement a socially extended mind picture with elements of enactivism (which is the approach Gallagher (2013a) takes in his most recent paper on the subject).

The first thing to note here is that (c) is not an issue if (a) and (b) are firstly rejected. That is, if (extended) functionalism can be shown to accommodate the body, affective spontaneity and intersubjectivity within interactions that move beyond regulated reasoning and rule-following, then the socially extended mind has no need to integrate itself with elements of enactivism. I will thus confine my criticisms to (a) and (b).

#### 4.2. Rigid Functionalism?

Looking at De Jaegher's (2013) complaints in a little more detail, her concern arises from aligning functionalism with (what she understand to be) the largely rigid realm of information processing, leaving little room for "[b]ody, spontaneity in reason and affect, idleness or play, and intersubjectivity" (p. 21). Whilst she considers a functionalist explanation of interactions with institutions as potentially applicable to cognition involving the rule-following of pre-existing institutions, she does not consider it applicable to cognition involving the more blurry, malleable and openly interpretable norms such as how far to

stand from one another in conversation, how to queue for a bus, or how to greet a stranger. Whilst functionalism flounders in this respect, it is with regards to these sorts of norms that De Jaegher evinces the strengths of the enactive account of cognition, as enactivism sees cognition (i.e. sense-making) as an inherently relational and dynamic process. It can thus easily accommodate the ongoing reciprocal modulation of norm-constituted individuals and community-constituted normative structures. Functionalism, on the other hand, tends to focus on “propositional attitudes, representations and informational states” (Gallagher, 2013a, p. 7), and any reciprocal modulation between systems will require a complex explanation of the circular relaying of information. The dynamic relationality of enactivism therefore seems to triumph where the more structured, sequential processing of functionalism seems to falter.

However, it is not clear that a functionalist account of individual interactions with institutional and cultural norms is resigned to failure. The assumption from De Jaegher (2013) is that functionalism is committed to a framework which obscures “the underlying fluidity[...], the messy meaning-making, the gurgling underbelly of society and mind” (p. 24). Yet the normal application of functionalism to the processing of representational data does not in itself mean that the theory must commit to the notion of rigid and non-malleable functional states, nor must a socially extended functionalist account be confined to interactions between individuals and rigid, ‘patriarchal’ structures. In the case of Alexis, for example, the legal system from which she receives her rules and expert opinions is not a completely static institution. Legal norms are readily subject to variable interpretations and adjustments, hence the long history of messy and contentious court cases. The processing of information that unfolds between an individual, such as Alexis, and a normative structure – even a seemingly rigid one such as the legal system – does not have to involve fixed rules and set algorithms for making decisions. Certainly there will be particular norms that are more steadfast and unchangeable, but there will always be scope, in any institution or social structure, for bending the rules or deciding on a unique interpretation of them. Indeed, from an extended functionalist perspective, the fluctuating nature of institutional norms and other social normative structures may resonate more harmoniously with the fluctuating functioning of organic cognitive states than do seemingly simpler cases of cognitive extension involving tools such as calculators or iPhones. This is not an attempt to invoke some sort of internally biased ‘parity principle’,<sup>[30]</sup> but rather to point out that once one accepts the indisputable cognition of neural states as involving nonlinear, fluctuating and seemingly ‘messy’ processes, and one accepts that it is possible to give a functionalist account

---

<sup>[30]</sup> The ‘parity principle’ was introduced by Clark and Chalmers (1998) as a means of explaining and defending their hypothesis of extended cognition. In its original formulation it claims that “[i]f, as we confront some task, a part of the world functions as a process which, were it done in the head, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world is (so we claim) part of the cognitive process.” (ibid., p. 10).

of these processes (see below on ‘microfunctionalism’), then the ‘messiness’ of extended processes should be unproblematic for functionalism. One of the central ideas of functionalism is that cognitive states or processes can be brought about by multiple media (i.e. they are *multiply realisable*); that is to say, functionalism contends that the designation of a system as ‘cognitive’ should not hinge on the ‘stuff’ that the system is made of, as all manner of stuff could facilitate the same undeniable cognitive performance so long as it is appropriately context-sensitive and systemically structured in the right way. The rigidity and ‘patriarchy’, or flexibility and messiness of the processes that a substrate generates are mostly irrelevant. Thus, the inevitably questionable and fluctuating nature of all forms of social normative structures is not as great an obstacle to functionalism as De Jaegher suggests; whether from an enactivist or functionalist perspective, cognitive processes in which social structure are involved can be limitlessly flexible and ‘messy’.

To further absolve functionalism (and therefore its involvement in the formulation of the hypothesis of the socially extended mind) from the accusation that it cannot satisfactorily deal with messy and fluid cognitive interactions, we can turn to the *microfunctionalism* (Clark, 1989) that is adopted in connectionist accounts of cognition. As its name suggests, *microfunctionalism* claims that the cognitive integrity of a functional state should be assessed at a fine-grained or ‘micro’ level. This is in contrast to ‘classical’ functionalism, in which theorists tend to assess the integrity of functional states in accordance with the relatively coarse-grained level of conceptual psychology (i.e. in accordance with concepts such as beliefs, desires, sensations and occurrent emotions). Due, in part, to acting at the level of conceptual psychology, classical functionalism takes as its model the abstract structure of human natural language, in which words (i.e. ‘cognitive representations’) are connected into meaningful expressions (i.e. ‘conceptual thoughts’) courtesy of grammatical (i.e. formal algorithmic) rules (Wheeler, 2005, p. 8). In other words, classical functionalism involves computational processing of cognitive states as though they are systemic data that are subject to a set of programmable instructions, analogous to the manner in which natural language functions. A sequential and formally structured functional economy is thus implicated, in which specific inputs (such as a specific pain stimulus) are dealt with by relevant parts of the system and produce specific outputs (such as avoiding the pain-causing stimulus) in line with the manifest functional state(s). For some (including, we can assume, De Jaegher), this linguistically styled sequential processing produces a functionalist framework that is suitable for ‘higher’ cognitive tasks such as logical reasoning, but less adept at more commonplace cognitive achievements such as making generalised inferences to deal with new experiences or coping with limited information. Rather than focusing on psychological concepts, microfunctionalism assesses the integrity of a functional state according to mathematical relations between ‘units’, which are canonically analogous to, or

abstract versions of, neurons within the human brain. Each unit has a unique ‘activation level’ in accordance with its connections to other units, with these activation levels being (typically) changeable (within initial configurative limitations) as the system ‘learns’ to map common and generalisable input-output connections. With its structure relying on relations between units, a microfunctionalist system is one of globally interconnected cognitive architecture that can dynamically adapt to myriad inputs. In other words, the units do not act as fixed data within an instructional (deterministic or probabilistic) programme, but can flexibly interact across the system in accordance with the relations that hold between them. By setting out its stall at this fine-grained level, microfunctionalism holds as relevant a functional profile that produces dynamically weighted, flexible and context-sensitive patterns across its parts, in lieu of the seemingly more rigid and sequential processing of classical functionalism. Further analysis of microfunctionalism is not necessary here. The point is that, unlike their classical counterparts, microfunctionalist systems can process information in such a way that context-sensitive flexible generalisations and coping with limited information are ‘naturally’ exhibited. As Wheeler (2005) puts it, the microfunctionalist seems “to get for free” the kinds of ‘natural’ cognitive processing that the classical functionalist has to pay for “in a currency of computational time, effort, and complexity” (p. 11).

The present relevance of this microfunctionalist excursion is that the focus on systemic flexibility and global relations amongst units means that we have a version of functionalism which is paradigmatically ‘messy’ and flexible. Once more, De Jaegher’s (2013) claim that functionalism is only relevant for rigid interactions thus seems unwarranted; the socially extended mind could make use of a microfunctionalist platform in order to preserve inherent fluidity.

#### 4.3. Affectless Functionalism?

In addition to the above arguments regarding flexible functionalism, it is also worth considering that in the legal case example introduced by Gallagher (2013a), Alexis remains the same affectively and rationally unpredictable entity whether an enactivist or functionalist perspective is taken. That is, De Jaegher (2013) seems to suggest that functionalist accounts of cognition treat persons, institutions, material objects and any other entity with the potential to play a constitutive role in cognition as affectless cogs within an indiscriminate cognitive system that approaches the world from a ‘neutral’ base. Such neutrality and indiscriminate stem from a conception of functionalism as only concerned with lone, reasoning, information-processing systems that can, at best, ‘extend their cognitive reach’ from an internalist basis (ibid.), instead of embracing the affective (and therefore discriminate) colouring of cognition. A consequence of this view is that De Jaegher also

claims functionalism masks the importance of *concern*, in that what cognitive states *mean* to the bodily and affective individual that ‘has’ them is of no interest, and *responsibility*, in that interactions with patriarchal institutions result in rule-determined propositional attitudes (p. 24). Yet functionalism (‘classical’ or ‘micro-’ versions) does not necessarily require individuals to be seen as bodiless, isolated, affectively neutral entities that are poised to impartially engage with social normative structures. Nor, therefore, does functionalism necessarily mask the importance of *concern* or *responsibility*, both of which are aspects of cognition that emerge from an individual being bodily and affective.

For instance, Stapleton (2013) has recently described steps that have been taken towards a ‘properly embodied’ cognitive science from within the confines of traditional, individualistically inclined cognitive science. These steps unite the (classical) functionalist processing of representational and informational states with inherently embodied – that is, bodily and affective – cognisers (in the case of humans, if nothing else). ‘Proper embodiment’ rests on the claim that “internal bodily affect is crucial to cognitive systems and an “emotion chip” [...that is somehow added into the cognitive picture...] just won’t suffice” (Stapleton, 2013, p. 7). Stapleton (2013) highlights Damasio’s (2010) recent work as indicative of an approach that could unite our understanding of affect with more traditional scientific understandings of cognition. In this work, Damasio (2010) argues that we have “primordial feelings” courtesy of brainstem structures – notably the nucleus tractus solitarius and parabrachial nucleus – that combine with the periaqueductal gray and the superior colliculus brain regions so as to neurally represent any interoceptive changes in the endocrine or autonomic systems (pp. 80-86). These internal bodily systems are constitutive of how we subjectively ‘feel’; for example, a surge in adrenaline from the endocrine system will make us feel energised, as will the invigoration of an autonomic increase in heart rate or respiration rate. Of course, such changes do not arise in isolation, but are responses to stimuli, so that the processes of these internal systems are intrinsically coupled to external stimulation of the body. Damasio’s (2010) hypothesis is thus that the aforementioned brainstem structures, which are attuned to endocrinal or autonomic changes, ensure a unity between neural systems and the ‘feeling body’. Continual changes in the endocrine and autonomic systems are continually represented in the nucleus tractus solitarius and parabrachial nucleus so that all neural processes are fully integrated with “primordial feelings” (Damasio, 2010, p. 84). As Stapleton (2013) explains, “[t]hese primordial feelings are constant and provide a background to all cognition” (p. 8). Thus, the integration of the nervous system with the autonomic and endocrine systems means that any neural process, no matter how abstract and analytical, is coloured with a background feeling that is a result of the body’s present state.

Whilst the above description is just a brief indication of the complexity of Damasio's (2010) work, it reveals the kind of route that is available to a functionalist who wishes to ground cognitive functions in an affective body, rather than treating cognition as the information processing of an isolated and affectively neutral system. With the neural system being so resolutely integrated with other bodily systems, any extended systemic state will also encompass affective bodily elements. Such encompassing of an affective body can be given either a strong or weak functionalist interpretation: the strong interpretation would be that the entire system of an individual's cognitive processes – incorporating neural machinery, the affective body and the environment (such as the norms of a social structure) – should be given a functionalist treatment, such that any cognitive explanation is cashed out in terms of the system's holistic functioning; the weaker interpretation would be that a functionalist treatment should be given only to cognitive processes of classical (i.e. non-affective) embodiment, but that these processes would then be contextualised in affect. On either interpretation, one is left with a functionalist account of cognition that is certainly not 'affectless' and seems more than equipped to cope with the affective perturbations that cognition is subject to. Crucially, once the affective body is integrated into the cognitive system in this kind of way, De Jaegher's (2013) worry about functionalism omitting the importance of *concern* is nullified. That is, any cognitive system that is sufficiently integrated with (or constituted by) affective components will produce a cognitive agent who is affectively biased towards certain aspects of the world and will experience variations in concern for decisions across different situations. Drawing on Stapleton (2013) and Damasio (2010), I would thus argue that De Jaegher (2013) is wrong to suggest that functionalism implies that cognitive agents are affectless information-processing systems; instead, cognitive states are inherently *concernful* for any agent that is appropriately integrated with affective components.

It is fair to assume that once cognition is imbued with concernful affect, there is an implication that cognitive agents can also take *responsibility* for their actions. To be concerned, and to act in accordance with one's concerns, suggests that one does not blindly follow norms – no matter how seemingly rigid and patriarchal – but that one is always open to questioning norms. Put simply, concern imports the capacity to question and rebel, and this entails *responsibility* for one's actions.<sup>[31]</sup> Once we have rejected De Jaegher's (2013) claims regarding the rigidity and affectlessness of functionalism, there thus seems no reason to suspect that functionalism lacks any of the elements of cognition that enactivism brings to the table when describing our interactions with institutions.

---

<sup>[31]</sup> See Protevi (2009) for an excellent account of how responsibility can be preserved in the case of soldiers who 'extend' their cognition into a military unit.

#### 4.4. A Functionalist Triumph?

Across the two previous sections, I have elucidated opposition to De Jaegher's (2013) criticisms of functionalism, which are allegedly damaging to the theory of the socially extended mind. In my view, functionalism is perfectly capable of providing a fluid and affective account of cognition. And if functionalism is absolved of these flaws, the socially extended mind is vindicated in its use of functionalist elements in its theoretical construction. Applying this functionalist defence to the chapter's original issue regarding the pervasive constitution of cognition by social norms, one can claim that functionalism, like enactivism, can produce a plausible explanation of the circular dependence and (co-)modulation of the relationship between individuals and normative structures. What is predominantly required is that functionalism embraces the flexibility of institutions and their norms, as well as conceding that cognising systems are not isolated, logical, rule-determined beings, but are engaged, bodily, affectively biased and questioning beings.

Having proffered and defended the capacity of an alternative to enactivism in capturing the idea of a dynamic interaction between an individual and a social normative structure, I am now going to make the tactical decision to retreat back within the comfort of the enactive framework. Rather than cowardice, this is because I will now demonstrate that an enactive perspective is bolstered by sociological and evolutionary stances on cognition that are complementary to the idea that all human cognition is social cognition.

### **5. Final Remarks on Social Constitution**

#### 5.1. Luhmann's 'Social Autopoiesis'

To this point in the chapter, I have been outlining evidence for the claim that we, as humans, are both behaviourally and cognitively constituted by social norms at all times. I think that this claim should be incorporated into any theory of cognition (with functionalism having been demonstrated as an option), but it is the theory of enactivism that is perhaps best placed to carry the claim forwards. In this section, I wish to very briefly analyse more radical and expansive enactive viewpoints that can be employed to add further weight to the notion of our pervasive social constitution.

In its original, biological conception, enactivism was generally limited to the autopoietic behaviour of molecular or organismic entities (Maturana & Varela, 1980, 1987). If it moved beyond this, it tended to remain within the biological realm, describing the self-organising and self-producing behaviour of ecosystems (Maturana & Varela, 1987) or, at its most extreme, the self-organising and self-producing behaviour of our planet (Lovelock & Margulis, 1974; Lovelock, 1979; Margulis, Sagan and Elredge, 1995). The application to supra-individual systems took a while to develop (Fleischaker, 1992; Mingers, 1994) and a specific focus on the benefits of an enactive approach to social cognition is a relatively recent

move (e.g. De Jaegher & Di Paolo, 2007). However, a slightly different track has been taken by Luhmann (1986/1989; 1984/1995; 2002/2012), who approaches enactivism from a sociological perspective. Whilst the notion of an enactive system at the level of society is hinted at in Varela, Thompson & Rosch's (1991) seminal book, *The Embodied Mind*, it is Luhmann who has expounded a rigorous concept of an autopoietic society. For Luhmann, social autopoiesis occurs within its own domain that cannot be reduced to biological or psychological elements. His argument is that the basic processes of social systems are not biochemical processes or 'psychic' processes, but *communicatory* processes.<sup>[32]</sup> A social system will normally have biological roots, in that it requires communicating subjects for its preservation, but these subjects also depend on the social environment for their own autopoiesis and, at the social level, their individual preservation is a sufficient, but not necessary, condition for the autopoiesis of a social system (theoretically, communicatory processes could be maintained by AI or non-human creatures, that is, in the absence of individual humans). Rather than society being a simple aggregation of individuals, and therefore an aggrandisement of organismic autopoiesis, it is the *communication between subjects* that is the necessary condition for generating and self-producing the emergent global order of a social domain (Luhmann, 2002/2012, pp. 190-204). Any communicatory attempt from an individual perspective can only come about if there is another entity to communicate with, from within "the network of communication" (ibid., p. 79). Thus, communicatory processes are not taken to be an individual accomplishment, nor can they be conceived of as an asymmetrical transfer of information from one being to another; instead, they are "an emergent property of the interaction between *many* (at least two) psychic [i.e. cognitive] systems" (Seidl, 2004, p. 8). In this way, Luhmann claims communication is a *genuinely social* autopoietic process: generating and maintaining a social system with a discernible structure, which, in turn, generates and maintains the communicatory processes. For Luhmann, these communicatory processes are the basic elements of "a *sui generis* reality" (ibid., p. xi), irreducible to psychological or biochemical processes of organismic or cellular levels.

The primary point of contention with Luhmann's theory is likely to be that the 'autopoiesis' he describes is in fact not autopoiesis at all, as a social system lacks a clear material boundary. However, for Luhmann, the definition of autopoiesis as the biologically grounded recursive reproduction of a system's structure and processes through its own elements – or, more simply, as "autonomy plus materiality" (Wheeler, 2011b, p. 152) – is too restrictive. He believed it worthwhile to abstract the notion of autopoiesis from biology so as to create a

---

<sup>[32]</sup> For Luhmann, communication is a tripartite phenomenon involving 'information', 'utterance' and 'understanding': "the social comes about whenever information, understanding and utterance are produced as a unity that has feedback effects on the participating psychic systems" (ibid., p. 191). Notably, this idea of communication being its own unity that modulates its participants is naturally congruent with De Jaegher & Di Paolo's (2007) notion of an interaction having its own autonomy that is both modulated by and modulates its participants.

trans-disciplinary, “truly general theory of autopoiesis” (Luhmann, 1986/1989, p. 172) that could then be applied to various domains, including the social domain. Once the concept of autopoiesis is opened up to non-materially bound systems in this way, Luhmann is free to view the self-production and self-organisation of social systems as irreducible to any ‘lower’ level, thereby interpreting societies as autopoietically maintained and specifically *social* phenomena. This is undoubtedly an interesting claim and one that I believe deserves considerable attention within sociology and cognitive science. However, the point I presently want to raise is that Luhmann’s theory of societies being autopoietic strengthens the notion of pervasive normative structures being interactively available to society-embedded individuals, as well as solidifying the mutual co-dependence of individual and society. Endowed with their own autopoietic structure, society-wide norms have a resilience that the relational dynamics of dyadic and small-scale interactions will rarely manage. For example, if the autonomy of one of the individuals in a dyadic interaction is completely eradicated, the autonomy of the interaction will also cease, in that the involved individuals are exhaustively responsible for generating and maintaining the interaction. With regards to interactions with social normative structures, however, the eradication of a single individual will not (under normal circumstances) end the autonomous (or ‘autopoietic’) self-organisation of the structure. Communications amongst other individuals will continue to generate and maintain the social normative structure, which will continue to reproduce the normative communications (i.e. its constitutive elements) amongst the individuals. The robustness of deeply ingrained social norms is thus more in keeping with the robustness of organismic autonomy than with the fleeting autonomy of small-scale interactions. This provides a natural explanation as to why the *participatory* aspect of an interaction between an individual and a large-scale social structure is perhaps less immediately evident than is the *participatory* aspect of small-scale interactions, such as those within dyads. Society-wide norms will typically be more robust and, in cases such as the legal system, will have been explicitly institutionalised so that it takes a gradual transformation in common social attitudes, or a particularly powerful event (e.g. Rosa Parks’s behaviour), to change them. The opposite end of the scale involves fleeting interactions between couples or small groups, in which, generally speaking, the norms are highly malleable and organised through the behavioural minutiae of the individuals, which are part of their communications. Between these extremes, there are interactions with norms present in medium-sized social groups, such as trends within teenage cliques or the attitudes of sports teams. It is thus not the case that individuals are *not* participating in the generation and maintenance (including transformation) of more rigid society-wide norms, but that the autonomy (or ‘autopoiesis’) of these social structures is more durable in the face of individual behaviour. Nor is it the case that the relational dynamics between an individual and a society-wide structure are

ontologically different from the relational dynamics between an individual and a small-scale social structure such as that of a dyad; it is simply that the robustness of the former tends to diminish the transformative power of individual participation. The coexistence of the regulatory drive of organismic autonomy and social heteronomy – of small-scale and large-scale structures – is a basic feature of human existence, and by considering societies as autopoietic systems in themselves (Luhmann, 1986/1989, 2002/2012), we can better understand the pervasive influence of social normative structures. Irrespective of any individual persistence, societies and their norms are robustly self-maintained through the communal activity of collectives. Reciprocally, these collectives are constitutively regulated by the social norms, such that their sense-making is a participatory act of interacting with the normative structures that they generate.

## 5.2. Back to Biology?

Taking Luhmann's notion of autopoietic society to its most radical conclusion, one could even attempt to conceive of social systems as a sort of living organism. As noted already, Luhmann distinguished social autopoiesis from 'psychic' or biological autopoiesis, so he would undoubtedly resist this idea. Indeed, it would be a highly contentious stance to take from within any school of thought. Yet societies do seem to display *adaptive* behaviour, generating events that will develop or solidify communicatory processes – such as the production of new technologies, the assimilation/protection of cultural traditions and the movement into various factions that generate more fine-grained norms – and moving away from threatening events such as the loss of political order or solidarity (i.e. the cohesion that arises from the interdependence of individuals (Durkheim, 1893/1997)). Johnson (2001) describes such societal adaptation through the formation of neighbourhoods:

[N]eighborhoods are patterns in time. No one wills them into existence single-handedly; they emerge by a kind of tacit consensus: the artists go here, the investment bankers here, Mexican-Americans here, gays and lesbians here. The great preponderance of city dwellers live by those laws, without any legal authority mandating that compliance. It is the sidewalk – the public space where interactions between neighbors are the most expressive and the most frequent – that helps us create those laws. (p. 91)

Societies thus seem to pulse with their own viable adaptations, independently of any individual's or central government's control. Coupled with the concept of irreducible social autopoiesis, this adaptivity would seemingly accord with the definition of a minimal living system (Thompson, 2007, p. 159). Notwithstanding the fragility of this claim, it leads to what I believe to be a particularly useful way of thinking about the ongoing interactive relationship between individuals and social normative structures; namely, the concept of *symbiosis*.

Symbiosis is broadly defined as the living (and thus interacting) together of different organismic species. When both of the involved species are entirely dependent on one another for their persistence, the relationship is one of *obligate symbiosis* (Douglas, 1994). In such relationships there is often interdependent co-evolution. This seems applicable to human individuals and their societies: as humans have evolved and means of communication have developed, so, too, have the societies which we create and live through, not simply as a consequence of individual behaviour, but as the medium through which such behaviour is executed. The environmental entities that are available to us for cognising have increased in complexity and variability as societies have increased in complexity and variability. Indeed, there is even growing evidence that modern culture is significantly altering our ontogenetic development, such as the claim that human memory adapts to the availability of technological information and technological means of communication (Sparrow, Liu and Wegner, 2011). This suggests that we, as modern humans, are not merely structurally coupled to a ‘natural’ environment, but are structurally coupled to the societal environment – the infrastructure, technology, conventions and beliefs – that we reciprocally interact with.<sup>[33]</sup>

Recent evolutionary theories support this hypothesis. According to traditional ‘adaptationist’ theories, gradual and cumulative changes will occur in a species so that they become better fitted to their environment and thus have a greater chance of reproductive success. This, in a nutshell, is the process of natural selection: the cross-generational “preservation of favourable variations and the rejection of injurious variations” within a species (Darwin, 1859/1996, p. 176). As Gould and Lewontin (1979) have pointed out, however, this adaptationist picture seems to treat the environment as separate from the organism, as a kind of puzzle that species must try and adapt to in the best possible way in order to survive. Treating the environment and organism as independent in this way directly contravenes the co-dependence that is essential to enactivism (Thompson, 2007, pp. 202-205). Thus, for *enactive evolution*, organisms are not passively subject to environmental manipulations, but are continuously adapting as a contingency of their *active* autonomy. That is to say, “the condition of adaptation is an invariant of life” (ibid., p. 205) – it is not the case that organisms optimise their adaptations to an environment, but simply preserve their adaptive capacities, which are necessary for their viable continuity and are a result of their autopoietic organisation and structural coupling (ibid.). Adaptation, for the enactivist, differs from traditional adaptation in that it is a process of *satisficing* – “taking a suboptimal solution that is satisfactory” (Varela et al., 1991, p. 196) – rather than optimising. It is the *conservation* of this condition of adaptation that is central to enactive evolution. Mapping

---

<sup>[33]</sup> Structural coupling generally refers to the co-dependent structural congruence of organism’s physical structure and its environment in virtue of a history of recurrent interactions between the two (Maturana & Varela, 1987).

this belief to the societal environments in which humans live, one can claim that our construction of a society that constitutes how we think and act is a means of continuing the satisfaction of our (autonomous and) adaptive capacities. The conservation and development of society is thus inherently tied to the conservation and development of our cognitive and behavioural potential, which results from our continuing capacity to autonomously exist in such a way that we adapt. As far as humans are concerned, both organism and social structures seem to symbiotically satisfy one another's integrity to persist: societal adaptations are constituted by collectives of individuals and facilitate individual preservation, and individual adaptations are societally constituted and facilitate societal preservation. Both humans and their social environment co-evolve in virtue of preserving their adaptive potential; both mutually *participate* in the enactment of life.

Moving forward with Luhmann's (2002/2012) concept of social autopoiesis, it is thus possible to uncover a new perspective of the interactive relationship between human individuals and their societal structures. The infrastructure, technology, mores, conventions and beliefs of society canalise our physical and cognitive activity. Every human plays a participatory role in the generation and maintenance of our social surroundings, and these surroundings constitutively shape how we think and act. What's more, we have co-evolved with these social surroundings, generating the very same structures to which we are coupled. The self-organisation of social structures helps to preserve conditions for individual-level viable adaptations and the self-organisation of collectives of individuals at various levels helps to preserve conditions for societal-level adaptations. Without the social structures that we generate, we simply would not exist as we do.

## **6. Conclusion**

In the previous two chapters, analyses of two recent 'interactive' theories of social cognition (*we-mode cognition* and *participatory sense-making*) eventually led to an ostensible theoretical crossroad. Down one path of this crossroad, which we can call 'the left turn', is the idea that all cognition is dynamically uniform. More specifically, the idea is that any cognitive event involves participation in the enactment of social norms, which results in the assertion that all cognition is 'social'. It is this idea that I have considered and elucidated throughout this chapter, considering human social constitution from a behavioural perspective (section 2), the cognitive perspectives of enactivism (section 3) and functionalism (section 4), and, all too briefly, from sociological and evolutionary perspectives (section 5). The conclusion that is hopefully apparent is that our very existence as humans is contingent upon our continuous interaction with social normative structures that we generate, maintain and transform. These social structures reciprocally generate, maintain

and transform our cognitive lives. From the moment we are born into a familial society, until our death, we are irrevocably *ensocialled*. Indeed, to be human at all is to be social.

In the next chapter, I shall investigate ‘the right turn’ of the theoretical crossroad. The idea that lurks down this path is that there is something unique about our cognising with and about co-present others. Whilst I wholeheartedly endorse the claim of pervasive social constitution of human existence that has been expounded thus far, I also believe there is some truth to this ‘right turn’. The ‘crossroad’ is thus not an either/or junction, but one in which both ‘turns’ have value and one (‘the right turn’) must work within the other (‘the left turn’). The position I am putting forward is that within the overarching framework of ‘all cognition being social’, there are still reasons to carve out human cognising with and about other present humans as a unique cognitive phenomenon. All cognition may be social, but (*anthro-*)social cognition involving present persons still merits its own disciplinary investigation.

**Chapter 6 – Finding the**  
**Social within the Social: a**  
**New Framework for**  
**(Anthro-)Social Cognition**



## **Finding the Social within the Social: a New Framework for (Anthro-)Social Cognition**

### **1. Introduction**

As is hopefully evident from earlier chapters, ‘social cognition’ is a hugely exciting and exponentially growing field of research that is benefitting from the welcome trend for inherently interdisciplinary work. Studies of society and culture no longer need to be confined to humanities, economics or anthropology, and studies of cognition no longer need to be confined to psychology or neuroscience. The inextricably connected nature of these subjects is gradually becoming more and more appreciated, and biologists, computer scientists and theologians are making further contributions to the melting pot of ideas. Social cognition, perhaps more than any other relatively ‘new’ field of research, has been embraced by scientists and theoreticians from across all of the aforementioned fields. Yet there is still a pressing need for greater interdisciplinary cohesion – a need that is often overwhelmed by the vast diversity of ‘social’ events. Consequently, in many cases, social cognitive research suffers from the affliction of running before it can walk (see chapters 3 and 4). In response to this, I intend in this chapter to take a step back and address the *philosophy* of social cognition; that is, how we should delineate social cognitive events and how social cognitive studies should proceed.

Importantly, such philosophising on social cognition will take place within the world of ‘all cognition being social’ – in virtue of cognition involving interaction with social normative structures – that was presented in the last chapter. This overarching social world is the phenomenological bedrock from which I believe all cognitive studies should emerge. Within this, I will henceforth take ‘social cognition’ to be *the transformation of cognising that unfolds when human agents are immersed in the activity of a physically co-present other (or others)*. Not only will this delineation provide us with a robust domain for answering the question of ‘how we understand others’, but it will also, amongst other benefits, serve the purpose of parsing instances of humans thinking *with* and *about* other present humans from all other forms of cognition.

To clarify, across this and the previous chapter I am putting forward the claim that there are two interpretations of ‘social cognition’. The first interpretation of social cognition encapsulates *all* human cognition, due to our co-constitution with social normative structures (see previous chapter). Naturally, however, there is no theoretical benefit in applying the title ‘social cognition’ to all cognitive phenomena. Therefore, henceforth I will only use the phrase ‘social cognition’ to describe events in which agents experience a transformation of their cognising due to immersion in the activity of co-present others (this is the second interpretation). Accordingly, ‘cognition’ will henceforth be taken to designate the ‘all-cognition-is-social’ notion that was explicated in the previous chapter.

Whilst my definition of social cognition will hopefully become clear throughout the chapter, there are five points that I will swiftly elucidate further. Firstly, to qualify as a ‘human agent’ one must be appropriately autonomous, embodied and ensocialled, so as to partake in the normativity of humanity (there will be more on this in section 4, but the idea is present already in chapters 2 and 5). By ‘ensocialled’ I mean that an agent is always and irrevocably social. Unlike the terms ‘socialised’ or ‘enculturated’, which suggest the assimilation or transformation of an agent within a specific socio-cultural domain, ensocialled is intended to convey a fundamental feature of one’s existence. That is, an ensocialled agent does not grow *into* the social world, nor can she shed her social nature by living as a hermit; instead, an ensocialled agent is a constitutively social being.<sup>[34]</sup> With this in mind, it will be shown throughout section 4 that being ensocialled involves an intersubjective normative dimension which persists even in private life. Importantly, to be ensocialled is integrated with being autonomous and embodied. Autonomy is a fundamental feature of subjective experience and is a necessity for an agent to engage with the relevant normative spheres that allow one to cognise through interactions with, or immersive observations of, others. Embodiment is a necessity in that one’s body is integral to social existence; conversely, as we will see in the next chapter, one’s ‘ensocialment’ is integral to bodily existence, as one’s bodily expression is socially dependent.<sup>[35]</sup> To cognise as a human agent is thus to engage with social normative structures as an autonomous, embodied and ensocialled being.

The second point to elucidate is that the notion of ‘immersion’ is intended to capture an existential ‘dwelling’ in another’s activity, such that one’s experience is that of a ‘we-self’. As I will demonstrate in section 5, a we-self takes hold when one’s experience is actively structured by another’s physically co-present activity, such that a self-other domain of normativity is generated and modulated. When acting as a we-self one is psychologically predicated on the other’s current and unique activity.

Thirdly, the rhetoric of the chapter is primarily aimed at social cognition during ‘interactions’ between agents. However, ‘interactions’ are not limited to face-to-face

---

<sup>[34]</sup> To exemplify how an agent can be ‘ensocialled’, consider that the fictional characters of Robinson Crusoe (from Defoe’s (1719) eponymous novel) and Robert Neville (from Matheson’s *I Am Legend* (1954)) may be considered ‘de-socialised’, as they are completely isolated from other humans, but they are still ensocialled (and always will be) because they continue to live (and think) in accordance with recognisably anthro-social norms. Ensocialment is thus a fundamental feature of human existence (see section 4.1); it captures the fact that any first-person cognitive account is brought to fruition through a relation to social normative structures.

<sup>[35]</sup> The enmeshed nature of ‘body’ and ‘social’ that I am proposing naturally invites questions regarding the possibility of disembodied social interaction (e.g. through teleconferencing or instant messaging). Whilst it is certainly possible to ‘socially interact’ to an extent during a teleconference, phone call or real-time text conversation, these interactions are lacking an important social dimension in virtue of the body’s absence. As Dreyfus (2008) points out, a bodily co-present interaction involves a subtle alliance of gestures, posture and eye and head movements, along with sensitivity to vocal tone, physical appearance and contextual sights, smells and sounds; such subtleties simply cannot be broken down into context-independent media channels of visual images, audio and haptics (p. 56). The body thus constitutes a means of context-dependent ‘background’ communication during any social event. Moreover, as embodied agents, the body is the primary means for attaining an ‘optimal grip’ on the world (Merleau-Ponty, 1945/2012), which provides each of us with “our sense of the reality of what we are doing and are ready to do” (Dreyfus, 2008, p. 70). With the body missing from a social interaction, we thus have a diminished sense “of our power and of our vulnerability to the risky reality of the physical world” (ibid.), which, in social interactions, involves our power and vulnerability in relation to others.

encounters but are to incorporate instances of *immersive observation*, in that an agent who is immersively observing (and thus cognising about) another present agent is interacting with the social norms of the other's behaviour (see section 6). What matters is not so much the type of social engagement, but the immersion in a unique domain of self-other normativity. The significance of this is that the framework for social cognition that I am putting forward provides a robust concatenation between social interactions and some social observations.

Fourthly, any reference to 'present' or 'co-present' others should be read as physical (i.e. bodily) co-presence. This is not to deny the significance of social interactions that do not involve bodily co-presence (see footnote [35]), but to highlight the phenomenological importance of bodily presence in social situations (see, in particular, sections 4.2.3 and 6 for expositions of the modulatory capacity of the body within interactions).

The final point to note is that my definition of social cognition should not be considered a formulation of a theoretical competitor to 'mindreading' theories, or the likes of we-mode cognition and participatory sense-making. Rather, the definition derives from the normative *framework* for social cognition that I will be presenting, with the motivating thought being that many theories of social cognition can benefit from this framework. For instance, any theory which strives to address *human* social cognition will require the kind of framework that this chapter presents in order to account for the unique normative connection that is manifest during human-human engagements. As we will see in section 3, without normative framing, social cognition theories will be prone to misdiagnosing the social nature of the phenomena that they seek to explain.

My plan for the chapter will proceed as follows: in the next section, I will consider what it is that we want from a theory of social cognition; in section 3, I will outline the broad problems that plague current theories of social cognition; sections 4 and 5 will occupy the bulk of the chapter as I put forward new criteria for a satisfactory study of social cognition (namely, the criteria of *fundamentals of intersubjectivity* and *immersion*). Section 6 will briefly explain how these new criteria for social cognition provide common ground for various social interactions and social observations. I will conclude by describing the advantages of grounding a theory of social cognition in the philosophical framework that I put forward.

## **2. What do we want from a Theory of Social Cognition?**

As noted in the previous chapter, the many interpretations of social cognition lead to claims that it should involve a vast and varied range of phenomena including the behaviour of ant colonies, wolf packs, ape communities, domesticated pets and their owners, traffic jams and empathy between humans (Steiner & Stewart, 2009, p. 528). Divergence of this kind is largely due to the vague view of social cognition as a theory to capture "any cognitive

process that involves conspecifics” (Blakemore, Winton and Frith, 2004, p. 216). Even when it is theoretically confined in accordance with mechanisms or dynamic processes that are allegedly involved (e.g. as in *participatory sense-making* (De Jaegher & Di Paolo, 2007)), there is still regular ambiguity regarding the separation of social cognitive events from ‘non-social’ cognitive events (see chapter 4 and the next section). Fuelling such ambiguity and vagueness is, in part, confusion over what exactly we should want a theory of social cognition to explain.

What we should *not* want is for a theory of social cognition to focus only on the deliberative individualistic deployment of certain cognitive mechanisms to make inferences about others; as we shall see in section 4 (and as described in chapters 1 and 2), humans are not de-contextualised entities, but are intersubjectively open to one another in a fundamental manner that is not characterised by deliberative inferences. Nor should we want a theory of social cognition to focus only on the emergence of supra-individual dynamics, in which the concept of an individual is reduced to that of a subsystemic component within a novel organisation. Instead, as we shall come to see, there is a middle path to be taken between the purely individual and purely social perspectives, and ‘individual’ and ‘social’ should not be viewed as dichotomous poles.

In many ways, as Gallagher (2009, 2010) points out, the question that social cognition should address is simply ‘how do we understand other persons?’. Yet this is not as straightforward as it seems. For a start, the question can be construed as individualistically biased, in the sense that it attends to subjects (‘we’) targeting their subjective understanding towards nondescript ‘others’. At the least, the question seems to imply a unidirectional approach to social cognition (i.e. from subjects to targeted others), whereas it was demonstrated in chapter 2 that most social encounters have an implicitly reciprocal nature. An alternative would be ‘how do persons understand one another?’. Yet this, too, could be criticised for implying that social cognition always involves some sort of collaborative pursuit of understanding. Gallagher’s initial question may thus be the best we can hope for without making the question of social cognition a hair-splittingly laborious one (which would no doubt result in more restrictively targeted theories (see section 3)). However, for the understanding of others to be appreciated as occurring within a range of normatively immersive interactions, observations and contexts, I believe that two implications of Gallagher’s question need to be explicitly drawn out. Firstly, in responding to ‘how do we understand others?’, one must define *who* both the ‘we’ and ‘others’ are. That is, there must be consensus regarding the ontological nature of the subjects that social cognition is addressing. It is this issue that I turn to in section 4, demonstrating that the unique connection amongst humans justifies a disciplinary demarcation of *anthro-social cognition*. Secondly, the ‘how’ within Gallagher’s question must not restrictively address a specific sub-

personal or personal mechanism that fails to appreciate the many forms of social engagement in which we come to understand others, nor must it ignore the qualitative importance of our understanding of others. As we will see below, there are many social situations in which we understand others and the dynamics of these situations vary hugely in accordance with *who* is involved. By appreciating both the ontological make-up and situational variations of human social behaviour, a more acceptable framework for a theory of social cognition can be formed.

With the question of ‘how do we understand others?’ as motivation, I believe the correct place to start when defining a framework for social cognition is with the phenomenological transformations that take place when directly engaging in thinking with or about co-present others. Such transformations will emerge during an extensive range of social engagements, encompassing diverse phenomena that include: face-to-face encounters and the subtleties involved therein, such as the cognitive divergence in interacting with a loved one as opposed to a stranger; the mimicking of the actions of an observed agent (think of football spectators simulating the heading of a ball from their seats, even though they are not actually on the pitch and heading the ball); empathising with others interactively and observationally (a notion which will be revisited in section 4.2.1); the ‘crowd cohesion’ of agents sometimes acting with striking harmony in an ostensibly spontaneous manner; and the manner in which certain behaviours seem to be socially contagious. My belief is that if we wish to address our understanding of others then these kinds of intuitively graspable phenomena should be the bread-and-butter explananda of a theory of social cognition – not only *why* do these everyday phenomena occur, but *how* do they come about? Motivating this belief is the idea that prior to progressing to any technical or in-depth philosophical investigation of social cognition, our common-sense understanding of the fact that there is something unique about our thought processes when *interacting with* (or *immersively observing*) others should not be ignored.

By pursuing this phenomenological bedrock to a theory of social cognition, one immediately uncovers certain benefits. For example, a phenomenological perspective rids us of the dangers of cognitivistic-style individualism (in which, amongst various issues, individuals’ cognitive states are de-contextualised and solitarily closed off from those of others), whilst also allowing us to take a fundamentally subjective stance on the issue. This is because phenomenology is both inherently concerned with studying the structures of first-person experience and appreciates that this experience is grounded in aspects of others’ existence. For phenomenology, ‘self’ and ‘other’ are, for the most part, united as constituents of first-person experience. Moreover, phenomenology entails an immediate appreciation of the ‘persons’ who are engaged in social cognition, in virtue of being concerned with subjective *experience*. As the experience of all animals fundamentally differs, *human* social

cognition deserves delimitation from other forms of social cognition. This may seem anthropocentric, but we will see in the subsequent sections that such delimitation is justified by the fact that there is clearly something unique taking place when humans understand other humans, as opposed to non-human entities. One could also turn to the fact that the human mind is the archetypal object of study for cognitive science and so any form of *social* cognitive science should be addressing the social engagement of such minds, which, on a straightforward sense of ‘social’, would involve engagement with other like-minds.

With a phenomenological bedrock in place, we will be able to establish robust foundations on which a theory of social cognition can be built.

### **3. A Major Problem for Current Theories of Social Cognition**

As we saw in chapters 3 and 4, current theoretical approaches to social cognition can be broadly separated into two camps: ‘mindreading theories’ (also called ‘theories of mind’) and ‘interactive theories’. In sweeping terms, the former camp is troubled by its focus on agents making predictions from a detached, observational viewpoint, such that social cognition is treated as a phenomenon that is abstracted from the bodily, affective and socially constitutive aspects of being human (for a range of diverse criticisms of mindreading theories, see Fuchs and De Jaegher (2009), Gallagher (2001, 2008), Gallagher and Varga (2014), Gangopadhyay and Schilbach (2012), Di Paolo & De Jaegher (2012), Przyrembel, Smallwood, Pauen and Singer (2012), and Schilbach (2014)). In similarly sweeping terms, the latter camp is troubled by its focus on the dynamics between subjects, which results not only in any notion of an individual-in-interaction being lost, but often also in the loss of the possibility of cognising in a genuinely ‘social’ manner from a spectatorial point of view (see Bohl and van den Bos (2012), Hutto (2009) and Koubová (2014) for criticisms and developments of interactive theories).

However, perhaps more fundamentally troublesome than any approach-specific problems is the fact that neither mindreading nor interactive theories successfully explain why there is something phenomenologically special about humans socially engaging with other humans. For example, Abell, Happé and Frith (2000) and Castelli, Frith, Happé and Frith (2002) have suggested that we are capable of deploying mindreading capacities to the observation of non-human entities that are as simple as animated polygons (specifically triangles). If we can mindread animated triangles in the same way that we can mindread humans, then the mere activation of mindreading mechanisms seems insufficient in accounting for what is truly *social* about this kind of cognition. Similarly, one can easily entertain the possibility of sense-making in a participatory manner, or cognising in the we-mode with/about inanimate entities (see section 4.1 in chapter 4 and section 4 in chapter 3). Why, then, are these interactive approaches claimed to be theories of *social* cognition? Neither mindreading nor

interactive theories seem to want to address the fact that other humans occupy a unique position within human cognitive life. Yet we have seen in chapter 2 and we will see further in the next section that *as a matter of fact* other humans import phenomenological significance that non-human entities cannot (currently) match. Moreover, once we have accepted the undeniable fact that humans occupy a unique position within human cognitive life, it seems that due attention must then be paid to the *human subject* itself. Doing this involves ensuring that social cognitive mechanisms remain grounded within the invariant structures of subjective experience, so that the study of social cognition is not a science that is abstracted away from human involvement, but is properly integrated with a first-person understanding of our lived world (this idea will be revisited shortly).

Let us now briefly consider how common it is amongst social cognitive research to ignore the unique ontological presence of other humans. Firstly, in the field of A.I., Chaminade and Cheng (2009) have recently put forward examples that allegedly show that humanoid robots – robots that have been created to physically resemble humans – can “provide testbeds for hypotheses pertaining to natural social interactions” (ibid., p. 287). They reach their conclusion that robotics can be used to enhance studies into social cognition on the evidence of a humanoid robot that executes a certain action being sufficient to elicit activity (“resonance”) in the same neural structures that are activated when observing a human participant execute an analogous action (ibid.). Whilst I am strongly sympathetic to their conclusion, it is only a very narrow cognitive ability (what we could call the ‘resonance ability’) that is being tested, and it is far from clear how significant we should consider this ability in a full-blown account of social cognition that addresses how we engage in, sustain, modulate and are modulated by social interactions and observations. The starting premise of Chaminade & Cheng’s research is of testing a precise and abstractable cognitive capacity; they thus focus only on a narrow similarity between our neural response to humanoid robots and our response to humans, rather than focusing on the broad differences, such as the intuitively graspable transformations that take place when cognising with and about other humans. It would be an obstinately anthropocentric stance to take if I were claiming that robots are theoretically incapable of playing a role analogous to that of humans within my framework of social cognition; however, the point is that at this moment in time, robots (and any other non-human entity) can seemingly only contribute to very narrow cognitive criteria within social situations, such as basic interactive coordination or involuntary neural resonance. The wide spectrum of intuitively graspable intersubjective subtleties (which will be discussed in greater detail in section 4) are simply untouched. Indeed, both of the authors later observe in a further study that the human brain regions which are activated when making inferences about other humans are significantly different to those that are activated when making inferences about humanoid robots or random computer agents (Chaminade,

Rosset, Fonseca, Nazarian, Lutchter, Cheng and Deruelle, 2012). Thus, whilst there is nothing strictly *wrong* with Chaminade & Cheng's (2009) conclusion, I believe their work is suggestive of the kinds of theoretical abstractions and lacunae that are too frequently left unbridged in studies that are supposedly central to our *social* cognition.

In a similar vein to Chaminade & Cheng (2009), Tylén, Allen, Hunter and Roepstorff (2012) have recently assessed fMRI and eye-tracking results of participants' brain activity in both social observations and social interactions. Once more, I am sympathetic to their conclusion that different brain regions are activated during interaction in comparison to brain activation when observing an action (*ibid.*). However, the participants were only shown video clips, rather than observing or interacting with actual people. We can safely assume that several of the phenomenological transformations that we experience when interacting with and directly observing co-present others will be missing when one is engaging with pre-recorded and fictional video clips. There is thus another significant leap in what is being treated as *socially* relevant. Whether interacting with a humanoid robot or video clips, it seems prudent to be wary of empirical studies that fail to address the social complexities that are involved when cognising with or about other humans who are physically present.

Of course, I am not claiming that studies such as these have no value in contributing to our understanding of social processes – looking at subsets of social cognitive abilities can obviously yield a wide array of benefits. The point is that by focusing on cognitive subsets such as robotically generated neural resonance or neural activation from observing video clips, the scientific abstraction that the studies seek is a move away from human subjects and, crucially, the phenomenological impact of subjects engaging with (and understanding) others. A piecemeal view of social cognition is thus delivered, with the central role of human nature being obscured.

One final example should suffice to demonstrate this tendency of social cognitive analyses to discount the unique impact of human nature. In a special 'Perspectives on Social Cognition' issue of *Cognitive Systems Research* (9, 1-2), Marsh and Onof (2007) claim in the 'Introduction' that the question a social cognition theorist must address is "How does one apportion the extent to which individuals' cognitive states are dependent upon their social milieu?" (p. 2). On the one hand, this seems like a reasonable approach and also justifiably broad in order to welcome the diverse approaches to social cognition that the editors were clearly seeking. However, there are two fundamental errors in this summarising objective. Firstly, it is never a *single* social milieu that an individual's cognitive states depend on, but *social milieux*. Any person is fundamentally constituted by numerous social groups – cultures, families, occupations, gender, friendship groups and generations, along with many more – and these comprise the milieux which may come to bear, in any number of combinations, on social situations. Referencing a "social milieu" could be construed as *too*

broad and vague, in the sense that one could simply contend that individuals' cognitive states are always dependent on sociality (thereby resisting any attempt to consider dependence on more specific social happenings). Secondly, it is not just the *extent* to which individuals' cognitive states are dependent upon their social milieu that matters, but also *the manner in which the individuals' cognitive states are dependent upon their social milieu*. That is, a social cognitive analysis should not just be a quantitative assessment of which social milieu are predominantly involved in an interactive or observational social situation, or which cognitive mechanism is activated; it should also be a qualitative assessment of how another person, or a given situation, can induce specific qualitative experiences within a subject. As noted in section 2, even if we replace Marsh & Onof's question with Gallagher's (2009, 2010) simpler one – 'how do we understand others?' – there is still the need for the 'how' of our understanding to be appreciated from a qualitative perspective.

Recent work in *neurophenomenology* provides a radical indication of how such qualitative assessment can be effectively integrated with third-person empirical data. The premise of the neurophenomenological approach is "to marry modern cognitive science and a *disciplined approach* to human experience" (Varela, 1996, p. 330) by integrating expert first-person reports with third-person data in a mutually illuminating and co-determining exchange. A study by Lutz, Lachaux, Martinerie and Varela (2002) put this approach into action by using phenomenal categories identified by subjects to establish neurodynamical patterns during a simple visual task. Their findings highlight that the integration of first-person experience (i.e. the identification and reporting on phenomenal categories of preparation and perception) and third-person data (i.e. electroencephalogram neuro-images) can reduce "opacity" in neural responses and can lead to the detection of "original dynamical categories of neural activity" (Lutz and Thompson, 2003, p. 42). Applying this methodology to social contexts could lead to an illumination of the cognitive mechanisms that are called into play during the phenomenological transformations that take place when interacting with or observing others.

However, neurophenomenology remains the exception to the rule, and the reluctance to seriously consider phenomenological reports leads to a continuously abridged study of cognition. Certainly within social cognitive research it is notable that even the purportedly more radical and liberal interactive theories (see Gallagher (2012) and Overgaard & Michael (2013) for recent discussions) retain a tendency to overlook the fundamental experiential significance of another human's ontological involvement. As mentioned earlier, for example, one of the flaws with the theory of we-mode cognition is its failure to account for why "co-representations" are induced in a human agent only by the collaborative presence of another human, rather than by some non-human entity (see chapter 3, section 4). Similarly, the theory of participatory sense-making doesn't explicitly clarify why sense-making with

another autonomous agent differs from sense-making with inanimate entities (see section 4.1 in chapter 4). These failures seem to highlight the need to take seriously the phenomenological transformations that are intuitively evident in engagements with other persons.

By elucidating some of the *fundamental intersubjective qualities* that a human agent imports to a social situation and then considering how agents become *immersed* into novel intersubjective organisations, I believe that a more thorough and robust framework for social cognition can be created.

#### **4. New Criteria for a Theory of Social Cognition**

##### **4.1. The Importance of Human Ontology**

Thus far, I have alluded to the ‘intersubjective subtleties’ or ‘fundamental intersubjective qualities’ (henceforth, FIQs) that a human agent brings to a social situation. These ‘qualities’ are bound up with the unique ontological status of being human and I will shortly detail some of them (section 4.2). Firstly, however, I will briefly explain why I believe these FIQs matter.

To ignore the significance of a human’s presence in a social situation is to miss out on what makes the situation *social* in the first place. Sociality is not just a matter of being able to interact, as is possible with other (non-human) animals or artificial machinery, nor is it simply a matter of conspecificity. It is instead a matter of *ontological connectedness* in virtue of belonging to and living through the same normative world (see section 5 of chapter 2 for more on this). Being ‘human’ is about more than merely imitating human qualities – it includes being *ensocialled* in a network of norms that are communally shared with other humans.

It should be noted that, for some philosophical accounts, the notions of embodiment and autonomy already comprise this idea of being ensocialled (e.g. Battersby, 1998; Kyselo, 2014; Merleau-Ponty, 1945/2012; Varela et al., 1991). The link that these accounts share is the idea that a bodily human “is not only saturated by the social, but is entirely inconceivable without it” (Kyselo, 2014, p. 2). That is, for humans, being embodied simply *is* being social and to be fully (en)social(led) necessarily requires having a body. What these accounts should also emphasise, however, is the normative dimension of the social. To be human at all is to modulate and be modulated by normative structures that are shared with others. Put another way, the claim is that the nature of any human is reciprocally amalgamated with the nature of other humans, with all humans contributing to a collective normativity, which, in turn, normatively canalises individual activity. Crucially, this view should not just be a concern of metaphysical philosophers. The fact that we are bound to collectively (self-and-other) generated norms (see McGann & De Jaegher (2009)), and thus bound to *being-with-others*

(see Heidegger (1927/1962) and Haugeland (2013)), is integral to how we cognise. As I will explain in section 5, the self-and-other generated norms that emerge in interactions with (and observations of) other present persons provide the grooves along which our cognitive processes flow. Thus, to be human is to be part of socially generated normative structures and these social normative structures constitutively guide our cognitive processes. To study the social cognition of humans without considering this aspect of human nature will, at best, confine us to a skeletal and abstracted view of cognitive mechanisms. Conversely, by properly considering human nature, we can root social cognition in our fundamental intersubjectivity and the associated normative structures which mould our cognition. In this way, a skeletal, mechanistic theory of social cognition can be fleshed out with a more complete understanding of the agents to which the theory applies.

What's more, as I have demonstrated in section 3, many of the problems with current theories of social cognition arise from the fact that empirical work is targeted at restrictively specific social capacities whilst other, more general and subjectively graspable capacities are largely ignored. By focusing on the normative nature of humans, one has a framework from within which many diverse social capacities can be studied in a common currency. That is, if one wishes to focus on the influence of transcendental conditions on social cognition, then these conditions can be explained in terms of co-existing normative structures. At the same time, if one wishes to focus on neuroscientific mechanisms then these, too, can be explained in terms of sub-personal correlates of normative structures. The benefit is thus that an overarching normative framework can guide the amalgamation of sub-personal and phenomenological social cognition theories by translating diverse approaches into normative terms. Of course, one could claim that existing frameworks can already accomplish this. For instance, folk psychology provides a 'common currency' for diverse approaches to social cognition. However, frameworks such as folk psychology are already entrenched in existing literature and empirical work, such that, for example, folk psychology is closely tied to mindreading theories. As such theories have been shown to be flawed, this established connection is detrimental to the strength of the folk psychological framework. The normative framework that I am espousing thus provides a clean slate, away from problematic theories of social cognition and towards a phenomenologically sensitive stance. As we will see in the next chapter, a normative framework also provides a robust picture of selfhood that neither reduces our self-constitution to some 'internal' quality nor inflates it to a supra-individual 'group' level.

#### 4.2. Human Nature and Fundamental Intersubjective Qualities

When I interact with another human, I am not merely making theoretical inferences about their cognitive states, nor am I merely making simulative inferences, nor am I merely caught

up in the momentum of coordinated movements and expressions. Amongst all (or any) of these cognitive skills, I am also sensitive to the ontological specifics of the person with whom I am interacting. Whilst each person is obviously unique, there are certain FIQs which are generalisable to the nature of all humans. These qualities are 'fundamental' in that they are pre-noetic; they are rudimentary constituents of how one relates to the world and thus of what it is to exist as a human. To lack these FIQs is to exist in a non-human way. Below are some of the most significant of these qualities.

#### 4.2.1 A Sense of Otherness

As discussed in chapter 2, prior to any engagement in an interactive or observational act of social cognition – both of which are crucially constituted by a normative component, as we will see in sections 5 and 6 – one must firstly relate to another as an agent to whom one can meaningfully attribute mental states or with whom one can meaningfully interact. That is, one must relate to the other's autonomous agency in such a way that there is an acquiescence towards the other's capacity to generate an observational or interactive connection. Incorporated in such acquiescence is a registration of the other's capacity to generate and modulate social normative structures (again, there will be more on this in section 5). Relating to the other in this way and acquiescing to their autonomous agency are not deliberative or representational acts; they are non-reflective, instinctive transformations in one's experience, which result as a phenomenological consequence of the other's meaningful existence as a human agent. Once such acquiescence is manifest, one is disposed towards the other as someone to whom one can attribute mental states or with whom one can meaningfully interact. In other words, one is imminently attuned to the presence of another in virtue of the phenomenological transformation in one's world that is induced by the other's presence as a being with whom one can forge a unique normative connection. This aspect of a being's nature is what I will refer to as *otherness*, and awareness of it involves a *sense of otherness*. Thus, in order to properly understand another, one must primordially experience his or her *otherness*.

One may worry that this preliminary description of 'sensing otherness' parrots the role that some social theorists believe *empathy* plays in human life. That is, empathy, which, broadly speaking, is the ability to experience the feelings (and, therefore, part of the cognitive activity) of others, is thought to facilitate any cognising about or with present others. In more precise terms, the phenomenological tradition sees empathy as the projection of oneself into some external entity, such that the entity is suffused with one's own existence (Lipps, 1909, pp. 224-225; Zahavi, 2014, p. 104; cf. Rameson, Morelli and Lieberman (2012) and Walter (2012) for neuroscientific approaches to empathy). However, not only is empathy a concept that is open to numerous interpretations (see Engelen and Röttger-Rössler (2012) for an

overview), but it's also not clear that it provides the kind of complete picture for social cognition that is currently being sought. For example, just as mindreading and interactive theories of social cognition fail to explain that prior to any engagement in an observational or interactive act of social cognition one must firstly relate to another as an agent with whom one can meaningfully interact or to whom one can meaningfully attribute mental states, 'empathy theories' fail to explain that one must firstly relate to another as an agent about whom one can empathise. A *sense of otherness* thus fundamentally underlies one's ability to mindread, interact or empathise.

In order to better understand this sense of otherness, it is perhaps easiest to firstly consider it in action. For instance, think of the shifts in feeling that one has when thinking that there is an observer present, only to realise that it is in fact a waxwork or model person (Ratcliffe, 2013b). Similarly, think of the shock one receives upon hearing a noise that is indicative of another person, such as a creaking floorboard or a slow exhalation of breath, when in a room that one assumed was empty. In such scenarios there is a discernible phenomenological movement in one's conception of the world, from a kind of 'bare' subjectivity to one that is coloured by the co-presence of another. This phenomenological movement recalls the qualities of *relevance*, *re-configuration* and *reification* that the physical co-presence of others imports to a situation (see chapter 2). Our sense of otherness encapsulates these phenomena, rendering one uniquely disposed to cognise with or about another from the foundation of the (transformative) phenomenological connection that is forged through one's co-presence with the other. That is, in sensing otherness, one is drawn into an experiential domain that is, in part, modulated by the other's presence. One's actions in such a domain are no longer those of a solitary 'I', but of an 'I *before this* other'. In keeping with the exposition of chapter 2, a phenomenological connection of this kind is manifest in the incipient moments of a social encounter and as such one's sense of otherness can be considered as a conceptually *prior* phenomenon to full-blown engagement in a social cognitive act. As I will explain in section 5, to socially cognise in the proper sense is to be *immersed* in an interactive or observational engagement with another such that the further FIQs of *status* and *pedagogy* are also implicated.

To almost all intents and purposes, the sense of otherness aligns with what Ratcliffe (2013b) has described as a "sense of personhood": "a primordial "sense of others [which...] is inextricable from a distinctive way of experiencing *possibilities*" (ibid., p. 221). However, I favour the phrase 'sense of otherness' for two reasons. Firstly, Ratcliffe (2013b) declares himself "agnostic" towards the possibility of non-human persons (p. 222), whereas I believe only humans can sense the otherness of other humans. This is not to deny the *theoretical possibility* of non-human beings having a sense of otherness; it is simply that non-human animals, robots or aliens would need to be established as participants in the normative

background of humanity if they were to induce the relevant phenomenological transformations in a human agent. Sensing otherness is not merely a question of sensing autonomous animacy, but also sensing the other's potential for interactive and observational engagement in virtue of belonging to the collective normativity of humanity. Secondly, Ratcliffe (2013b) is concerned only with our primordial "sense of the personal" (p. 226) and not with how this primordial sense of others diversifies during social cognition in conjunction with the other FIQs that are unique to humans. Thus, whereas Ratcliffe's 'sense of personhood' is focused on the phenomenological transformation of mere encounters with other persons, my 'sense of otherness' is intended as a foundational layer within a social cognitive account of how we interact with and observe others.

Arguably, this sense of otherness is the most important of the FIQs that I will discuss, in that it facilitates any subsequent interpersonal engagement and social cognising. Indeed, one could claim that the further FIQs I will go on to discuss – *status* and *pedagogy* – could be incorporated into the notion of a sense of otherness. However, I have refrained from doing this as I want to highlight the significance of specific human features in any given situation, not just the generalisable significance of being human.

#### 4.2.2. Social Status

In the vast majority of instances of social engagement and in all instances of social cognition, we do not merely recognise otherness, but also remain sensitive to the unique manifestation of otherness. That is, once we have non-reflectively appreciated another as a generalisable 'someone', we then appreciate them *as someone*. Rather than encountering some sort of universal other, the persons with whom we socially engage in cognitively relevant ways are always unique and thus generate unique cognitive operations in accordance with their biosocial embodiment. Whilst the cognitively unique individuation of each person is peripheral to a general theory of social cognition, there are certain agential characteristics – what I will henceforth refer to as *status traits* – which provide the 'tracks' along which cognition unfolds and thus should not be ignored by cognitive disciplines. For example, when an agent engages with another, it is not simply a mass of physical flesh that is being engaged with; the other's body is a physical manifestation of biosocial traits such as gender, age and ethnicity, along with, arguably, traits such as health, motor coordination (including athleticism and strength), confidence and social class (although this is perhaps more dependent on the body's context than just the physical body). It is these immanently conveyed social aspects of our embodied nature that I intend 'status traits' to capture.

On a first reading, this claim regarding the import of status-trait-defined cognitive 'tracks' may seem to harmonise with traditional approaches to social cognition, none of which would deny that cognitive content and/or mechanisms can vary due to the persons involved.

However, part of the motivation for defining a new framework for theories of social cognition is that current approaches are generally afflicted by a common methodological issue; namely, that they disconnect cognitive processes from the lived experience of cognising subjects. This means that agential characteristics, which are cognitively essential, are encountered either (a). from an individualistically biased perspective, taking the scientific starting point of a de-contextualised individual endowed with specifically appropriate cognitive apparatus, or (b). from a group-biased perspective, in which first-person experience is largely lost to the cognitive mechanisms of emergent relational dynamics. The perspectives of (a) and (b) then involve the application of their implicated 'default' mechanisms (such as theorising or simulation) to various incarnations of 'other'. My argument is that culturally recurring status traits should be integrated into the invariant experiential structures which science strives to investigate, such that status traits support and enhance investigations into cognitive content and mechanisms. In other words, agents should not be treated as biosocially 'blank' (i.e. without gender, ethnicity, age or cultural belonging), but as *ensocialled* beings who import cognitive significance (for self and other) through their lived presence.

On an initial impression, the cognitive significance of status traits may seem entirely straightforward, in the sense that the specific cognitive transformation a given agent undergoes when directly engaging with another present (*ensocialled*) agent is bound to depend on the *status* of each agent (i.e. *who* the agents are). But what exactly can be said about the impact status traits have on social cognitive events? In order to answer this question, let us consider an everyday scenario of an individual entering a café in which several strangers are already seated. In thinking about this entrant to the café, an orthodox cognitive scientist would claim that the seated patrons would make theoretical or simulative inferences, perhaps along the lines of 'I believe this person desires a drink or snack' (if the entrant looks at and heads towards the counter), or 'I believe this person is looking for a friend' (if the entrant looks around at the seated persons). Similarly, a cognitive scientist of an interactionist persuasion may claim that the seated patrons can directly perceive, in a non-inferential manner, that 'this person desires a drink or snack' or 'this person is looking for a friend'. On a superficial reading, these both seem like reasonable claims to argue for. However, both perspectives work from the initial assumption that the entrant to the café is a 'standardised' person. That is, both inferential and non-inferential explanations assume an application of cognitive processes to a *generic person*. But of course social cognising does not cater to such superficiality; *who* the person is directly impacts *how* they are thought about. And this is not just a case of variations in the content of the observers' cognitive states. A stranger's behaviour – even simple behaviour – may need to be inferentially assessed due to an agent's unfamiliarity with this other's normative comportment, whilst a

loved one's behaviour may be immanently graspable in a non-inferential manner. Whether or not a cognitive inference is needed may thus depend on *who* the café entrant is (i.e. the *status traits* that the entrant conveys). As Goffman (1956, pp. 1-6) explains, as soon as a person enters into a situation – that is, as soon as their otherness has been sensed – there is a projection of information that observers and potential interactors acquire (inferentially and/or non-inferentially). Some of this information will be an intentional effort by the individual (qua entrant) in question, but other information (i.e. the status traits) will be unintentionally conveyed by their embodied character. In both its projection and acquisition, this information helps to define the situation, so that the involved persons are cognitively disposed to one another in a specific manner. From the simple entrance of an ensocialled agent into a situation there can thus emerge “a kind of interactional *modus vivendi*” (Goffman, 1956, p. 4), which is generated by both an ‘actor’ and an ‘observer/interactor’. Importantly, this *modus vivendi*, which is negotiated and modulated by the expressive (and directly or indirectly interactive) behaviour of the involved persons, provides the systemic scope within which the persons’ cognitive states unfold.

A brief survey of empirical work supports the idea that persons’ status traits can alter our cognition. For example, in separate studies, Xu, Zuo, Wang and Han (2009), and Avenanti, Sirigu and Aglioti (2010) have demonstrated how another’s status – specifically, in this case, another’s race – can alter one’s empathic reactions to pain. That is, whereas participants experienced increased anterior cingulate cortex (Xu et al., 2009) or increased corticospinal system responses (Avenanti et al., 2010), both of which align with the feeling of pain (and empathy for the pain of others), when observing an ‘in-group’ same-race subject have his/her hand or face pricked by a needle, they did not experience this “empathic sensorimotor resonance” when observing other-race subjects suffer the same painful stimulus (ibid.; Xu et al., 2009). Similar studies support the same broad conclusion – that we experience more empathy for in-group subjects and less for out-group subjects – when the ‘group’ in question is decided by subjects’ preferences for sports teams (Leach, Spears, Branscombe and Doosje, 2003) or political parties (Combs, Powell, Schurtz and Smith, 2009). In other work, Molenberghs, Halász, Mattingley, Vanman and Cunnington (2012) found that the neural regions that are activated when judging the performances of one’s (arbitrarily assigned) in-group team members during a competitive situation differ from those neural regions that are activated when observing one’s competitors complete the same action. There is also evidence that subjects are better at recognising members of their own ethnicity than at recognising members of other ethnicities (Malpass and Kravitz, 1969; Sporer, 2001). From a more subjective perspective – that is, the traits of one’s own body – there is further evidence that status helps determine how we cognise. Farmer, Maister and Tsakiris (2014), for example, have demonstrated that multisensory-induced ownership of an

‘outgroup’ hand (in their study, a dark-skinned hand) reduces implicit bias towards the implicated outgroup. From a more speculative point of view, one can also consider a recent study that describes how the cognition of German-English bilingual persons is context-bound (when categorising motion events) by the language in which they are operating (Athanasopoulos, Bylund, Montero-Melis, Damjanovic, Schartner, Kibbe, Riches and Thierry, 2014). Once one considers that language conceived of more broadly – including postural gestures – differs across cultures (Argyle, 1975; Matsumoto and Kudoh, 1987; Matsumoto, 2006), one could extrapolate these findings to suggest that the (linguistic) status conveyed by one’s body has a self-modulating role in cognition, along with affecting how one is engaged with by others in interaction.

Aside from these empirical studies, the cognitive importance that I believe a person’s *status* carries can be further expounded by considering the status trait of *gender*. For most people, across the vast majority of global societies, gender is an unchanging and immediately evident aspect of one’s embodied existence.<sup>[36]</sup> Rather than a biological dichotomy, gender is most pronounced as a kind of social normative structure. That is, males will tend to behave and think in certain ways, in accordance with the social norms of masculinity, and females will tend to behave and think in certain ways, in accordance with the social norms of femininity. Such a claim is not intended to paint *all* males and *all* females with distinct brushes, nor should it be considered an unalterable feature of human existence. The point is simply that gender – as a status trait and conspicuous bodily characteristic that is a product of societies in which gendered practices and mores persist – produces differences in cognition and behaviour that are phenomenologically and empirically evident. Consider, for example, findings suggesting that men and women adopt differing cognitive strategies, resulting in the activation of differing brain regions, when faced with tasks requiring creative or generative responses (Abraham, Thybusch, Pieritz and Hermann, 2013; Abraham, 2015). Further studies of gender differences suggest: that neurological activation in men is more “world-focused”, whereas for women it’s more “self-focused”, during affective experiences (Moriguchi, Touroutoglou, Dickerson and Feldman Barrett, 2013); that the same bodily sensations can have a contrary impact on humour experiences for each gender (Kaspar, Jurisch and Schneider, 2015); that young girls tend to rate themselves lower than young boys in self-estimates of strength (Gross, 1968); that women show greater sensitivity to physical pain (Wiesenfeld-Hallin, 2005); and that women generally show lower satisfaction with body appearance (Feingold and Mazzella, 1998). Although physiological differences may have some role to play here, such gender divergence is primarily generated through the

---

<sup>[36]</sup> There are, of course, hermaphrodites, androgynes and transgender persons for whom gender is not necessarily an unchanging or immediately evident aspect of one’s embodied existence. However, such people will still convey a ‘gender status’ in virtue of bodily convergence to gender stereotypes, or in virtue of the androgyny that their body constitutes (that is, their body conveys the trait of being neither/both female nor/and male).

*normativity* that pervades societal conceptions of male and female roles and capacities. As Young (1980) points out in her seminal paper ‘Throwing Like a Girl’, women generally ‘carry’ their bodies in a notably different way to men and will often greatly underestimate their bodily abilities, approaching physical tasks “with timidity, uncertainty, and hesitancy” (p. 142). There is little genetic justification for such a notable divergence in behaviour and it is something that women can transcend under the right circumstances. As Young (1980) explains, however, the feminine modalities that women enact are normatively generated by western societies’ tendency to situate female bodies as *mere bodies*; that is, “as shape and flesh that presents itself as the potential object of another’s intentions and manipulation, rather than as a living manifestation of intention and action” (p. 154). Young’s overall point is that there is a significant phenomenological difference in accordance with gender. It is little wonder, then, that there are gender differences not only in overt behaviour and cognitive strategies, but also in how one *thinks with* and *about* each gender.

With gender differences being enacted, in various forms, across all societies, along with other normatively maintained biosocial status traits such as ethnicity, race and age, it is striking that no theory of social cognition consistently accounts for the cognitive differences that these traits generate. One might argue that I am being too picky in my contention that theories of social cognition – and cognitive science in general – should heed the FIQs of human ontology, such as a *sense of otherness* and *status traits*. As Protevi (2009, p. 30) points out, an opponent could easily argue that a cognitive scientist should not be blamed for abstracting principles of cognition away from a more nuanced analysis of the specific subjective (and intersubjective) qualities that cognisers bring to a situation. However, once one has progressed from the “classical sandwich model” (Hurley, 1998, p. 401) of strict representationalist cognitive science, it seems that there is no choice but to take notice of the many possible ways in which FIQs significantly alter our cognitive states and processes. Consider, for example, that part of the reason why some 4E approaches to cognition (i.e. embodied, embedded, extended and enactive approaches to cognition (see Menary (2010) for an introductory overview)) oppose orthodox representationalist accounts of cognition is that representationalism disconnects cognitive processes from the cogniser’s body and immediate environment, both of which have, at the very least, a scaffolding effect on cognition. Yet once body and environment are rightly given due attention in 4E cognitive studies, it seems strange that they are then readily standardised across the vast array of daily situations and subjects for whom cognition clearly varies. Agents’ bodies and environments bring to bear the status traits of each agent’s unique history and sociality, which are central to how the agents cognise and are cognitively engaged with by others. And, crucially, the cognitive relevance of bodies and environments simply isn’t uniform from one person to the

next or from one situation to the next. Focusing once more on gender, Protevi (2009) explains:

If all concrete subjectivities are gendered – or at least have developed via gendering practices and have to navigate a world where gender matters – can we be satisfied with abstract principles of cognition that ignore gender effects? (p. 31)

The answer is ‘no’. Turning one’s back on bodily manifest biosocial characteristics such as gender is to abstract a mechanistic conception of *the* physical body away from cognitively relevant status traits of agents’ situated bodies. Whilst the detailed individuation of specific agents’ bodies and environments is a concern primarily for psychiatrists and therapists, the omnipresent status traits that environmentally situated bodies constitute cannot be ignored if a full account of cognition is sought. This is especially true for social cognition because of the flexible and modulatory role of bodies in social encounters. Thus, rather than focusing on generalisable and synchronic cognitive processes of *the* human subject in any given situation, there is a pressing need to consider the status-partial and diachronic processes of embodied and ensocialled cognisers.

Some may worry that this call for consideration of the status-partial and diachronic processes of embodied and ensocialled cognisers could lead to an abandonment of scientific *explanation*. However, as has already been alluded to, this is not my intention in putting the spotlight on the FIQs of human ontology. Rather, it is an enhanced and honest form of scientific explanation that I am seeking – one which, for instance, could properly integrate third-person observational data with the phenomenological insights of first- and second-person experience. One approach that is taking steps in the right direction, neurophenomenology, was mentioned earlier towards the end of section 3. An alternative is simply to ensure justice is done to the likes of status traits as bona fide scientific categories that are relevant to both experimental participants (i.e. subjects) and experimental targets (i.e. others). For example, it is well established that cultural differences between Westerners and East Asians result in differing neural processes during the perception of objects and concepts (Goh and Park, 2009), as well as resulting in differing neural responses to evaluative threats (Park and Kitayama, 2012) and even differing neural patterns for the “embodied construction of the self and identity” (Kitayama and Park, 2010, p. 111). Not only could these cultural differences be incorporated into experimental set-ups, such that conclusions are drawn relative to the culture of the participants (and ‘targets’ in social studies), there should also be attempts to refine these cultural differences at micro-cultural levels through investigations into neural differences resulting from regional, contextual and even dyadic variations. I believe it would also be beneficial if first-person reports concerning

culture could be re-inserted into studies as phenomenal categories, as per Lutz et al.'s (2002) neurophenomenological approach. It would obviously be highly challenging to distil motivational cultural factors into phenomenal categories in the same way that Lutz et al. were able to do so for preparatory and perceptual factors in a simple visual task; however, if there were discernible consistencies in first-person reports of cultural determinants in outlooks or decisions, then including these reports in further neural assessments of outlooks/decisions would undoubtedly yield novel insights into cultural neuroscience. Similarly, for social cognitive gender studies, the integration of first-person reports of self- and other-conceptions of femininity and masculinity with neuroimaging data would also undoubtedly produce highly illuminating results. At the very least, carefully considering the neural differences that emerge in tandem with variations in biosocial status traits (encompassing gender, ethnicity, age and culture) will produce a more expansive, detailed and experientially relatable science of cognition.

Put more broadly, if our embodiment and contextual embeddedness is to be taken seriously, then our ensocialment – including the status traits harboured therein – needs to be carefully included in empirical investigations. Through such an approach, a scientific explanation of social cognition will not only be retained, it will also be ameliorated in such a way that it harmonises with our intuitive understanding of the (varying) significance of others.

#### 4.2.3. Pedagogy – the Transmission of Norms

Implicated within the claim that theories of social cognition should heed pervasive *status traits*, alongside a *sense of otherness*, is an idea that has been breached at several points of this thesis: that inherent human connectedness constitutes a normative dimension. In order for this normativity to persist in its generation and regulation, we need a concept that accounts for the transmission of norms across the vast diversity of the human population.

This brings us to the third of the FIQs of human ontology that I wish to expound, that of *pedagogy*.<sup>[37]</sup> In order for our pedagogical nature to be made apparent, one must consider that the normativity which pervades human life and canalises our cognitive processing is not distinct from everyday human behaviour, but is regulated by it, and, for such regulation to persist and be constitutively attended to, norms must be transferred amongst agents. In other words, for normativity (in which our cognition unfolds) to persist, we must *learn* norms from one another. Once again, this is an aspect of human nature that is straightforwardly grasped and related to on a cursory consideration. For example, whilst propositional knowledge is generally gathered via explicit education or, at least, through the

---

<sup>[37]</sup> Pedagogy is to an extent already implicated by the FIQ of status traits. That is, without the (pedagogical) ability to be transferred from one agent to another, no meaning would be immanently conveyed through status traits. In this way, pedagogy could be viewed as a fundamental intersubjective *mechanism* by which status traits are mobilised. However, I also think it is worth considering as a 'quality' (i.e. an FIQ) in its own right, in that any social engagement involves agents embodying roles as (non-reflective) normative teachers and/or learners.

explicit attention of the learner, there is a great abundance of further knowledge – ‘know-how’ – that is implicitly amassed through experience. One is rarely directly taught, nor could many claim to have ever studied, how to execute the nuances of everyday tasks such as how to greet a stranger, how to walk with a ‘normal’ ambulatory gait, how loud to speak across different situations, or how far to stand from someone during a conversation. Such everyday tasks are consistent within societal groups because they are dependent on social normative structures which are digested and dispersed by the (modulated and modulating) expressive behaviour of individuals. Transmitting norms in this way is essential to being human: it provides the glue for our *ontological connectedness* (see section 4.1). The very possibility of having societies and cognitive dispositions of a meaningfully relational form is dependent on the ability “to imitate and borrow information and then pass it on to another by non-genetic means” (Bloch, 2005, p. 7.). This ability is what I mean by *pedagogy*.

The pedagogical nature of adult humans is hopefully evident from the fact that customs, attitudes and outlooks not only vary but constantly change across and within cultures and generations. All cultural transformations – progressions and regressions – result from people learning from one another and assimilating certain behaviours into a collective ‘way of life’ (Haugeland, 1998). However, describing this as a *fundamental* aspect of human nature is perhaps not as immediately evident from a developmental perspective. Yet there is support for the idea that pedagogy is an essential part of human development. For example, Csibra and Gergely (2006, 2009; Gergely and Csibra, 2014) have argued for *natural pedagogy*: a specialised social cognitive mechanism that has been evolutionarily selected to facilitate the “transfer of uniquely human forms of cognitively opaque, generic, and shared cultural knowledge” (Gergely & Csibra, 2014, p. 130). Their claim is that human infants are disposed towards such natural pedagogy during communicative acts in a tripartite manner:

- (i) by being sensitive to ostensive signals that indicate that they are being addressed by communication, (ii) by developing referential expectations in ostensive contexts and (iii) by being biased to interpret ostensive-referential communication as conveying information that is kind-relevant and generalizable (Csibra & Gergely, 2009, p. 148)

These three conditions allow human infants (and, indeed, adults) to acquire new, generalisable information and skilfully deploy it in later contexts, thereby completing a successful transmission of generic knowledge from one individual to another (*ibid.*). The necessity of communication for such pedagogical transmission of information arises because a naive infant employing strictly observational learning strategies would have difficulty in (a). picking out the causal relevance of a knowledgeable action, (b). comprehending the generalisability of the obtained knowledge, and (c). comprehending whether or not the knowledge is widely available to members of a community (Gergely, 2010; Gergely & Csibra,

2014). Communication overcomes these difficulties because a ‘teaching’ individual actively behaves in such a way that the relevant generalisable information is selectively manifest *for* the ‘learning’ individual (ibid.). Thus, there is a demonstrative intersubjective aspect to a ‘teaching’ individual’s behaviour which is secondary to their primary subjective purpose of acting in such a way. Gergely & Csibra (2014) stress that not all communication is of a ‘natural pedagogical’ form and it is also not the case that all teaching is a kind of natural pedagogy. Rather, they contend that it is *ostensive communication* which imports natural pedagogy, drawing support from empirical findings which suggest that infants will ignore seemingly irrelevant actions when imitating causally efficacious means to achieve goals unless their demonstrator makes it clear (i.e. *ostensively communicates*) that they should do otherwise (e.g. Gergely et al., 2002; Brugger et al., 2007; Nielsen, 2006). What’s more, Gergely & Csibra make a compelling case for the developmentally “basic” nature of natural pedagogy by describing the sensitivity of infants to ostensive communication. So, for example, newborns are known to show preferences for a direct gaze (Farroni, Csibra, Simion and Johnson, 2002) from the upright face (Farroni, Menon and Johnson, 2006) of an interactor, as well as preferring ‘infant-directed speech’ (‘motherese’) to adult-directed speech (Cooper and Aslin, 1990). Such preferences suggest that newborns are sensitive to being communicatively addressed and can disambiguate a communicative act from a non-communicative one (Csibra & Gergely, 2009), as well as having the prior sensitivity to disambiguating human entities from non-human entities (Gallagher and Meltzoff, 1996; see Gallagher (2005, pp. 69-85, 225-230) for more on neonate sensitivities). As newborns become young infants, they display the tendency to follow another’s gaze only when ostensive eye contact or an infant-directed greeting has been established (Senju and Csibra, 2008), along with displaying an expectation of a referent object when following another’s gaze in an ostensive context (Csibra and Volein, 2008). Relying on such claims, Gergely & Csibra (2014) propose that human infants are fundamentally disposed “to being at the receptive side of verbal as well as pre-verbal communication from the beginning of their lives” (p. 3).

For present purposes, the significance of Csibra & Gergely’s theory of natural pedagogy is that it provides a developmental basis for the claim that learning and transmitting knowledge is a *fundamental intersubjective* aspect of human life.<sup>[38]</sup> For Csibra & Gergely (2009), the demonstrative pedagogy of ostensive communication is distinct from the communicatory functions of cooperation, manipulation and gossip, amongst others. However, I believe one can posit a pedagogical dimension to all forms of communication. As discussed already, to cognise and act in a properly *human* way, one implicitly (and, at times,

---

<sup>[38]</sup> An alternative to this developmental story could be provided through the notions of *primary intersubjectivity* and *secondary intersubjectivity* (see Gallagher (2005) and the next chapter).

explicitly) conforms to social norms. These modulating norms are reciprocally modulated by individuals' behaviour, with such co-modulation requiring a continuous assimilation and dispersion of norms across and within various social structures. This continuous assimilation and dispersion of norms suggests that we do not only learn through ostensive communication, but also as a natural upshot of living in accordance with the normativity of humanity.<sup>[39]</sup> Although the pedagogy of infants may need to be ostensive in certain circumstances, the actions of non-infant humans seem to be imbued with a non-ostensive pedagogical dimension through the initial foundation of humans' pervasive phenomenological acquiescence to one another's presence. In this way, pedagogy can be considered an importantly broad category, one that is not limited to certain forms of communication or a specific ontogenetic story, but is instead a fundamental feature of *being human*. That is, the pedagogical transfer of norms is an implicit feature of our living through the enactment of norms, or, from a broader perspective, it is an implicit feature of human nature. In the subsequent section I will expound this idea further but, for now, the key point is that the pervasive normativity of human behaviour necessarily incorporates the notion of a pervasive pedagogical dimension. Whilst such indirect pedagogy may differ from Csibra & Gergely's direct and ostensive 'natural pedagogy', the former can still affix itself to the latter's developmental story. Thus, the social ontology of being human is also constituted by the FIQ of *pedagogy*, alongside a *sense of otherness* and *status*.

As far as theories of social cognition are concerned, the primary significance of pedagogy is as the mechanism by which status traits manifest normative 'grip'. It can also be viewed as another FIQ from which a full and satisfactory account of social cognition should not be abstracted. We do not just encounter and engage with others as entities to be analytically discovered, nor as interactional partners with whom we can jointly uncover new ways to cognise; we also encounter and engage with others as repositories and recipients of generalisable knowledge, from and with whom our cognition unfolds, expands and transforms.

The importance of these FIQs is that we now have a more complete understanding of the nature of the human agents who are commonly the acting subjects and intentional targets of social cognitive studies, as well as how this nature lays a unique foundation from which agents come to understand one another during interactions and observations. A consequent of this is that it provides robust reasons for parsing *anthro*-social cognition – that is, one human cognising with or about another human during an immersive interaction or observation – from all other forms of cognition. Quite simply, non-human entities that are

---

<sup>[39]</sup> Empirical evidence of norm-learning outside of natural pedagogy comes from the cultural transmission of language through iterative learning (Kirby, Cornish and Smith, 2008) and through '(non/)conformist-bias' (Eriksson and Coultas, 2009), as well as through primary and secondary intersubjectivity (see previous footnote).

It thus seems clear that humans are predisposed to acquiring and dispersing normative knowledge from/to one another.

engaged with by humans do not import FIQs as a constitutive part of their ontology, whereas other humans do. In the next section, I will turn my attention to the dynamics of the normative connection that is generated during anthro-social cognitive events.

### **5. Social Cognition and Self-Other Immersion**

Across the previous three sections (4.2.1-4.2.3), I have argued that a satisfactory theory of social cognition should not be abstracted from the FIQs of *sense of otherness*, *status traits* and *pedagogy*. With these social qualities established, the question that must now be faced is when exactly one can be said to be *socially cognising*, rather than simply cognising? Motivating this question is the fact that merely encountering another seems to generate a phenomenological transformation which differs from that of directly interacting with another or (many forms of) observing another. Thus, if we are to claim (as most people do) that social cognition is distinct from non-social cognition, we need a new criterion to justify the distinction. My argument is that agents can become immersed in another's co-present activity such that a unique self-other domain of normativity in which one or both agents' worlds become phenomenologically transformed to the perspective of a 'we-self'.

As discussed in chapter 3, for most 'mindreading' theories, the processes of social cognition seem to be part of an internal mechanism that can simply be deployed when required. This results in social cognitive accomplishments often being treated as detached, predominantly observational abilities that largely ignore the fundamentally embodied and ensocialled aspects of human nature. For 'interactive' theories, there is a struggle to explain exactly when one is *in* an interaction. The theory of participatory sense-making comes close to diffusing this problem by appealing to the intuitive (and philosophically supported) notion of an interaction 'taking over' from strictly individual autonomy with its own momentum (De Jaegher & Di Paolo, 2007; Di Paolo & De Jaegher, 2013). However, this interactive 'taking over' can become very unclear when considered more closely (see chapter 4), and it also seems to preclude the possibility of socially cognising from outside of an interaction (i.e. from an observational stance), when any relational dynamics between individuals are less apparent. Both 'mindreading' and 'interactive' theoretical approaches also struggle to account for the intuitive distinction between human-human cognising and human-nonhuman cognising. Throughout the next two subsections (5.1-5.2), I will elucidate the manner in which the concept of *self-other immersion* overcomes these issues to provide a more satisfactory delineation of social cognitive engagement. In order to maintain an emphasis on the phenomenological illumination of cognitive science, I will consider the notion of self-other immersion from two complementary perspectives: firstly, I will outline the phenomenological notion of a 'we-self'; I will then consider the unique self-other domain of normativity that is generated by the lived experience of a we-self.

### 5.1. The Experiential Structure of the ‘We-Self’

In order to properly grasp the concept of ‘we-self’ and its experiential structure, it will be useful to re-consider the experiential transformation that takes place when *sensing otherness* (section 4.2.1). Recall that sensing otherness is an implicit acquiescence to another’s mere presence as a fellow ensocialled agent, which results, amongst other things, in a unique kind of phenomenological shift in one’s experience. The shift that takes place includes a *re-configuration* of the phenomenological scene, generating a *before another* structure to experiential options. Importantly, however, one preserves an appreciable degree of self-ownership over one’s action possibilities. That is to say, the other has a permeating influence on one’s phenomenological world, but during the incipient moments of an encounter it remains an influence that is ancillary to one’s usual experiential ownership. Thus, even as the world takes on a *before another* structure, it is still emphatically a self-owned ‘I’ that acts before another. In a generic form, the acquiescence that takes place when sensing otherness produces the experiential structure of ‘I am *Xing* before another’.

Preserving a self-owned ‘I’ whilst acting before another can be considered one mode of behaving as an ‘I-self’ – that is, retaining, for the most part, governance over one’s activity. I may be acting before another, but *I* am still the dominant initiator and executor of my own activity. In contrast, when one immersively interacts with or immersively observes another (so as to be properly *socially cognising*), one is no longer an I-self, but instead experiences the world as a ‘we-self’. As a we-self, one no longer has one’s world merely blanketed in self-other referentiality by the other’s passive presence; rather, one’s world is, in part, constituted by the other’s active presence. This constitution involves the other’s active co-presence framing an agent’s experience, such that the agent’s activity is funnelled through the particularities of the other’s behaviour. Much like the manner in which one’s mood can dispose an agent towards certain thoughts and behaviours, the active co-presence of another can existentially dispose an agent towards a specific mode of being: a mode of being that incorporates the other in a structurally fundamental manner. In terms of affordances, the phenomenological shift is no longer to a *before another* structure, but to a *with this (specific) other* or *through this (specific) other* structure. It becomes a matter of co-responding to affordances: ‘I (as a we-self) am *Xing with/through this (specific) other*’, instead of ‘I (as an I-self) am *Xing before another*’. In this way, one is *immersed* in another’s presence, rather than merely acquiescing to the other’s presence. Whilst the concept of *immersion* will be explicated further in section 5.2, it is worth noting that it inherently imports awareness of the FIQs of *status traits* and *pedagogy*, in addition to the *sensing of otherness* that occurs in a mere encounter.

If we take a two-person face-to-face conversation as an example of the we-self taking hold, the possibilities for action that arise within this interaction are not given to one individually as an I-self, but are partly ‘owned’ by the other, so that any subjective action is constituted by dyadic reciprocity. Neither interactor individually owns or generates possibilities for action, with each co-owning and co-generating a self-other domain of normativity that is unique to the two of them (again, there will be more on this normative generation in the subsequent section). As a we-self, one’s cognition is thus predicated on the other’s (inter)active involvement, such that one inhabits a specific world of possibilities with/through the other.

In order to add flesh to this conception of the experiential structure of the we-self, a delicate comparison can be made with aspects of an expert carpenter’s phenomenological experience of using a hammer, which was mentioned in chapters 1 and 2 (see Heidegger (1927/1962)). Recall that the carpenter engages with the hammer with such nuanced skill that the conspicuous presences of the hammer and of the carpenter as a hammer-user are amalgamated into a non-reflective performance (ibid., p. 98). The carpenter and hammer do not stand out from one another as distinct entities, but instead become united in the ongoing flow of experience. In Polanyi’s (1964) words, the carpenter existentially “accepts” the hammer by “dwelling” in it such that it is part of his subjective activity (p. 64). To ‘dwell’ in entities is to “assimilate them as parts of our own existence” (ibid.), so that one’s worldly possibilities are disclosed through the entities themselves. This dwelling is perhaps best exemplified by the passionate performances of professional singers or poets, who do not stoically recite words, but are moved by them in a kind of active interplay. That is, the singer co-generates her performance with a song’s lyrics (and music), letting the lyrics draw expressions from her as much as she is uttering them. She emotionally draws the lyrics into herself and is simultaneously drawn out by them so that she lives *through* the lyrics. Similarly, for the carpenter, the hammer draws out actions as much as it is itself moved by the carpenter. To dwell in an entity is thus to accept it as an active co-generator of experience. It is important to point out that engagement with inanimate or intangible entities can never be directly analogous to engagement with ensocialled agents; however, the idea of ‘existentially accepting’ and ‘dwelling’ in some entity that is outside of one’s bodily self is highly pertinent at this point of analysis. The feature that I wish to extract from the notion of a carpenter’s phenomenological engagement with a hammer is that when agents engage with one another, each agent can be said to ‘dwell’ in the other’s activity. This is not to say that the agents lose explicit recognition of one another or their subjective separation, but that each agent ‘pours’ herself into the other’s activity and assimilates the other as a structural part of her present experience (cf. Polanyi (1964, pp. 61-68) on the assimilation of tools and concepts). If we recall the idea that another’s presence can transform one’s agency, the notion of we-self immersion claims that the other’s co-present activity directly modulates

one's agential existence, such that one simply does not cognise or behave in the manner that one would in the absence of the other. It is, in the terms of Merleau-Ponty's (1945/2012, p. 215) 'intercorporeality', as if the other's intentions inhabit one's own body and thereby modulate one's behaviour. When interacting with or immersively observing another, one's experience is thus cultivated by the other's such that one temporarily dwells in the other's involvement, therein generating a we-self mode of being.

Notably, this we-self dwelling in another's presence is not always an amicable and harmonious connection. Social engagements can often be antagonistic, combative or cumbersome. Yet these can still be viewed as a kind of collaboration, in the sense that involved parties occupy a shared space of co-modulation. Two individuals having a heated argument, for instance, are still co-generating a unique normative domain in which their argumentative actions are absorbed. Each individual is partially owning the other's behaviour and, simultaneously, inhabits a specific world of possibilities by dwelling in the other's activity.

It is also important to note that the emergence of a we-self does not signal the loss of a given subject to group dynamics. The experiential structure within which one dwells as a we-self is inherently dependent on one's continued connection to the other (on whom the structure is also inherently dependent). This dependence means that each agent that is involved in the interaction, as a we-self, maintains a unique kind of subjective ownership over the interactive experience. However, this unique we-self-ownership is not the same as individual I-self-ownership as it is appurtenant to the reciprocal give-and-take of one's own and the other's collectively generated self-other domain of normativity (see the subsequent section). To be in the mode of a we-self thus accepts the importance of other-generated experiential modulation, whilst resisting the abandonment of any notion of 'self' within a collective domain. We-self toes the line between an individualistic 'I' and an I-less 'we'.

Crucially, this transformation to a we-self is not something that can be induced by non-human entities, nor can it arise through any fleeting or purely physical encounter. Due to the FIQs of human nature (section 5.2), social engagement amongst humans is bestowed with a unique normative component that carries the we-self as its phenomenological counterpart.

## 5.2. Self-Other Normativity

Throughout the previous chapter and section 4.2 of this chapter, the idea that human nature and cognition are constituted by a normative component has been a recurrent feature. In short, the proposal is that there is no such thing as (a). a non-social human being, or (b). a non-social human thought. Every thought is, in part, derivative of social normative structures that help to scaffold, transport and constitute cognitive content. Even a highly original thought, such as Einstein's envisaging a ride atop a beam of light on his journey to

developing the theory of general relativity, will be dependent on social norms of thinking and behaving in certain ways. And if thinking and acting are essentially socially normative accomplishments, then to be human in a meaningful way is to constitute and be constituted by social norms. These social norms are generated, maintained and modulated at various levels of analysis: globally, societally, familiarly, through cultural organisations (corporations, charities, sports teams etc.) and during everyday interactions. Diverse collectives of agents, through behaviour and cognitive performances, generate and modulate social normative structures which reciprocally modulate agential behaviour and cognition.

Let us imagine two friends having a drink together in a quiet bar. Each friend is reciprocally responsive to the utterances, body position, movements, countenance and mood of the other, such that their interaction is constituted by fluid and mutually maintained dynamics. Due to this reciprocal responsiveness, neither friend can be said to be individually dictating their own or the other's behaviour; they are each subject to the power of the co-generated interaction itself. That is, the dynamics of the interaction that are generated between them feedback to each friend and thereby constitutively determine their thoughts and behaviour. In phenomenological terms, each friend's experience is reciprocally structured by the other's in such a way that the normative canalisation of their activity is the product of we-selves – selves that are immersed in a co-constituted organisational structure that is unique to the two of them. Each friend, in virtue of their ongoing interaction, *lives through* the relational dynamics of their collectively transformed phenomenal worlds. In other words, the friends are generating and modulating a *self-other domain of normativity* that is reciprocally modulating how they are thinking and acting together. Suppose that one friend begins reminiscing about a past humorous event. The other friend, who remembers the event, enthusiastically joins in with the retelling. Sensing this enthusiasm, the first friend recounts his version of the tale with growing gusto, which the other corroborates with similarly increasing joy. The enthusiasm from both parties soon seems to give the story an impetus of its own, such that it carries along its narrators in a crescendo fashion towards its humorous climax. In other words, the two friends' retelling of the event generates a unique normative domain that reciprocally modulates their cognition and behaviour, which in this case amounts to incrementally enhancing the passion and enthusiasm with which the event is recalled.

This self-other domain of normativity and its constitution by we-selves may invoke comparisons to the theories of we-mode cognition (Gallotti & Frith, 2013a, b) and participatory sense-making (De Jaegher & Di Paolo, 2007; Di Paolo et al., 2010; Di Paolo et al., 2013), which were criticised in chapters 3 and 4. Both of these theories extol the virtues of psychological togetherness and the coordination of communicatory activity, in manners that certainly bear some resemblance to the notion of self-other immersion that has been

developed here. Indeed, self-other immersion in a normative domain may at first seem like a reformulation of the interactive autonomy of participatory sense-making, in which the reciprocal relational dynamics between individuals are said to ‘gather their own momentum’ and ‘take over’ individual cognising. However, by placing an emphasis on relational *normativity*, the present delineation of immersion in a unique self-other domain differs from both we-mode cognition and participatory sense-making in important respects. With regards to we-mode cognition, the initial difference is that the current delineation of social cognition is designed to capture all instances in which an agent is phenomenologically transformed through immersive involvement with a co-present other, whereas we-mode cognition is a unique form of social cognition that is most applicable to joint action. We-mode cognition also fails to explain what is unique about human-human interactions, which the present theory overcomes through emphasis on the FIQs of human nature. Finally, the present theoretical framework clarifies how a subject can become integrally involved in a transformative way with another (as a we-self), whereas we-mode cognition became incoherent in its attempt to make a similar claim (see chapter 3, particularly section 4).

With regards to participatory sense-making, the theory of social cognition that has been presented throughout this chapter differs in four notable ways. Firstly, we know that immersion in a self-other domain of normativity is unique to humans due to the FIQs that were discussed in section 4.2, so that relevant self-other relational dynamics only emerge during human-human social engagement. Secondly, mere encounters cannot generate relevant relational dynamics because the phenomenological transformation to a we-self within a self-other domain of normativity has been shown to differ from the acquiescent transformation that takes place (as an I-self) when sensing otherness. At this point, one would be entitled to worry that there is a kind of slippery slope between the acquiescence of mere encounters and the immersion of full-blown social cognition engagements. For instance, ‘video chatting’ through platforms such as Skype would seem to be more than simply acquiescing to another’s presence, but still lacks the full-blown immersion of the embodied engagement of physically present interactions. Such a slope would be problematic in that one of the criticisms levelled at participatory sense-making was its exposure to a ‘bloat’ worry (i.e. its failure to place a firm ‘frame’ around those situations that should count as instances of social cognition; see section 4.1.1 in chapter 4). However, the notion of immersion implicates FIQs and specific focus on normativity, both of which are lacking in the formulation of participatory sense-making. The overarching view that I am currently espousing thus provides a strengthened ‘frame’ for social cognition, although certain ‘borderline’ cases are inevitable. To evince this strengthened framework, the third notable divergence from participatory sense-making is that purely physical coordination, such as two individuals stepping around one another in a narrow corridor, will not generate a self-other

domain of normativity *unless* the individuals are attentive to one another and are thereby normatively affected by one another's status and pedagogical influence (in addition to sensing otherness). The fourth and final divergence from participatory sense-making, which will be addressed in the subsequent section, is that the notion of self-other immersion can account for the unity of interaction and immersive observation – that is, thinking *with* and *about* present others – within a theory of social cognition.

## **6. Observation as Interaction**

One of the drawbacks of existing theories of social cognition is that they rarely account for instances of both social interaction and social observation. 'Mindreading' theories tend to focus on observational inferences to the detriment of the collective dynamics that emerge in interaction, whilst 'interactive' theories tend to shun the social processes that obtain when cognising about another from a spectatorial point of view. There are obviously clear differences between social observation and interaction, both from commonsense and empirical perspectives (e.g. Tylén et al., 2012), but a theory of social cognition that operates within the framework of FIQs and self-other immersion can, at the very least, establish a solid area of unity between the two perspectives.

Firstly, recall that a self-other domain of normativity emerges during a dyadic interaction when each individual is attentively responsive to the other in such a way that one individual's actions will induce a modulated response in the other, which will then feedback and modulate the first individual's response, which will reciprocally modulate the other's response, which will reciprocally modulate the first individual, and so on. This mutual reciprocity generates a normative domain that is unique to the involved individuals, in accordance with their ongoing interaction. The continuous feedback loops between the two ensure that neither one of them can be said to be individualistically acting or cognising; they are each subject to the collectively generated normativity of the interaction itself.

Now let us consider one individual observing another physically present individual. There is no longer an obvious manifestation of feedback loops or mutual reciprocity between the observer and the observed, as only one of them is attending to the 'relationship' between them. For most current theories of social cognition, this issue is unsurpassable. However, as long as we operate within a cognitive framework which endorses ensocialled agents, who only cognise in accordance with social norms, there is in fact a route to explaining social observation as a kind of interaction.

In the last chapter, particularly sections 2 and 3, I explained how all human cognition can be conceived of as constituted by social normative structures. So, for example, when someone uses language, they are engaging with an existent, cultural body of language whose comprehensibility is presupposed (Steiner & Stewart, 2009; Luhmann, 2002/2012).

Similarly, the appropriate distance to stand from another during conversation requires the enaction of culturally (or more locally) variable standards (Dreyfus, 1991). The same goes for any everyday behaviour: social normative standards are engaged with in the performance of socially recognisable activity. Even iconoclastic behaviour is telling of social norms in virtue of its explicit rejection of them. Importantly, socially normative behaviour implicates socially normative cognition, so that even one's 'internal' ponderances are canalised by social normative thoughts or *ways of thinking*. Thus, when one individual observes another, we have still not escaped the bonds of normativity. The observed individual's actions pedagogically convey a certain manner of acting to the observer, either in accordance with certain social norms or as a manipulation of them. Crucially, this manner of acting is not discarded by the observer; even if it is not explicitly recycled, it is still assimilated into the observer's cognitive architecture of normative possibilities. If the observer sees this manner of acting again or opts to re-enact it herself, or even if it is never seen again and never re-enacted, the behaviour has been introduced as something that is poised to gather momentum as a social norm or something that will die out. Suppose the observer does re-enact the previously observed behaviour (although the same argument would hold if she does not): we would then have one individual behaviourally engaging with some norm, which modulates the behaviour of an observer, who would then similarly modulate the behaviour of others through her (observed) behaviour, with these others then modulating further others, and so on. All agents involved in this process of continuously modulating some norm (which socially constitutes how to behave in a certain situation) would be reciprocally subject to the modulation of the norm itself. We thus have a repeat of the relational dynamics that occur in dyadic interactions: (collectives of) individuals engage with and modulate a social norm (or norms), which reciprocally modulates individual behaviour. The differences between the case of social observation when compared to social interaction are temporal and demographic; that is, the normativity of observation will only feed back to an individual over time and the normative domain that is manipulated is generally pre-existing, rather than unique to the involved persons. However, as far as the relational dynamics are concerned, observation can still be considered a form of interaction.

Two further clarificatory points need to be briefly addressed. Firstly, for an appropriate normative connection to obtain between an observer and an observed other, both agents must be bodily present. This is to ensure that relevant FIQs are in-play so that the observer is suitably disposed to the other's normative influence (see section 4). Secondly, from a phenomenological perspective, when one is immersed in observing another, one is so as a we-self. That is, the normative modulation of the other is such that one's phenomenology is transformed to a '*through this other*' structure; the other is integrally involved as a structural part of the observer's present experience. As one individual immersively observes another,

the normative modulation that occurs prevents the observer from cognising as an I-self. She is specifically attending to *this other's* actions, which induce in her modulated cognitive processes (perhaps through the assimilation of motor resonance activity (see Gallese, Fadiga, Fogassi and Rizzolatti (1996) and Rizzolatti, Fadiga, Fogassi and Gallese (1999) for early assessments of motor resonance, and Gallese & Cuccio (2015) for an updated interpretation)). With this we-self-projection during some social observations, one once again protects against the kind of 'bloat' worry that troubles participatory sense-making, although, as stated previously, there will be some indistinct cases between we-self observation and I-self observation.

It is important to reiterate once more that the claim here is not that social observation is in every way the same as social interaction. Rather, the claim is that the normative framework for social cognition that I have put forth, drawing on the notion of *self-other immersion*, can elucidate a highly significant area of overlap between the two phenomena. Both interaction and observation function through individual immersion in self-other domains of normativity. This is relevant because both social interaction and social observation seem appropriate for envelopment within a theory of social cognition as they are both routes to understanding others.

## **7. Conclusion**

Throughout this chapter, I have put forth a new framework for theories of social cognition that focuses on *the transformation of cognising that unfolds when human agents are immersed in the activity of a co-present other (or others)*. It is through such immersive transformations that humans are disposed to understand others. I will now swiftly highlight the benefits of this framework:

- I. by focusing on FIQs and normative connectivity, the framework gives due attention to the ontology of humans, thereby justifiably confining social cognition to human-human interactions and observations.
- II. the approach bridges a gap between individualism and collective emergence. That is, it appreciates that social cognition is an *intersubjective* achievement, but does not lose the individual within this intersubjectivity. The phenomenological notion of a we-self captures how an agent *lives through* a modulating other, thus capturing the integral importance of the other's involvement, whilst emphasising that it is still subjective experience that is modulated.
- III. the need for *immersion* in a self-other domain of normativity protects theories of social cognition from being applied to fleeting encounters or instances of purely physical coordination.

- IV. contrary to the majority of existing theories, my proposed framework allows for the unifying principle of normativity to explain how both social interaction and social observation can be incorporated within a single approach to social cognition.

Underlying these benefits is the idea that phenomenology can not only form a mutually illuminating relationship with cognitive science, but can in fact provide foundations from which scientific explanations unfold. By making use of this phenomenologically grounded approach, I believe that a more resolute and suitably anthropocentric framework for social cognition has been delineated.



# **Chapter 7 – The Biosocial** **Self**



## The Biosocial Self

### **1. Introduction**

To this point in the thesis, I have made claims regarding the nature of our everyday engagement with the world, the phenomenology of encountering others, our pervasive constitution by social normative structures, issues with existing theories of social cognition and a new delineation of social cognition in accordance with the ontology of being human. Throughout all of this, I have made several references to ‘subjects’ or ‘agents’ as importantly embodied, autonomous, embedded, enactive and, in the last two chapters, ensocialled. For the most part, I have focused on the positive and overlapping aspects of each of these adjectival descriptions, while not discussing their many differences and conflicts. I have, in many ways, thus taken the idea of ‘subject’ (and ‘subjectivity’) largely for granted. In this chapter, I am going to bring together the several aspects of subjectivity that I have discussed in order to construct a picture of a *biosocial self*.

A preliminary caveat must be addressed, which is that I am taking the necessary connection between subjectivity and selfhood largely for granted. The exact relation between these two concepts is a highly complex and hotly debated one (see Zahavi (2005), Janzen (2007), Dreyfus (2011) and Ganeri (2012) for just some recent discussions of the topic), but a full engagement with the associated debate would warrant its own thesis and it is not something that imminently detracts from the novel notion of self that I wish to put forward in this chapter. Inasmuch as a clarificatory remark is required, I concur with Zahavi’s (2005, 2014) contention that any conscious episode, which one can reasonably assume is a subjective experience (broadly construed), simply cannot be selfless. A conscious experience must involve a conscious experiencer, and an experiencer is always a self, in at least a very basic sense of the word. One cannot experience something without a degree of first-personal involvement – an experience must always be *someone’s* – and it seems to me that conscious first-personal experience must entail selfhood, lest we abandon the notion of self as a reality (see Albahari (2006), Krueger (2011), and Hood (2012) for some diverse approaches to the idea that selfhood is an illusion).

This caveat aside, the structure of this chapter will be as follows: in the next section I will provide a detailed outline of Kyselo’s (2014) enactive approach to the self, with which my account of ‘biosocial selfhood’ is strongly connected; in section 3, I will introduce two criticisms that I believe Kyselo’s account must face; section 4 will respond to these issues from within the framework of Kyselo’s theory; section 5 will then elucidate biosocial selfhood as a more satisfactory alternative to Kyselo’s theory.

### **2. Kyselo’s Enactive Approach to a ‘Body-Social’ Self**

### 2.1. The Need for Clarity

In her paper, *The Body Social: an enactive approach to the self*, Kyselo (2014) is addressing one of philosophy's oldest questions: what is the human self? Her proposal is that converging on this question from the perspective of enactivism will overcome a persistent tension between embodied and social approaches to selfhood. As Kyselo explains:

On the one hand, embodied cognitive science risks a new form of methodological individualism, implying a dichotomy not between the outside world of objects and the brain-bound individual but rather between body-bound individuals and the outside social world. On the other hand, approaches that emphasize the constitutive relevance of social interaction processes for cognitive identity run the risk of losing the individual in the interaction dynamics and of downplaying the role of embodiment. (p.1)

The tension is thus one of preserving a genuinely first-personal selfhood whilst not individualistically confining this selfhood to the organismic boundary of the body. This tension is perhaps unexpected because the underlying motivation behind both the embodied and social 'turns' in cognitive science is to move away from the notion of the brain as a kind of executive controller of cognitive processing. Both embodied and social approaches wish to emphasise the dynamic and not uniquely neural interplay of experiential processes that are implicated in systemic cognition, so that cognition is, in part, an importantly worldly achievement. Yet if one endorses an embodied view of cognitive science – that is, the view that cognition is not uniquely 'in the head', but is also deeply dependent on bodily processes (which are themselves embedded in a scaffolding environment) – then there is the tendency to individuate the self in accordance with the living body. In other words, the self is designated as a bodily self, independent from (albeit heavily interlinked with) sociocultural norms and processes. Conversely, if one endorses a social view of cognitive science (also referred to as "the interactive turn" (De Jaegher, Di Paolo and Gallagher, 2010)) – that is, the view that cognition is primarily intersubjectively dependent – then there is a tendency to individuate the self in accordance with social processes, such that one can become 'lost' in the relational dynamics that are generated through interactions with others (there will be more on this shortly). However, once one moves 'beyond-the-brain' and embraces the integral dependence of cognition on worldly features, either of these tendencies seems to prejudicially carve the cognitive pie, relegating bodily or social processes to merely contextual significance. Indeed, even if bodily or social processes play a crucial causal role in the individuation of the self (beyond a merely contextual position), this will still not fully satisfy Kyselo's objective of integrating bodily and social dimensions into a coherent unity of selfhood (Kyselo, 2014). What is needed (and will be delivered in section 4) is a constitutive fusion of body and social such that neither can be extracted and leave the self intact.

As Kyselo explains, the question is how to coherently unite bodily and social processes with regards to selfhood so as not to conceptually isolate either set of processes, promoting one set to the status of harbouring the true self and relegating the other set to merely causal or contextual contributors to the self. In other words, we are faced with the matter of resolving the *body-social problem*: “the question for philosophy of cognitive science about how bodily and social aspects figure in the individuation of the human individual self” (Kyselo, 2014, p. 4; see also Kyselo and Di Paolo (2013)). If our embodiment is considered the ‘special’ aspect of selfhood then one risks a hollow stance towards social processes, whereas the reverse can also be true. Without conceptual clarity, these bodily and social processes may seem mutually exclusive and attempts to unite them cannot do so in an equilibrating manner (Kyselo, 2014). The self proper, as far as Kyselo is concerned, is a kind of unity of bodily and social processes such that one cannot be disentangled from the other (although, as we shall see in section 3, it is debatable whether she actually achieves a true unity of these processes).

One solution to this tension would be Gallagher’s (2013b) pluralistic approach to selfhood. Gallagher’s idea is that we have several aspects of selfhood that form interrelated layers of a cohesive self. These various self-aspects include *minimal embodied*, *minimal experiential*, *affective*, *intersubjective*, *psychological/cognitive*, *narrative*, *extended* and *situated*, and they coalesce into a coherent pattern of ‘self’ (ibid., p. 2) The appeal of this approach is that it sensibly acknowledges the great variety of self-related phenomena that seem to constitute each one of us, from having a basic bodily schema with its own sensorimotor contingencies to inhabiting a specific sociocultural world, or from incorporating worldly entities in one’s physical activities to conceptualising oneself as the protagonist of a life story (and much else besides). These diverse phenomena all seem to play an important role in how oneself and others conceive of individual persons with unique characteristics. There is thus something intuitively appropriate about viewing these phenomena as somehow uniting into a self-constituting pattern in which no aspect of self is given executive precedence over others. However, the ‘somehow’ in the preceding sentence is exactly the problem with Gallagher’s pattern approach to self. Once we have a diversity of self-relating phenomena, it is crucial that the manner of their relations to one another is elucidated (Kyselo, 2014). Many people would agree that selfhood is a multi-faceted phenomenon that coheres into a discernible pattern, but it is exactly how this coherence is achieved that really matters. As Kyselo rightly claims, Gallagher’s approach is simply too “laissez-faire” (Kyselo, 2014, p. 1). Not only does it leave us with an incomplete theoretical picture of selfhood, it is also unsatisfactory from a practical perspective. An explanation of how self-aspects coalesce into a patterned self is essential for how a doctor treats a patient, how a psychologist conceives of pathologies of the self, or how any number of academics conceive of experimental subjects (ibid.). Without the

coherence of self as a unity, there could be attempts to isolate and independently deal with the various self-aspects that comprise an apparent pattern.

A polar opposite approach to Gallagher's (2013b) 'patterned self' is to focus on one aspect of seemingly multi-faceted selfhood and designate this aspect as *the* self. For Kyselo, this kind of approach, which she designates 'essentialist', suffers from two insurmountable problems. Firstly, there is minimal consensus on what aspect of personal existence could qualify as *the* self (or, at least, the most fundamental aspect of selfhood), with a great variety of well-argued claims being made for the brain (Churchland, 2013), specific neural regions (Northoff and Bermpohl, 2004), the organic body (Johnson, 1987), the affective body (Colombetti, 2014), narrativity (Dennett, 1992), explorations of the environment (Neisser, 1993), sociality (De Haan, 2010), or a mystical soul. Secondly, within this discordant arena, each claim is inherently reductive and so restricts selfhood to some singular element of existence, which seems to dismiss the dynamic and protean ways in which lives are lived. Consider, for example, that one would generally consider oneself to be the same person, in at least some ways, even if one underwent a significant bodily transformation, or if one settled in a new culture, or if one underwent a 'life-changing' experience such as finding religion, swimming with dolphins or surviving a near-death event. Similarly, one tends to consider oneself as the same person, in at least some ways, as one was at birth, or as one perhaps will be at old age, even if suffering from memory loss. The 'self', whatever it may be, seems to withstand significant changes throughout a life-time and it is very difficult to pin-point how selfhood persists in the face of these changes. Rather than being reduced to a specific feature, a theory of self should thus strive to account for the multifarious, interactive and persistent nature of self-existence, in such a way that a cohesive unity is discerned and preserved.

## 2.2. Kyselo's Solution

As far as Kyselo is concerned, enactivism provides the perfect conceptual framework within which a theory of self can be constructed – a theory that both appreciates the multifaceted nature of self and delineates it as a cohesive unity. In this way, her theory of self bridges between essentialist and pluralistic theories of self, as well as dissipating the tension between embodied- and social-focussed approaches to selfhood. According to the kind of enactivism that Kyselo favours, a cognitive entity is one that is *autonomous*, *emergent*, *embodied* and *experiential*, so as to become a *sense-maker* (see Di Paolo & De Jaegher (2007) or chapter 3 for a detailed explanation of these qualities). If we accept that a self is a cognitive being then this means that we already have several aspects of selfhood that are assimilated within the very framework from which an investigation into selfhood is being made. What's more, enactivism inherently provides a "fruitful link" into analyses of selfhood because its conception of autonomy is one of individuation. An entity is autonomous when it self-

generates, self-preserved and self-identifies – and thus self-individuates – through the recursively propelling activity of its constituents, such that a systemic unity with established options for environmental interaction is manifest (De Jaegher & Di Paolo, 2007; Thompson, 2007; chapter 3). That is, the recursive interactions of constituent processes form an interconnected network in which every process both enables and is enabled by some other process, with this self-maintaining network forming a systemic identity whose survival is reciprocally dependent on the interconnection of its constituents (De Jaegher & Di Paolo, 2007). Central to this “perpetual self-renewal” (Jonas, 1966/2001, p. 80) is the idea that an organism’s constituting parts necessarily require material resources to survive, but the organism’s identity is a continuous “achievement” of emergence *from* these material processes (ibid.; Kyselo, 2014). To put this more simply, whereas the material resources of the environment provide substance from which an organism is constructed, these resources are organised and processed by the organism in the unique way that is necessary for its specific survival, so that it maintains its own “functional identity” that simply ‘passingly incorporates’ the outside world (Jonas, 1966/2001). Jonas refers to an organism’s dependence on material resources whilst simultaneously functionally distinguishing itself from its material environment as “needful freedom” (ibid., p. 80). Due to needful freedom, an organism is incessantly interacting with its environment in order to survive and continually construct its own identity. To do this, an environment must be seen as valenced, such that certain resources are more ‘appealing’ to the self-preservation of identity than others. Every organism thus has its own ‘perspective’ on the world in virtue of the valenced appeal of the environment and, in this way, organisms are inherently purposeful beings, striving to maintain an identity through environmental interactions (Jonas, 1996/2001; Varela, 1997; Kyselo, 2014). This perspectival engagement with an environment applies to all living systems; however, for a system to be living *and* cognitive, adaptivity is also required (Di Paolo, 2005). Adaptivity is a system’s capacity to actively engage with certain environmental perturbations such that the system’s viability is preserved (ibid.). Without this capacity, a system is passively subjected to environmental perturbations. So, for example, a mere living system such as a biological cell produces its own metabolic network and physical boundary (i.e. its membrane) through the recursive interdependence of its chemical processes, which are reciprocally constrained by the cell membrane, and therein specifies the environmental perturbations that will bring about its structural changes; in other words, the biological cell is a paradigmatic example of a living system for enactivism. Yet a biological cell does not exhibit adaptivity, in that it only passively ‘frees’ itself from its material surroundings. Cognitive systems, on the other hand, such as cats, dogs, crocodiles or sucrose-seeking bacteria (Varela, 1997), will adaptively engage with environmental perturbations in varying ways in order to survive. This means that a cognitive system goes

beyond merely physically living in virtue of generating its own 'sense' of the world, thereby cognising as well as physically persisting (Kyselo, 2014).

Enactivism's bedrock of biological autonomy is, for Kyselo, the perfect foundation for selfhood, in that it inherently explains how a cognitive entity self-individuates itself through interactions with its environment. However, as Kyselo points out, there is an ambiguity lurking in the enactive framework that must be ironed out if a satisfactory concept of self is to be developed. Looking at the theory of *participatory sense-making*, which is the enactive approach to social cognition (De Jaegher & Di Paolo, 2007), Kyselo contends that it is unclear exactly how the autonomy of a human, who regularly interacts with other humans, persists during interactions. Recall from chapter 4 that the central claim of participatory sense-making is that when two or more individuals interact with one another their intentional activity can become dynamically coordinated in such a way that a new relational system *with its own autonomy* can emerge between them (De Jaegher & Di Paolo, 2007; Luhmann, 2002/2012). As this participatory process that emerges between interacting individuals has its own autonomy, it can be said to have a systemic identity of sorts, due to the fact that autonomy must 'belong' to a discernible structure of some kind. Importantly, whilst the interacting individuals are generating this relational process with its own autonomy, they are also reciprocally having their individual sense-making capacities enhanced by the interaction itself. There is thus a kind of bidirectional modulation between the individuals as constituting elements of the interaction process and the individuals as constituents of a 'group' autonomous structure. Yet this bidirectional modulation raises the pertinent issue: where or what exactly are these interacting individuals? Should they be individuated in accordance with the biological autonomy of their organismic bodies, or should they be individuated in accordance with the autonomy of the social interaction that they are generating? We thus face a crystallised example of Kyselo's *body-social problem*.

Kyselo puts forward this body-social tension within the theory of participatory sense-making in the following way. On one interpretation, (a). the individuals are 'dissolved' as constituents of the interaction because their "intrinsic purpose" seems to be directed at the generation and maintenance of interaction dynamics, which renders them *heteronomous* rather than autonomous. The resulting individuation, if there is to be one for the interacting agents, is of 'parts' within a social process that has its own autonomous ('group') identity (Kyselo, 2014; Steiner & Stewart, 2009). Alternatively, (b). each of the interacting agents "is individuated from others *qua* being *embodied*" (Kyselo, 2014, p. 7), such that any delineation of 'individuals' is made in accordance with organismic bodies. Yet differentiating each individual from the other and from their jointly created interactive dynamics in virtue of their presence as bodily beings condemns the body as "a locus of isolation, not a means of

connection and engagement” (ibid.). How, then, can we fairly make sense of individuation within the midst of these embodied and social autonomies?

In response, Kyselo quickly dismisses two responses. It won’t do to accord the body with individuating ‘power’, but concede that it is socially mediated, because this reverts back to the imbalance that Kyselo wishes to avoid, namely that social processes are ‘trivialised’ and come to matter “only as a context” (ibid., p. 7). Indeed, remarks from De Jaegher & Di Paolo (2007) that suggest the co-regulation of activity between interacting individuals is down to “bodily variables[...] and neuro-physiological variables” (p. 492) seem to exacerbate the idea that it is the body at the root of any individuated engagement with the world and, therefore, of any generation of socially autonomous structures. Social mediation of these bodily variables does not quite go far enough for Kyselo, who claims that “there is nothing social about the organismic or the moving body *per se*” (p. 4), and any attempt to socially embed bodies cannot avoid reducing individuals to “ready-made [bodily] identities” (p. 7) who just so happen to socially interact in important ways. Conversely, individuating agents in accordance with the autonomy of self-other generated interactions, such that individual identity “is defined as being relational or a participant in social interactions” (ibid., p. 6), risks distributing any given agent across and within her interactions to the extent that any notion of a genuinely *subjective* individual is completely obscured. Within the framework of participatory sense-making, Kyselo claims that “the individual’s nature as a relational being is underdetermined with respect to its own identity” (p. 7), resulting in a conception of individuals that emphasises the role of heteronomous constituents but not of autonomous ‘constitutors’.

Kyselo’s view is that the remedy to these difficulties emerges from Jonas’s notion of *needful freedom*. Building on this notion, the self should be conceptualised as arising “*through and from a world*” (Kyselo, 2014, p. 8), whereby

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 (ibid.)

The core processes of this principle of ‘human needful freedom’ are those of *participation* and *emancipation*.<sup>[40]</sup> To *participate* in the social world is to be organisationally self-individuated in accordance with social interactions (“virtual or actual interpersonal engagements” (ibid.)), such that one’s identity as a self is inherently socially relational. When individual agents participate in the social world, as an intrinsic part of the principle of human needful freedom, they are constantly achieving an identity, “jointly relying on

---

<sup>[40]</sup> Kyselo frequently uses ‘distinction’ in place of ‘emancipation’, but I will retain the latter term throughout this chapter for purposes of simplicity.

behavior and action and on doing and being together with others” (ibid.). *Emancipation*, on the other hand, keeps an agent from ‘dissolving’ into this social world of participatory processes by transforming one’s relationship with the social world. Rather than engaging with social processes as a constituent of some ‘group’ identity, an *emancipated* individual ‘stands out’ against the social relations from which it is made. Analogously to Jonas’s biological needful freedom, Kyselo’s principle of individuation *through and from a world* construes social processes as the ‘material’ through which an individual identity is constructed, but also as that from which the individual distinctively identifies itself in virtue of its unique processing of material. Thus, “just as an organism’s metabolism continuously exerts a choice by taking in only particular processes, while avoiding others” (ibid.), so does a socially organised human agent continuously exert a choice by taking in only particular social processes at any given time.

In order to clarify this notion of human needful freedom, Kyselo elaborates on her analogy. Whereas the autonomy of most organismic systems is dependent on the adaptive evaluation of the environment according to its potential to maintain the organism’s metabolic network, the autonomy of humans also requires the evaluation of the environment according to its potential to maintain a human’s network of social organisation. Rather than just requiring physical nutrition, humans are in need of *social recognition* as “the nutrient required to co-construct the boundary of the self” (ibid., p. 11). Ikäheimo (2009) goes as far as to claim that social recognition is the means by which human persons can be distinguished from mere animals, in virtue of surpassing the desirability and avoidability of a ‘pure state of nature’ and instead behaving in accordance with “mutually mediated intentionalities” (p. 34). In other words, humans live beyond mere animal existence through recognition of one another as participants in the collective mediation of norms – norms which are “constitutive of the lifeworld of persons” (ibid., p. 36). The role of social recognition here thus bears apposite similarity to the *ontological connectedness* (maintained by *sensing otherness*) that was discussed in the previous chapter and it will bear relevance in section 5 when the idea of *biosocial selfhood* is put forward. Kyselo also points out that the ‘nutritional’ value of social recognition is perhaps evident in the severity of consequences that arise during social confinement (e.g. Guenther, 2013; Rasmussen, 1973) or in the difficulties that emerge with self-identification due to social exclusion (e.g. Williams, 2007; Stillman, Baumeister, Lambert, Crescioni, DeWall and Fincham, 2009). The idea is thus that just as an organism is generated and maintained through the ongoing metabolic processing of physical matter (of which it is itself a part), so is a human self (co-)generated and (co-)maintained through the ongoing organisational processing of social interactions (of which it is itself a part) (Kyselo, 2014). To be starved of social interactions in some way would be to lose the ongoing capacity to generate a self.

There are two crucial steps for Kyselo's principle of individuation through and from a social world that we need to briefly consider. Firstly, an individual's participation in and emancipation from social processes are alleged to not occur at ontologically distinct 'levels' of identification (ibid.). That is, in everyday interactions an emancipated individual is not acting as a solitary, socially void agent and a participating individual is not mindlessly 'lost' within social customs and practices. Rather, emancipation and participation should be seen as poles within an individuating spectrum that describes the extent to which certain social 'matter' is incorporated into one's present delineation as a self. Emancipation roughly equates to being oneself in one's own right and thus 'owning' one's actions (e.g. practising yoga, feeling nervous before an audience, romantic disconnection, or the pride or shame one feels at a successful achievement or embarrassing act); participation roughly equates to openly connecting with others as a constituent of some collective activity (e.g. falling in love, feeling partly responsible for a team or institutional achievement or failure, or being caught up in the momentum of a crowd). For an agent to be a self is to be present on this emancipation-participation spectrum. Secondly and interrelatedly, the fact that the principle of human needful freedom is relevant to emancipation from and participation in social processes does not, for Kyselo, mean that the role of the body is downplayed. Instead, the body "changes its status" during our social interactions (and emancipation from them) so as to become "the self's means and mediator" (ibid., p. 12). This involves organisationally regulating the various forms of social relations that an agent engages with through the social perspective and impetus that is manifest by the body's physicality and expressivity. The body acts both as "an interface for connection" (ibid., p. 12), allowing us to engage with the world and be engaged with by others in specific ways, and as "an imprint of social engagement" (ibid.), shaping our cognitive responses to social situations in accordance with previous interactive experience. So, for example, adopting an expansive and open bodily posture can regulate an interaction by reflecting (and producing) a feeling of social power and dominance in the poser, thereby illustrating the body's role as a social "interface" (Carney, Hall and Smith LeBeau, 2005; Hall, Coats and Smith LeBeau, 2005). Reversing this relationship, the fact that one may blush during the recurrence of a formerly embarrassing situation indicates how social occurrences can imprint themselves onto the body (whilst simultaneously regulating the subsequent interaction). In other words, the body's expressive characteristics help to facilitate and regulate social interactions, whilst also acquiring new meaning through these interactions. The body is therefore inherently tied to *how* agents socially interact, from both a subjective perspective of others and from others' perspectives of a subject. Whilst openness and connectivity with others during social interactions are the key conditions for achieving selfhood within Kyselo's enactive framework, it is essential to her theory that this openness and connectivity is only possible in virtue of the body's mediation.

In summary, Kyselo considers human autonomy, like that of all organisms, to be underlaid by the principle of *needful freedom*. Unlike other organisms, however, the needful freedom of humans requires recognition during interactions in order to preserve an identity as a self. This means that human autonomy is not only dependent on the metabolic preservation of the organismic body, but also on the interactive (self-other) preservation of a socially organised self-identity. By oscillating between the processes of emancipation and participation during social interactions, the self's social autonomy is continuously co-generated and co-maintained. Crucially, this social autonomy can only be preserved with the mediation of the body, which is not a "random vessel" but the "means and mediator" (Kyselo, 2014, p. 12) of social engagement.

### **3. Problems with Kyselo's Account of Selfhood**

#### **3.1. The Persistence of the Self – Overcoming a 'Social Death'**

Part of the motivation behind Kyselo's enactive approach to the self is a desire to unite the potentially disparate processes of bodily and social being such that neither one can be said to have primacy over the other with regards to individuation of the self. Such unity should rid us of any ambiguity regarding the status of an agent during interactions. However, it is unclear that Kyselo achieves the balanced unity of bodily and social processes that she craves. This is perhaps most apparent when considering the possibility of 'social death'.

As we have just seen in the previous section, Kyselo's enactive self relies not just on biological needful freedom, but also on *social needful freedom*, which describes the necessity of self-constituting social interactions and distinguishing oneself from these interactions. Preserving oneself as a social unity is achieved by negotiating the balance between participating in interactions with others and emancipating oneself from these interactions. However, this social unity is forever "at risk of dissolution" due to the potential to lose the balance of participation and emancipation, thereby becoming "stuck in the extremes of either of the two dimensions" (Kyselo, 2014, p. 10). If this occurs and one's unity of social selfhood is 'dissolved', then Kyselo claims that one suffers a *social death* (ibid.). Exactly how one would become 'stuck' at either the pole of participation or that of emancipation is not entirely clear, but Kyselo provides us with some indicators. Firstly, an individual could 'dissolve' into the extremes of participation if one were to embed oneself in intersubjective activity to the extent that one suffers a complete loss of individual agency (ibid.). We could perhaps entertain the possibility of such an event in the case of cult members, in which persons reside in cult customs and the instructions of a leader to the point where the subjective ownership of actions seems entirely lost. Cult members would perhaps be best described as largely heteronomous with regards to their actions, undergoing governance that is established outwith their personal activity. At the other end of the spectrum, one's self is

'lost' to the dimension of emancipation when one becomes isolated from self-constituting social structures (ibid.). A possible example of such an occurrence would be extreme social confinement or isolation, or perhaps the unreality of the social world that one can feel when depressed, so that one feels completely estranged from meaningful social ties (Ratcliffe, 2014, pp. 15-17). For Kyselo, such happenings would amount to a loss of the ability to oscillate between the poles of participation and emancipation, thus becoming 'stuck' in one of these polar dimensions.

However, there are difficulties with buying into the possibilities of 'extreme participation' or 'extreme emancipation'. With regards to the alleged loss of oneself within the domain of participation (e.g. within a cult), it is hard to conceive of a subject being truly excised of their conception of self-owned subjectivity. As long as bodily autonomy and phenomenal consciousness remain, it would seem that there are always aspects of experience that are inherently *for the subject*. As far as Kyselo is concerned, these aspects of experience would perhaps amount to some form of 'identity', but not genuine selfhood. However, this is itself problematic. Every subject seems to participate in certain social normative structures to an 'extreme' extent (e.g. the normativity of being a male or female, a parent, a partner, an academic, an athlete etc.), but such participation is part of the foundations of selfhood, rather than a loss of it.

In response, one could claim that the idea of 'dissolving' into the dimension of participation is losing one's *authenticity*, i.e. failing to resolutely stand behind one's self-owned behaviour and thereby simply doing what others do (Heidegger, 1927/1962). This seems like a useful concept for describing how one can properly distinguish oneself from social processes, but unfortunately it, too, is unsatisfactory for Kyselo's account. This is because the vast majority of people rarely behave authentically – we all mostly do things simply *how they are done*, without critically reflecting on them. In Kyselo's framework, this would seem to suggest we are either (a). consistently in the realm of 'social dying', in virtue of participating with others to the extent that we lose our authenticity, or (b). never actually succumb to fully losing oneself in the dimension of participation and instead always retain some degree of self-owned subjectivity. Neither option is viable for Kyselo, in that the former would suggest that the self is a series of synchronic social 'lives', punctuated by various deaths, and the latter would suggest that social death simply doesn't exist (which is the concession that I believe has to be made – see section 5.1).

Similarly, there are very good reasons for doubting the possibility of a 'social death' within the dimension of emancipation. For instance, even if one is completely socially isolated, one still lives in accordance with certain social norms; indeed, as long as one cognises one is bound to certain social ways of being (see chapters 4 and 5). This raises the interesting question of how exactly Kyselo is conceiving of 'participation'. She initially states that

participation encompasses “virtual or actual interpersonal engagements of at least two individuals” (Kyselo, 2014, p. 8), but how far ‘virtual’ can be stretched is not entirely clear. If a virtual interpersonal engagement can involve interaction with social norms, then it seems we can never be fully emancipated from social processes, in line with the aforementioned idea that meaningful existence is bound to certain social ways of being. If, on the other hand, a virtual engagement is limited to interactions in which individuals are not (directly) physically present to one another (e.g. via a video-link), or perhaps even to imagined discourse between (at least) two individuals, then it seems we are ‘fully’ emancipated whenever we are alone and not thinking of others. The idea that social deaths become a frequent occurrence (see previous paragraph) thus resurfaces.

Kyselo has a footnote (n.7) in which she claims that “social relationality need not translate to actual engagement, or actual interaction, with others”, and also that “*participation* does not mean to cease to exist as separate individual; were this the case, no *one* would be distinguishable as participant” (p.9). This would seem to assert that we are always bound up in sociality *as an individual* to some degree. Indeed, it would seem an excellent reinforcement of human needful freedom. However, exactly how this assertion can be reconciled with the possibility of ‘social death’ remains unclear for the reasons discussed above.

Even if we ignore the vagueness of exactly *how* one succumbs to a social death, it still seems that the notion is problematic for Kyselo’s account, in that it suggests an imbalance between bodily and social processes. What I mean by this is that even if one were to suffer a social death, one’s body could still survive. A cult member who is ‘lost’ in the dimension of participation still has their own body with its own biological autonomy and a socially confined individual who is ‘lost’ in the dimension of emancipation would similarly still have their own body with its own biological autonomy. Conversely, if one suffers a bodily death, it is hard to see how any social self could persevere. To put it simply, if one takes the view that ‘social death’ is possible, then there seems to be good reason to endow bodily processes with primacy when it comes to individuating the self, as bodily processes can persist in spite of a social death but social processes could not persist in spite of a bodily death.

The difficulty of social death also raises an ambiguity within Kyselo’s theory of an enactive self. As mentioned in section 2.2, what is arguably unique about human needful freedom is that we move beyond the biological autonomy of our organismic bodies and also live through the relational autonomy of our social interactions. However, the relation between these two forms of autonomy, or how they interrelate, is not entirely clear. For example, one could draw on the aforementioned issue surrounding ‘social death’ to suggest that one can lose one’s social autonomy and still survive, but one cannot lose one’s biological autonomy and survive. Again, this reformulates the claim that bodily processes (i.e. the biological autonomy

of the organismic body) should be given primacy over social processes with regards to the individuation of the self.

Alternatively, one could argue that when engaged in extreme participation it is actually biological autonomy that loses out; for instance, if cult members engage in mass suicide then the relational autonomy of their cult interactions seems to have ‘overridden’ the self-preserving purposiveness of biological autonomy. This would (re-)endow social autonomy with some of its grip on selfhood, but it still wouldn’t be entirely satisfactory: the suggestion here would be that biological autonomy and social autonomy are opposing poles on a spectrum, and any spectral polarity – particularly a polarity in which either pole can fully ‘override’ the other – undermines the notion of a genuine unity between bodily and social processes.

Two final points will serve to underline the confusion surrounding the issue of ‘social death’. Firstly, without a greater deal of clarity, one could be forgiven for supposing that the spectrum of emancipation and participation is supposed to map onto a spectrum of bodily (or strictly biological) and social processes as two poles within a domain of enactive autonomy. So when one is emancipated one is engaging with the organismic autonomy of one’s individualistic body and when one is participating one is engaging with one’s social autonomy. Yet this assumption would only further complicate matters, in that one is seemingly capable of emancipating oneself whilst being highly socially autonomous (e.g. ‘owning’ one’s actions as the star member of a team), or being highly participatory whilst being biologically autonomous (e.g. recalling and employing sociocultural norms or interactions in order to survive as an isolated shipwreck victim).

Secondly, even with the issue of polarity set to one side, there is further confusion in virtue of the fact that one’s social organisation can seemingly furcate within a situation. For example, dance performers on a stage need to collaborate in a coordinated and often intimate manner, exhibiting high levels of participation, but at the same time they may be sharply emancipated (or, at the very least, notably different in their participation) from their audience, as well as being potentially emancipated from or participating with various other groups such as stagehands, directors or accompanying musicians. The temptation would be to say that the on-stage interaction between the performers is ‘dominant’ and it is this participation that matters. Yet it seems undeniable that the presence and phenomenological relevance of the audience and other groups are certainly not negligible. Alternatively, one could claim that there are distinct spectra of self-organisation relative to the situation: that between the performers, and that between the performers and the various other groups (such as the audience). Yet this is also dissatisfying, in that the self either seems to be somehow split across constituting interactions, or the manner in which it integrates these forms of interaction is insufficiently explained.

We thus seem to have a number of messy and unsatisfactory options regarding participation-*emancipation*, body-social processes, biological-social autonomy, and how all of these concepts connect to one another. Without genuinely unifying the biological body and social structures, which will be my objective in sections 4 and 5, Kyselo's enactive approach to the self seems to lose grip on the cogency that at first seems apparent.

### 3.2. Losing the Body

A further issue for Kyselo's account of the self stems from her claim that the body is the "self's means and mediator" (p. 12). Whilst this clearly suggests that the body has an integral, non-trivial role in the generation and maintenance of selfhood, the situation becomes somewhat muddled when we look at things more closely. Consider the following claim:

the self is not just a lo[o]se collection of aspects but has boundaries that are generated through interacting and being related to others. The self in its most minimal sense, thus escapes the body. It is never fully separable from the social environment, but instead determined precisely in terms of the types of social interactions and relations of which it is, at the same time, a part. (Kyselo, 2014, p. 12)

The emphasis here is clearly on the social individuation of the self, so much so that the body can in fact be 'escaped' from. Whilst Kyselo stresses that the claim is not in favour of "a disembodied conception of the self" (*ibid.*), the cited passage is telling of the fact that Kyselo's self ultimately seems to be a social self that just so happens to have *this* body as its mediator, but could perhaps satisfactorily inhabit the world through any appropriate bodily medium.<sup>[41]</sup> She seems so committed to overcoming the claim that the self *is*, in some way, the body that she ends up pitching herself in the camp of 'the self is a social construction'. This obviously jars with her attempt to delineate a coherent notion of self that overcomes the 'body-social problem' by giving primacy neither to bodily processes nor to social processes.

In order to fully understand the imbalance that Kyselo eventually seems to favour, we can consider further her arguments regarding the body. Take as an example how Kyselo describes getting a haircut as not merely being "a change to some biomaterial that grows out of my head" (p. 12), but rather as a change to "how I saw and now see myself and[...] how others have seen and now see me" (*ibid.*). That is, a haircut is a bodily change that transforms how one relates to the world. Importantly, however, the transformation of this

---

<sup>[41]</sup> The ambiguity of the body's role in Kyselo's enactive approach to the self indicates an underlying confusion of the body as an ontological concept. At times, the body seems to be an integral constituent of existence – an existence that is permeated by socio-cognitive complexity. This existentially constitutive body is more than mere biological matter: it is an autonomous system without which cognition is impossible and through which all cognitive activity is manifest, as any cognitive process ultimately refers to the ongoing preservation of the body (see Di Paolo et al. (2010) for more on this). However, elsewhere, Kyselo seems to treat the body as aligned with the purely biological organism, which is distinct from (albeit importantly interlinked with) sociality. For example, the principle of a human individuating herself *through and from a world* involves "bracket[ing] for a moment any role the body might play in the individuation of human cognitive identity and[...] instead consider[ing] human individuation as a social process" (Kyselo, 2014, p. 8). Such 'bracketing' suggests that the body is no longer existentially constitutive in a full-blown sense and is instead akin to an organic, causally influential medium.

bodily change only becomes significant through an “act of relationality” (ibid.): either how one relates to oneself or how one relates to (or is related to by) others. This would seem like a perfectly reasonable description to endorse if it were not for the fact that the body no longer seems to be playing a truly *constitutive* role in the individuation of the self. Note, for instance, that the bodily change does not impact on one’s selfhood until it “acquires a social meaning” (ibid.) during some sort of interaction. The body is dislodged from the inherent constitution of the self and relies on integration with social processes for modulating the self. This is, of course, in keeping with Kyselo’s claim that the body is “the self’s means and mediator”, as it suggests that the body does not directly impact self-individuation, but instead only has a mediatory impact on one’s social relationality. The body is thus distanced from imminent self-individuating power, only having relevance as a medium of social activity.

In Kyselo’s defence, she does claim that a bodily change may be implicitly significant to one’s selfhood because she views relating to oneself as an inevitably social process (p. 12). She also takes the positive step of building on Thelen’s (2004) work to suggest that bodily activity, being induced by social interaction, can shape neural processes rather than neural processes having unidirectional control over bodily activity. However, the positive movements that she makes in favour of the body’s integral role in selfhood are undermined by its status as a *mediator* of a primordially important social self. As with the issue of ‘social death’, we once more seem to have just fallen short of achieving a genuinely coherent *unity* of bodily and social processes. What is perhaps most worrying is that whereas the issue of social death resulted in the body seemingly being given primacy with regards to the individuation of the self, the current issue suggests that social processes have primacy. The exact conjunction of Kyselo’s account of an enactive self is thus submerged in uncertainty.

#### **4. Can Kyselo Respond?**

The issues of social death (section 3.1) and the body being merely a mediator for selfhood (section 3.2) both emerge from a failure to satisfactorily unite bodily and social processes. In this section, I will (very) briefly offer potential responses to these issues from within the confines of Kyselo’s account (section 4.1). Subsequently (section 5), I will put forward a new claim for human selfhood that offers a more robust theory of individuation in virtue of emphatically expounding the ontological unity of ‘body’ and ‘social’ with regards to human existence.

##### **4.1. Brief Responses from within Kyselo’s Framework**

Firstly, let us re-consider the following difficulty: that Kyselo’s introduction of the possibility of a social death results in her account seemingly favouring the primacy of the

body in the individuation of the human self. Whereas bodily processes could seemingly persist in the face of a social death, it is hard to see how social processes could persist in the face of a bodily death (see section 3.1). However, if Kyselo were able to account for the persistence of social processes in the face of a bodily death, then the balance between social and bodily processes would be restored (albeit at the fulcrum of a seesaw). Such a solution puts selfhood at a precarious juncture between social and bodily 'deaths', but it is the only solution that I believe is available within the current framework of Kyselo's account.

The issue depends, in part, on exactly what Kyselo means by 'death', which seems to be the loss of the ability to 'oscillate' between the poles of participation and emancipation, i.e. the loss of the ability to properly exercise one's needful freedom (and thereby preserve a social identity). 'Death', in this sense, is not a full-blown cessation of existence, but a cessation in the specifically human mode of social existence. Working with this esoteric notion of death, one could make the argument that just as a bodily identity can survive a social death, so, too, can a social identity survive a 'bodily death', meaning that neither bodily nor social processes have primacy over the other in the individuation of the human self. 'Bodily death', here, would be the loss of the ability to properly exercise one's bodily autonomy (with a continuously preserved identity). On this view, even in the cases of a complete enslavement of one's body or a comprehensive surgical transformation of one's body, one could still preserve some sense of being a social self. For example, an individual whose physical appearance is completely altered would still retain some sense of being the same (social) person. Indeed, people who undergo gender reassignment often claim that the process is one of finding a body to suit the self they have always been; an extreme version of 'reassignment', in which one undergoes a complete physical overhaul, could thus be easily construed as a transformation to accommodate one's pre-reassignment conception of self, rather than becoming a new self. Further to this, as Kyselo points out, sufferers of locked-in syndrome, which is an occurrence of complete bodily paralysis (except for eye movements), also retain a sense of self in spite of their physical incapacity. This is due to patients' preservation of the ability "to engage with others, be recognized and experience themselves as subjects" (Kyselo, 2014, p. 4; see also Gosseries, Bruno, Vanhaudenhuyse, Laureys and Schnakers (2009), and Lulé, Zickler, Häcker, Bruno, Demertzi, Pellas, Laureys and Kübler (2009)). Inasmuch as a bodily identity or bodily sense of self can seemingly persist in the face of a 'social death', one can similarly claim that a social identity or social sense of self can persist in the face of a 'bodily death'. The problem of social death resulting in a favouring of the body in the individuation of the self – as the body could survive a social death – can thus be counter-balanced by the idea that a social identity could analogously survive a 'bodily death'.

With regards to the second issue of the body being *merely* a mediator of selfhood (see section 3.2), rather than inherently constitutive of selfhood, Kyselo could respond by

emphasising the idea that the mediation of the social self by the body is in fact *reciprocal*, so the bodily self is also mediated by the social. For example, in an analogous fashion to the body being an ‘interface’ for social interaction, the social world is an interface for bodily expression. That is, how one carries oneself (i.e. posture), how one moves and how one talks are all mediated by one’s social situation. Further to this, how one’s body develops or changes is mediated by social factors, in virtue of the daily activities that one engages in, the nutrients that one ingests and the clothes that one wears. In enactive terms, the idea is that one’s biological autonomy is mediated by social aspects of self-preservation and, reciprocally, one’s social autonomy is mediated by biological aspects of self-preservation. The idea should thus not be that the body is in some way removed from the direct constitution of selfhood and is relegated to a mediatory role, but that bodily constituting processes are themselves also mediated by social processes. In this way, the balance of ‘primacy’ between bodily and social processes is preserved by the reciprocal co-mediation of biological and social autonomy.

This final idea – that bodily and social processes are *co*-mediatory – is perhaps the most important to a proper account of selfhood, as it naturally implies a genuine unity between body and sociality. However, the above attempts to resolve the ‘primacy imbalance’ between bodily and social processes remain contentious because they are delivered from within Kyselo’s framework for an enactive self; that is, the ‘balance’ that is achieved would seem prone to perennial fluctuations between bodily and social processes, which are still treated as polar measures that must be integrated, rather than as genuinely consonant. In spite of Kyselo’s efforts, the ‘tension’ that she seeks to overcome is not eliminated: bodily and social processes still seem to be in opposition and their struggle is simply delicately navigated. It is thus debatable whether the responses offered on behalf of Kyselo in this section actually dismiss the problems that were outlined in section 3. In order for the sought-after unity between bodily and social processes to be unquestionably achieved, I believe that a fresh approach needs to be taken to how humans navigate their existence. It is this approach that I will turn my attention to in the remainder of this chapter, delivering an outline of the *biosocial self*.

## **5. The Biosocial Self**

In order for there to be genuine unity between bodily and social processes, ‘body’ and ‘social’ need to be treated not as distinct aspects of selfhood that just so happen to be closely entwined, nor as opposite poles within a spectrum of selfhood, but rather as a single ontological feature. The *co-mediation* between bodily and social processes that is mentioned above is a positive step towards achieving a unified conception of body and social. Consider, for example, how seemingly basic bodily features such as one’s size and shape, one’s heart

rate, one's skin condition and one's reflex speed are all, in part, socially modulated by what one does on a daily basis. Even seemingly private bodily functions will be conducted within a sphere of social normativity, at the very least in virtue of the fact that one sees such functions as suitable for privacy. Reciprocally, the body is the site of enaction for any social encounters, attitudes or inferences; it is, as has been argued in many quarters (Johnson, 1987; Merleau-Ponty, 1945/2012; Shapiro, 2014; Varela et al., 1991), an essential constituent in all cognitive processing. However, in spite of the clear importance of sociality to bodily processes and the body to social processes, conceiving of the two sets of processes as connected in such a way that they cannot be disentangled is rare. As we saw earlier on, the problem is that bodily and social processes tend to be seen as 'poles', or in some way distinct from one another, such that one set of processes merely mediates the other primordially important set.

To overcome this, I am going to address three aspects of human existence in order to detail a theoretical picture of the *biosocial self* – a mode of being in which biological-bodily and social events are singular ('biosocial') ontological processes. The first step is to re-formulate the unique experiential space in which humans dwell (section 5.1). I will then show how dwelling in this space renders humans *biosocial* entities, from both a developmental perspective (section 5.2) and into the matured stages of life (section 5.3).

### 5.1. The Biosocial Experiential Space

In this section, I will explain why 'body' and 'social' cannot be disentangled from a human point of view, in virtue of humans inhabiting a unique experiential space. The first step in making this claim will be to consider the uniqueness of human experience through the works of McDowell (1994) and Maturana (1975, 1990, 1995, 2002). Rather than building on these works, I introduce them simply as a means of establishing the idea that human experience is distinct from the experiences of the rest of the animal world. For both McDowell and Maturana, our experiential uniqueness is due to our use of language, whereas for my account of biosocial selfhood it is due to the unity of bodily and social processes, such that there is reciprocal co-modulation of our bodies as socially expressive and our sociality as bodily expressive.

The idea that human experience is unique is not, in itself, a unique claim. Most people would agree that the demonstrable intelligence of humans suggests that we experience the world differently to other animals. McDowell (1994) explains this by claiming that humans have *rationality* that other animals lack – we can conceptualise aspects of the world and decide how to think about and act upon them. For non-human animals, on the other hand, "sentience is in the service of a mode of life that is structured exclusively by immediate biological imperatives" (ibid., p. 115). This means that non-human animals exist in a mode of being that is characterised by responding to problems and exploiting opportunities

connected to “immediate biological forces” such as brute survival and reproduction (*ibid.*). Meanwhile, humans “rise above the pressure of what impinges on us from the world [...into a] free, distanced orientation” of conceptual possibilities (Gadamer, 1960/1992, pp. 444/445). This is not to suggest that we are in some way separate from the natural world, but that the space of rationality emerges within it, such that our lived subjectivity is always conceptually structured. On this view, there are no non-conceptual processes that exist as precursors to, or components of, conceptual ones – human experience *always* unfolds through conceptual processes. Whether a human is contemplating an abstract dilemma or performing a routine bodily movement, the experiential characters of such acts belong to the same space of rationality.

Maturana, whose autopoietic theory (Maturana, 1975, 1990, 1995, 2002; Maturana, Mpodozis and Letelier, 1995; Maturana and Varela, 1980, 1987) is an antecedent form of much modern enactivism, also contends that humans experience the world in a unique way. Rather than specifically separating humans from non-human animals as McDowell (1994) does, Maturana (1990, 1995) claims that every living species occupies a unique experiential domain. Through a recent exegesis of Maturana’s work by Villalobos and Ward (2016), we can see that this claim begins with the positing of two basic domains of existence for all living beings: (i). the *physiological* domain, which is constituted by the structural dynamics of the modalities and divisions of an organism’s sensory systems, and (ii). the *relational* domain, which is constituted by the behavioural interactions of an organism (as a totality) with its environment (Villalobos & Ward, 2016). The subsequent claim is that (i) provides an organism with basic sensory experience and (ii) provides an organism with a “characteristic mode of life” (Villalobos & Ward, 2016, p. 209). This mode of life is generated through an organism relating to its surrounding world in a specific way and, consequently, living through specifically structured relational *experience*. Just as certain sensory systems are assumed to give rise to certain sensory experiences – for example, visual systems will generate visual experience, regardless of the exact mechanics of the involved system<sup>[42]</sup> – so, too, will an entity’s behavioural dynamics give rise to certain relational experience. This claim draws on the enactivist tenet that every living entity is *structurally coupled* to its environment, such that there is a history of recurrent congruence between the dynamic changes enacted by a living entity and related changes in the entity’s environment (Maturana & Varela, 1987, p. 75; Thompson, 2007, p. 45)). Resultantly, the relational space occupied by a cat generates ‘cat-experience’ that is incommensurable with the relational experience

---

<sup>[42]</sup> This is not to deny the potential for plasticity in how experience is generated (e.g. see Bach-y-Rita, Collins, Saunders, White and Scadden (1969) on the possibility of substituting tactile sense for vision), or the possibility of experiential variation through neurological abnormalities (e.g. see Harrison and Baron-Cohen (1997) on synaesthesia). The point is rather that (natural) visual systems are assumed to generate experience that belongs to the domain of visual phenomenology, whether these systems belong to fish, reptile, mammals, birds, insects etc. (Villalobos & Ward, 2016, p. 209). The modalities of visual systems will not (in general) bring about olfactory or auditory experience, regardless of the vast architectural and mechanistic differences in visual systems across the range of living entities.

enjoyed by other living entities, just as visual phenomena brought about by visual systems are incommensurable with sensory phenomena brought about by other sensory systems (ibid.). Accordingly, humans enjoy a unique kind of 'human-experience' that is distinct from the experience of all other living entities. Rather than defending these claims, I simply wish to appropriate the idea, as I am doing from McDowell (1994), that there is a unique experiential world that humans inhabit.

A notable harmonious connection between the accounts of McDowell and Maturana is the fact that they both endorse language as the mechanism by which human experience obtains its complex character. That is, it is through initiation into language that humans enter into the domain of conceptual rationality that is unique to our species (McDowell, 1994; Maturana, 1995, 2002; Maturana & Varela, 1987; see also Gadamer (1960/1992)). However, it is here that I wish to make a clear departure from McDowell and Maturana. I do not think that our ability to form and act on conceptual thoughts – thereby occupying a unique experiential domain of rationality – rests on our ability to use language. Instead, I believe such abilities, *including* the ability to use language, are subserved by the inherent communicatory potential of our bodies. As we will see in the next section (5.2), humans are potentiated from birth to recognise and interact with other humans, with our bodies being the means by which such social accomplishments are achieved. Such bodily engagement in social activity is the first indication of human cognitive engagement with the world. For instance, newborns have been shown to respond to human interaction within minutes of birth, often imitating facial gestures (Meltzoff and Moore, 1977; Gallagher and Meltzoff, 1996). Importantly, they only imitate facial gestures from other humans (Johnson, 2000), suggesting that there is an inherent human disposition towards parsing the world into human and non-human entities (Meltzoff and Brooks, 2001). After these nascent abilities to recognise and interact with others are established, newborns gradually begin to display further capacities such as gaze-following at approximately 9 months of age (Senju, Csibra and Johnson, 2008), joint attention at 9-14 months (Phillips, Baron-Cohen and Rutter, 1992) and comprehension of goal-directed behaviour in a specific context at 18 months (Meltzoff, 1995; Meltzoff and Brooks, 2001). Collating these empirical studies gives us two crucial pieces of information: firstly, human newborns seemingly have inherent embodied capabilities *for* social interaction; secondly, through these embodied interactions with others, infants begin to build a repertoire of increasingly complex cognitive capacities.

The point here is that whereas McDowell and Maturana contend that human occupation of a unique experiential domain rests on language (in the traditional sense of verbal language), I am claiming that such experience is present from the earliest moments of life through our nascent capacity for bodily-social communication. As 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*.<sup>[43]</sup> In the terms of the social normativity that has played a key role across the last three chapters, the idea is that any bodily activity is socially normative and these social norms then feed back into individuals' canalised range of potential bodily actions which will, in turn, generate and modulate further norms which will then instantiate further feedback (and so on). We are distinct from 'mere animals' because we inhabit an experiential world of social normativity that is in no way separate from our organismic bodies, but is instantiated and maintained by them. This normativity, which is literally embodied across collectives of individuals and to which we are inherently disposed, serves as a constant store of historically accumulated cognitive possibilities that are permanently in a state of dynamic modulation by intersubjective activity. On this view, there can be social normativity without natural language because our bodies take up the linguistic mantle and convey meaningful norms in the absence of verbal accompaniment. We are human because we engage with this normativity, living through a biosocial world in which embodiment and ensocialment are unified.

To understand this claim in more detail, recall that any bodily activity – even something trivial such as walking down the street or asking a stranger for directions (see chapter 5) – will potentially modulate the normative possibilities of how that activity is executed, for oneself and for others. To move one's body is not to merely perturb the physical world, it is also a communicatory perturbation of the socio-normative world that we inhabit. So the idea is that any bodily activity has a normative component that is an expressive modulation of possibilities for action for others. Such normativity first takes hold through the primordial communicatory capacity to sense otherness that was discussed in chapter 6 (i.e. the capacity of non-reflectively acquiescing to another's ability to be an active generator and modulator of the social norms through which life is lived). Acquiescing in this way is a kind of foundational structure on which all other normativity (and thus cognition) is subsequently built. Importantly, the social act of sensing otherness is only possible in virtue of agential physicality. For example, the capacity to acquiesce to another's ability to generate and modulate social norms only makes sense if there is a relational medium in which such acquiescence takes place. For humans, and, indeed, any hypothetical entity that one could think of, this necessarily requires a (specific) body (with specific normative status). One cannot sense and acquiesce to another's agency, or be reciprocally acquiesced to oneself, if there are not specific bodies in which the other's and one's own unique normative potentials are manifest. The very bonds of social normativity in which our existence is rooted are thus bodily manifest at a foundational level and then continuously modulated through bodily

---

<sup>[43]</sup> I favour the use of 'biosocial' over 'bodily-social' for two reasons. Firstly, 'biosocial' is a more compact and evocative term. Secondly, with the idea of biosociality established, 'body' is already a social concept, so to speak of bodily processes is, strictly speaking, to simultaneously speak of social processes. This renders 'bodily-social' an over-determining phrase.

expressivity and interactions. The sociality with which a body is permeated is, in the first place, bodily generated in tandem with the social ability to sense otherness. Humans are thus *biosocial* entities because we occupy a relational domain of experience in which our bodies are (socio-)normatively laden, so as to generate a holistic network of norms through which behavioural and cognitive dynamics are looped.

Running with this idea, one could claim that our bodies are close to being ‘linguistic’ in the sense employed by Cuffari, Di Paolo and De Jaegher (2014), in that we are bodily sensitive to the social world in such a way that intersubjective activity is habitually rendered intelligible. Indeed, my view of human bodily existence being biosocial chimes happily with Cuffari et al.’s (2014) claim that our “world-engagement is an integrated whole of embodied interpreting[...] embedded in horizons of social normativity” (p. 1115). However, care needs to be taken in not underselling the fundamental nature of the biosociality that I am proposing. It is important that embodiment and ensocialment are not separable à la Kyselo’s theory of selfhood. Unlike Kyselo’s enactive approach, in which bodily and social processes occupy separate spheres of activity that must be reconciled with one another, the biosocial approach to existence rejects the possibility of fundamental separation for bodily and social processes. Having a normatively laden body is not something that we gradually develop as we grow into a sociocultural world; rather, by occupying a relational domain of biosociality, it is something that we ontologically *are*. Our biological bodies and social environment belong to one and the same experiential space – a space that makes us unique amongst living beings.

With this in mind, it is no longer enough to say that bodily existence involves a living body *in* a social world; instead, it is *being a living social body*. Human corporeality is always more than mere flesh: the body is ensconced in social purpose so that it simply *is* a social entity. The mistake that Kyselo (2014) rightly wishes to avoid (see section 2) – that social or bodily processes are thought of as merely contextual – is now fundamentally bypassed. Without the body’s entanglement with sociality, the meaning that an agent generates through its environmental activity would be little different from the basic self-perpetuation of sucrose-seeking bacteria or sunflowers. It is the imbrication of sociality and a lived body that distinguishes human ontological make-up from that of non-human animals, plants and basic autopoietic structures. For a body to be just a body in the strictly physical sense is for it to be extracted from the collectively mediated social norms that are constitutive of humanity’s lifeworld (Ikäheimo, 2009, p. 36). And without the body, there is no relational medium in which the generation and modulation of social norms can take place. There is thus ongoing and reciprocal iterative feedback between our embodiment and ensocialment, which are co-constitutive features of our biosocial existence.

Having outlined biosocial existence, we can now consider how this concept can respond to some of the issues facing Kyselo's enactive approach to selfhood. Firstly, there is the fact that the concept of a 'social death' (and the connected potential for a 'bodily death' (section 3.1)) is now nonsensical. The idea that there cannot be non-socially normative bodily processes means that a 'social death' is simply a straightforward death. In the terms of Kyselo's theory, to lose the capacity to socially interact would simply be to cease bodily activity (and, conversely, to lose the capacity to bodily interact would be to cease social activity). The ideas of a cult member who is 'lost' in the social dimension of intersubjective participation or a shipwreck victim who is 'lost' in the dimension of emancipation, so that each is suffering a social death of sorts, are now untenable. Both cases preserve the unity of bodily mediated social processes and socially mediated body processes, so that the temptation to view bodily and social processes as part of the polarity between emancipation and participation actions is now removed. Emancipation and participation both involve enacting (body-)social norms through one's living (social-)body. 'Body' and 'social' are not opposing poles on a spectrum, nor are they in tension. They are conjointly foundational to the manifestation of one another, so that any attempt to distil the ontological 'essence' of humanness cannot be done without the imbricated notions of both body and sociality.

Secondly, the idea of the body being a non-specific mediatory vehicle of the primordially social self (section 3.2) can now also be rejected. Due to the fact that we inhabit a biosocial domain of existence, the body is never merely a causal or contextual passenger to social experience. It always has a characteristic of potentiated normativity that renders movements socially determinable. The body is therefore an integral constituent of our social experience just as social relations are an integral constituent of bodily experience. As mentioned previously, the body is a social entity and so any bodily change is simultaneously a social one.<sup>[44]</sup> Kyselo's claim that a haircut is a bodily change that "*acquires social meaning*" (2014, p.12; my italics) should now be replaced with the idea that it is a direct biosocial change. There is no intermediate step between bodily and social changes as there is no longer any experiential wedge that can be driven between them. It is not just that the body is indispensable to human cognition, as it is for any entity, but that human bodies only ever take form in an intra-specific social world through which we generate meaning and make sense of personal existence. The extraction of bodily processes from social processes is

---

<sup>[44]</sup> Some people may raise the following concern at this point: surely there are minor, unnoticed or internal bodily changes that are *only* bodily occurrences, i.e. not social at all? My simple response is 'no'. To ask this question is to fail to appreciate the fundamental entwinement of bodily and social processes that I am proposing. Once the body is reconceived as a social entity there is no such thing as a purely physical change. Some changes may be too subtle for others to notice – or, indeed, for a subject to notice within herself – but this does not mean that such changes aren't taking place, either from a physical or socio-normative perspective. Just as a haircut is directly a biosocial change, with no mediation between bodily alteration and social appreciation of this alteration, so, too, is any internal alteration a biosocial change. Proper appreciation of a *living social body* involves appreciating the fact that every aspect of the body contributes to one's phenomenological experience of the world. The notion of a gendered body (section 5.3) is useful to illustrate this point, in that every bodily change is filtered through one's gendered conception of how this change *should* be experienced.

simply impossible from a human perspective, in that the loss of either must entail the loss of the other. With the simple step of elucidating the *genuine* coalescence of ‘body’ and ‘social’ into biosociality, we thus find that the problems facing Kyselo’s account – ‘overcoming a social death’ (section 3.1) and ‘losing the body’ (section 3.2) – can now be avoided.

In order to further strengthen this idea of human biosocial existence, I will now look at biosociality from both developmental and everyday perspectives.

## 5.2. Biosociality in Development

For Kyselo, human selfhood emerges as biological bodies are individuated through social interactions – “it is an achievement[...] *between* individuals” (2014, p. 8). According to this view, we “become a bodily identity” (ibid., p. 12) at birth, but we do not achieve selfhood until we have individuated ourselves from other human subjects by oscillating between the interactive processes of participation and emancipation (ibid.). According to my view, however, any bodily identity is a biosocial self, in that human bodily processes always occur within the experiential space of normative biosociality. Selfhood is thus not ‘achieved’ on the back of a nascent bodily identity, but is present with the earliest indicators of individuation. Indeed, many early-life studies of neonates support the idea that bodily activity is, from the burgeoning moments of being, socially meaningful and receptive.

Firstly, let us consider the straightforward claim that neonatal bodily behaviour is inherently socially meaningful. It is a simple fact that neonates are entirely dependent on social interaction for survival. Without the nutritional and practical support of caregivers, neonates simply wouldn’t survive. Whether or not a neonate *intends* her crying, writhing or other bodily activity to convey a certain message, they all contribute to the normativity of the human experiential domain. As long as a neonate is bodily present with others, she is a modulator of human existence. Conversely, without immersion in the sociocultural world, the persistence of a body simply wouldn’t be a genuine (human) self: it would be purely the persistence of biomatter in a strictly material world.

Importantly, neonate behaviour is responsive to that of others and, from very early on, neonates will meaningfully interact with other humans. For example, there are a plethora of studies that endorse the neonatal ability to imitate tongue protrusion and mouth opening (Anisfield, 1996; Meltzoff and Moore, 1977, 1983; Nagy, Pilling, Orvos and Molnar, 2013), as well as imitating index finger protrusion (Nagy, Kompagne, Orvos, Pal, Molnar, Jansky, Loveland and Bardos, 2005) and emotions such as happiness, sadness and surprise (Field, Woodson, Cohen, Greenberg, Garcia and Collins, 1983; Katiz, Meschulach-Sarfarty, Auerbach and Eidelman, 1988), with some of these abilities coming within the first few hours of birth (Kugiumutzakis, 1985; Soussignan, Courtial, Canet, Danon-Apter and Nadel, 2011). If we endorse the widely accepted view that imitation requires a body schema – that

is, pre-reflective sensorimotor abilities for enabling and constraining appropriate bodily movement in response to environmental stimuli (Gallagher, 2005, pp. 24-38) – then these studies would suggest that human neonates possess “a primitive body schema from the very beginning” (Gallagher, 2005, p. 72). One could be tempted to use this evidence to rejuvenate the suggestion that bodily processes should be considered primary in the individuation of the self, in that we seemingly have an innate body schema that underlies subsequent development. However, I believe it is wise to forego this suggestion for the following reason. Firstly, Gallagher (2005) describes a body schema as the operation of sensorimotor systems “without awareness or the necessity of perceptual monitoring” (p. 24), which suggests that such schemata should be viewed as pre-reflective constraints on experience, rather than features *of* experience. That is, body schemata are simply a result of being bodied and are thus available to myriad bodily beings, human or non-human. The imitative abilities of neonates, on the other hand, are intersubjective achievements as much as bodily ones and are indicative of the kind of biosocial experience that I am claiming is unique to humans. Neonates display not only the presence of a bodily schema, but also an intermodal awareness (achieved via proprioception and perception) of the structural similarity between oneself and a human other (*ibid.*, p. 75). Whilst non-human animals are able to enjoy body schemata, in virtue of having functioning sensorimotor systems, it does not follow that they have the biosocial capacity to (pre-reflectively) employ such schemata for imitative purposes. To put this another way, if one’s concern is human selfhood, as it is in this chapter, then displaying motor activity in response to sensory excitations is presumably not enough – even plants demonstrate a kind of motor responsiveness in virtue of their various tropisms (i.e. directional growth responses towards or away from specific stimuli). What is needed is some feature of relational existence that is unique to human experience (see previous section), which I have argued is captured by our biosociality and which, in this case, is evidenced by bodily imitation. The neonatal empirical data thus supports the idea that bodily activity is socially imbued, in that neonates are inherently sensitive to social interaction from the first tentative moments of post-pregnancy life. Indeed, the above studies all bolster the notion of humans having the fundamental biosocial ability to sense otherness through bodily-social interaction (see also Gallagher (2005) and Ikäheimo (2009) for similar claims regarding our basic ability to recognise the animacy of other humans, as well as Farroni et al. (2002, 2006) for evidence of neonatal preferences for human direct gazes from upright faces).

Taking this idea to an even earlier developmental stage, there are suggestions that fetuses increase bodily activity – arm, head and mouth movements – in response to maternal speech and maternal touch of the abdomen, with the responses increasing in conjunction with foetal maturation (Marx and Nagy, 2015). Such evidence would suggest that even prior to birth

there is a significant entanglement of bodily and social processes, with foetal bodies being interactively responsive to social stimuli over other worldly events.

The exact details of when a foetus or neonate becomes a human self are not my concern here. The point is that at whatever time one wishes to claim a neonate develops awareness of self, which is typically taken to emerge with basic bodily self-awareness (see, for example, Gallagher (2005, pp. 72-85) on minimal embodied subjectivity), one must also acknowledge that such awareness is simultaneously a modulatory normative process and, therefore, a social activity. The claim is not that bodily (self-)understanding emerges *within* a social world, but that it is concurrently a kind of social (self-)understanding of one's own presence (and modulatory capacity) that belongs to the normative experiential world of biosociality.

What we can now turn our attention to by looking at the issue of gender is the manner in which our biosociality results in our bodies being more than merely physical entities in our everyday lives.

### 5.2. Everyday Biosociality

Recall from section 4.2.2 in the previous chapter that gender is best viewed as a social construction that is not rigidly aligned with biological sex (Laurie, Dwyer, Holloway and Smith, 1999). In Simone de Beauvoir's (1949/2011) famous words, "One is not born, but rather becomes, woman" (p. 330). There is, in short, no fixed gender on the pure basis of an agent's physical make-up; gender is instead something that is always "culturally mapped" onto a body (Valentine, 2001). Typically 'feminine' actions are, therefore, simply the consequence of socially generated norms that persist through the ongoing gendered treatment of what men and women are capable of and how they 'should' act. This is not to completely deny the importance of genetics in one's lived world; as Grosz explains, our bodies are both "involuntarily marked" (1994, p. 142) by our genetic inheritance and "tattooed" (1993, p. 12) by our socio-cultural existence. What matters for present concerns is that bodily 'tattooing' is, in some form or another, inevitable and ubiquitous in human existence. The purposefulness of one's world – what it affords and how one is poised to act on it through one's living social-body – is contingent on one's 'tattooed' existence. For example, as Young (1990) explains in response to the work of Straus (1952), the fact that young girls will often throw a ball with little lateral arm movement and minimal rotation of the body is not due to any innate 'femininity', but rather due to the situatedness of young girls' bodies within patriarchally gendered society. Such situatedness results in young girls (and, indeed, women more generally) existentially relating to the world in a manner that differs from that of men: as burdened by their bodies, inhibited in intentionality and discontinuous in activities (Young, 1990, pp. 145-148). To put it simply, women often live through their social-bodies as 'mere objects' – as fragile entities that must be cajoled into

activity and exist “*as looked at and acted upon*” (ibid., p. 148) – instead of as transcendent subjectivities. In other words, women are ‘tattooed’ with the social normativity of gendered modalities. Their bodies become something other than a mode of lived subjectivity: they become normatively laden by societal expectations and mores. In many societies across the world, there are marked distinctions arising in female and male behaviour due to modal expectations and enactments of ‘feminine’ and ‘masculine’ activity, with little biological justification for such activities to persist. De Beauvoir (1949/2011) provides a telling example of this, commenting upon the fact that young boys engage in all sorts of “games, sports, fights, challenges, and exploits” (p. 341), whereas for a young girl:

her ears are filled with the treasures of feminine wisdom, feminine virtues are presented to her, she is taught cooking, sewing, and housework as well as how to dress, how to take care of her personal appearance, charm, and modesty[...] posture is imposed on her[...] she has to repress spontaneous movements, she is told not to look like a tomboy, strenuous exercise is banned, she is forbidden to fight; in short, she is committed to becoming, like her elders, a servant and an idol. (ibid., p. 343)

Whilst the dated nature of this extract renders aspects of it relatively extreme, it still beautifully highlights the way in which gendering practices condition bodily subjectivity, such that young girls often reinforce uniquely feminine modalities (and young boys often reinforce uniquely masculine modalities) through their daily activities. It should also be noted that such modalities are *enacted* in the strictest sense; that is, the persistence of such modalities is due to the ongoing modulation of such norms through expected behaviours. This is crucial in that this discussion should not be misconstrued as endorsing the idea that gender is some sort of fixed social construction, but as highlighting the *active* role that individual agents play in generating and modulating social norms – in this case, gendering norms – that they are then modulated by.

Gender, then, provides a canonical example of the biosocial agent, in which one’s lived social-body is both a generator and generation of social norms. For gendered agents, bodily subjectivity is simultaneously social subjectivity, with both bodily maintained social normativity and socially maintained bodily activity being iteratively looped through one another. Importantly, there are innumerable other modes of existence that are entwined with and alongside gender, from engrained cultural dogmas to more subtle passing trends, all of which combine and channel the bodily activity of each biosocial agent into a unique pattern. One such example of our bodies being more than mere flesh is fashion. Every society, from the high couture of London’s Savile Row to the tā moko markings of Māori, promotes the adornment of bodies with specific clothing and/or accessories, all of which are normatively laden. At a more fundamental level we find that largely fixed physical traits are also laced with social normativity. For instance, perceived physical attractiveness will elicit

variations in social assessments and consequent achievements (Umberson and Hughes, 1987), with even less holistic social assessments such as perception of tooth colour producing variations in favourability (Kershaw, Newton and Williams, 2008). Even an isolated physical attribute such as height can alter social assessments of various qualities such as leadership potential (Re, Dzhelyova, Holzleitner, Tigue, Feinberg and Perret, 2012), political popularity (Sorokowski, 2009), workplace success (Judge and Cable, 2004) and anxiety levels (Melamed, 1994). Whilst one could interpret these findings as simply demonstrating correlation between distinct bodily characteristics and social responses, the interpretation that should be favoured in light of earlier discussions is that these bodily characteristics are direct modulators of the socio-normative domain of human experience. Every bodily trait and every bodily movement is loaded with social meaning, potentially regulating the sphere of normative possibilities for oneself and others. Simultaneously, the generation and modulation of social norms is carried through the interconnectedness of subjective bodies.

In this way, humans are inescapably *biosocial*. We are always more than mere physicality or mere sociality, instead being permeated by bodily-social power and purpose from the very earliest moments of life until its end.

## **6. Conclusion**

Across sections 2–4 of this chapter I have outlined Kyselo's (2014) enactive approach to selfhood, along with issues that arise from her theoretical formulation. Whilst I agree with the value in her objective of discerning the relationship between the body and sociality, so as to resolve the problem of how bodily and social processes relate to one another in the individuation of the human self, I do not believe that she satisfactorily offers a resolution. The persistent issue is that by initially considering bodily and social activities as distinct processes, Kyselo is unable to satisfactorily marry the two together in the balanced manner that she seeks. However, in section 5, I have demonstrated that there is a route to unifying bodily and social processes by elucidating the presence of a unique experiential domain for human existence. I refer to this domain, which is propelled by the ongoing iterative reciprocity of bodily generated social norms and socially generated bodily norms, as a *biosocial experiential space*. It is inhabiting this space that makes us what we are. Human selfhood is being a bodily-social modulator of the biosocial experiential space.

# **Conclusion**



### Conclusion

At the outset of this thesis, I highlighted two broad questions that would be playing a motivational role across each chapter. These questions were:

- i. How, and to what extent, is human selfhood constituted by social processes and social structures?
- ii. How should we conceive of human social cognition?

We are now in a position where these questions have been answered. With regards to the first question, I have argued that human selfhood is constituted by social normative structures from its earliest inception through to its cessation. This holds true whether we are considering human selves as individual subjects or as culturally embedded beings that are part of various societal domains. Indeed, our cultures are themselves constituted by social norms, such that there is ongoing reciprocal modulation between ‘individual’ and ‘societal’ orbits. Part of the relevance of this argument is that a proper appreciation of either human individuals or human societies requires viewing the two as inseparably entwined – ‘individual’ and ‘society’ are two sides of the same coin.

Whilst the claim that human selves are socially constituted is not novel to philosophy, I hope that, firstly, I have presented it in new light, as straddling a juncture between phenomenology and cognitive science, and I hope that, secondly, I have presented it more forcibly than is typical. This latter hope stems from the fact that my claim regarding the social constitution of human selfhood is not just that we are *with-others* in a social world, but that we *are*, in part, others. I mean by this that the active presence of others plays a modulatory role in the life of any given human individual, channelling behaviours and cognitive processes in such a way that one’s existence is prodded, pulled and massaged in various normative directions during interactions (across numerous social ‘levels’). On this view, even our bodies are socially codified, allowing us to rise above biological imperatives through the inherent communicative potential of our ‘physical’ movements.

An upshot of this view is that when one considers all those qualities that commonsense tells us make us unique – such as love, reason, or imagination – it is not enough to say that they are social contingencies; rather, it is that, within each of us, these qualities are only manifest through the constitutive modulating activity of others.

With regards to the second question that has motivated much of this thesis, several suggestions have been covered as regards a more satisfactory understanding of human social cognition. Firstly, deficiencies with existing theories of social cognition – even more other-dependent ‘interactive’ theories – have been demonstrated. These deficiencies highlighted the need for theoretical sensitivity to the ‘subjects’ that are involved in social cognition. As such, I have argued for the phenomenological grounding of any theory of social cognition, so

that the enworlded, context-sensitive and constitutively social nature of human existence is not ignored during scientific studies. I also believe that a phenomenological grounding for a theory of social cognition allows one to value the supra-individual dynamics of social engagements, without 'losing' individuals to these dynamics. This delicate navigation is achieved through the notion of a 'we-self', which captures the immersive transformation of a socially cognising individual within a unique self-other domain of normativity. Combining these phenomenological insights, we should conceive of social cognition as underpinned by the anthropically unique transformation of one's experiential world that unfolds when one is immersed in another's co-present activity.

Of course, a phenomenologically grounded theory of social cognition would still have much to answer to. For instance, connections to neural mechanisms, across the diverse spectrum of social engagement, would need to be drawn. There would also be a need to consider the differences involved in dyadic interactions as opposed to interactions within large groups, as well as considering in more detail the connections and differences between social observations and social interactions. Exactly what a phenomenologically grounded science would look like is also a fruitful avenue of investigation. The framework that I have proposed is thus only a preliminary breach of a whole panorama of potential further analyses.

Similarly, the picture of selfhood that I have proposed is by no means complete. To mention just two further avenues of interest: how the self is modified during various interactions and how it develops across the stages of a typical lifetime are both areas of potential analytical profit.

In keeping with the tone of these questions, one could consider a conclusion to this thesis to be as follows.

Within the kinds of phenomenologically sympathetic approaches to cognitive science (and phenomenological theories themselves) towards which this thesis is favourably disposed, there is a tendency to think of human cognition (and selfhood) as overcoming any arbitrary brain-body-world boundaries. Part of the point that I am making is that a proper appreciation of this cognitive cascade is to not merely focus on individuals' 'leaking' into the world, but to focus on how individuals leak into one another and into other social normative structures such as institutions, cultures and family networks. In this regard, this thesis is merely a foundation – a stepping stone – on which further socially grounded analyses of human cognition and selfhood can be made. By no means do I claim to have provided an abundance of answers, but I hope to have made some sort of start – a tentative modulation of thought – towards a properly social understanding of human life.

## Bibliography

- Abell, F., Happé, F. and Frith, U., 2000, 'Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development', *Cognitive Development*, 15(1), 1-16
- Abraham, A., 2015, 'Gender and creativity: an overview of psychological and neuroscientific literature', *Brain Imaging and Behavior*, 10(2), 609-618
- Abraham, A., Thybusch, K., Pieritz, K. and Hermann, C., 2013, 'Gender differences in creative thinking: behavioral and fMRI findings', *Brain Imaging and Behavior*, 8(1), 39-51
- Albahari, M., 2006, *Analytical Buddhism: The Two-Tiered Illusion of Self*, Palgrave Macmillan
- Anisfield, M., 1996, 'Only Tongue Protrusion Modeling Is Matched by Neonates', *Developmental Review*, 16(2), 149-161
- Argyle, M., 1975, *Bodily Communication*, Routledge
- Arkin, R., 1998, *Behaviour-Based Robotics*, MIT Press
- Aristotle, c.340BC/2002, *Nicomachean Ethics*, trans. S. Broadie and C. Rowe (eds.), Oxford University Press
- Athanasopoulos, P., Bylund, E., Montero-Melis, G., Damjanovic, L., Schartner, A., Kibbe, A., Riches, N. and Thierry, G., 2014, 'Two Languages, Two Minds: Flexible Cognitive Processing Driven by Language of Operation', *Psychological Science*, 26(4), 518-526
- Avenanti, A., Sirigu, A. and Aglioti, S., 2010, 'Racial bias reduces empathic sensorimotor resonance with other-race pain', *Current Biology*, 20(11), 1018-1022
- Bach-y-Rita, P., Collins, C., Saunders, F., White, B. and Scadden, L., 1969, 'Vision Substitution by Tactile Image Projection', *Nature*, 221, 963-964
- Baron-Cohen, S., 1995, *Mindblindness: An Essay on Autism and Theory of Mind*, MIT Press
- Battersby, C., 1998, *The Phenomenal Woman: Feminist Metaphysics and the Patterns of Identity*, Polity Press
- Baumeister, R., DeWall, C., Ciarocco, N. and Twenge, J., 2005, 'Social Exclusion Impairs Self-Regulation', *Journal of Personality and Social Psychology*, 88 (4): 589-564
- Bechtel, W. and Abrahamsen, A., 1991, *Connectionism and the Mind: an Introduction to Parallel Processing in Networks*, Basil Blackwell Ltd
- Blakemore, S., Winton, J. and Frith, U., 2004, 'Social Cognitive Neuroscience: where are we heading?', *Trends in Cognitive Sciences*, 8(5), 216-222
- Bloch, M., 2005, *Essays on Cultural Transmission*, Bloomsbury
- Block, N., 1996, 'What is Functionalism?', in Borchert, D. (ed.), *The Encyclopaedia of Philosophy Supplement*, Macmillan
- Bohl, V. and van den Bos, W., 2012, 'Toward an integrative account of social cognition: marrying theory of mind and interactionism to study the interplay of Type 1 and Type 2 processes', *Frontiers in Human Neuroscience*, 6, 274
- Bourassa, K., Knowles, L., Sbarra, D. and O'Connor, M., 2015, 'Absent but not gone: Interdependence in Couples' Quality of Life Persists After a Partner's Death', *Psychological Science*, 27(2), 270-281
- Bourassa, K., Memel, M., Woolverton, C. and Sbarra, D., 2015, 'A dyadic approach to health, cognition, and quality of life in aging adults', *Psychology and Aging*, 30(2), 449-461
- Bratman, M., 1979, 'Practical Reasoning and Weakness of the Will', *Noûs*, 13, 153-171
- Brook, A., 2007, *The Prehistory of Cognitive Science*, Palgrave Macmillan
- Brooks, R., 1991, 'Intelligence Without Representation', *Artificial Intelligence*, 47: 139-159
- Brotman, D., Golden, S. and Wittstein, I., 2007, 'The Cardiovascular Toll of Stress', *Lancet*, 370(9592), 1089-1100
- Brugger, A., Larivière, L., Mumme, D. and Bushnell, E., 2007, 'Doing the right thing: infants' selection of actions to imitate from observed event sequences', *Child Development*, 78(3), 806-824
- Cappuccio, M. and Wheeler, M., 2012, 'Ground-Level Intelligence: Action-Oriented Representation and the Dynamics of the Background', in: Z. Radman (ed.), *Knowing without Thinking: Mind, Action, Cognition and the Phenomenon of the Background*, New Directions in Philosophy and Cognitive Science, Palgrave-Macmillan, 13-36
- Capra, F., 1996, *The Web of Life: a new synthesis of mind and matter*, Anchor
- Carney, D., Hall, J. and Smith LeBeau, L., 2005, 'Beliefs about the nonverbal expression of social power', *Journal of Nonverbal Behaviour*, 29(2), 105-123
- Carruthers, P. and Smith, P. (eds.), 1996, *Theories of Theories of Mind*, Cambridge University Press
- Castelli, F., Frith, C., Happé, F. and Frith, U., 2002, 'Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes', *Brain: a journal of neurology*, 125(8), 1839-1849
- Chalmers, D., 1994, 'A Computational Foundation for the Study of Cognition', *Journal of Cognitive Science*, (2011) 12: 323-357
- Chaminade, T. and Cheng, G., 2009, 'Social cognitive neuroscience and humanoid robotics', *Journal of Physiology*, 103(3-5), 286-295
- Chaminade, T., Rosset, D., Fonseca, D., Nazarin, B., Lutchter, E., Cheng, G. and Deruelle, C., 2012, 'How do we think machines think? An fMRI study of alleged competition with an artificial intelligence', *Frontiers in Human Neuroscience*, 6, 103
- Chartrand, T. and Bargh, J., 1999, 'The Chameleon Effect: the Perception-Behavior Link and Social Interaction', *Journal of Personality and Social Psychology*, 76(6), 893-910
- Churchland, P., 2013, *Touching a Nerve: The Self as Brain*, W.W. Norton & Company
- Clark, A. and Chalmers, D., 1998, 'The Extended Mind', *Analysis*, 58(1), 7-19
- Clark, A., 1989, *Microcognition: Philosophy, Cognitive Science and Parallel Distributed Processing*, MIT Press
- Clark, A., 1997, *Being There: Putting Brain, Body and World Together Again*, MIT Press
- Clark, A., 2001, *Mindware: An Introduction to the Philosophy of Cognitive Science*, Oxford University Press
- Clark, A., 2003, *Natural-Born Cyborgs: Minds, Technologies and the Future of Human Intelligence*, Oxford University Press
- Clark, A., 2007, 'Pressing the Flesh: A Tension in the Study of the Embodied, Embedded Mind?', *Philosophy and Phenomenological Research*, 76 (1): 37-59
- Clark, A., 2010, 'Coupling, constitution and the cognitive kind: A reply to Adams and Aizawa', in Menary, R. (ed.), *The Extended Mind*, MIT Press, 81-99
- Clark, A. and Lutz, R. (eds.), 1992, *Connectionism in Context*, Springer Verlag
- Colombetti, G., 2014, *The Feeling Body: Affective Science Meets the Enactive Mind*, MIT Press

- Combs, D., Powell, C., Schurtz, D. and Smith, R., 2009, 'Politics, *schadenfreude*, and ingroup identification: The sometimes happy thing about a poor economy and death', *Journal of Experimental Social Psychology*, 45, 636-646
- Cooper, R. and Aslin, R., 1990, 'Preference for infant-directed speech in the first month after birth', *Child Development*, 61(5), 1584-1595
- Coultas, J., 2004, 'When in Rome...an evolutionary perspective on conformity', *Group Processes and Intergroup Relations*, 7(4), 317-331
- Crisafi, A. and Gallagher, S., 2010, 'Hegel and the Extended Mind', *AI and Society*, 25(1), 123-129
- Csibra, G. and Gergely, G., 2006, 'Social learning and social cognition: the case for pedagogy', in Munakata, Y. and Johnson, M. (eds.), *Processes of Change in Brain and Cognitive Development*, 21, Oxford University Press, 249-274
- Csibra, G. and Gergely, G., 2009, 'Natural Pedagogy', *Trends in Cognitive Sciences*, 13(4), 148-153
- Csibra, G. and Volein, A., 2008, 'Infants can infer the presence of hidden objects from referential gaze information', *British Journal of Developmental Psychology*, 26, 1-11
- Cuffari, E., Di Paolo, E. and De Jaegher, H., 2014, 'From participatory sense-making to language: there and back again', *Phenomenology and the Cognitive Sciences*, 14(4), 1089-1125
- Currie, G., 2007, 'Narrative Frameworks', in Hutto, D. (ed.), *Narrative and Understanding Persons: Royal Institute of Philosophy Supplement*, 60, Cambridge University Press, 17-42
- Damasio, A., 1994, *Descartes' Error: Emotion, Reason, and the Human Brain*, GP Putnam
- Damasio, A., 2010, *Self Comes to Mind: Constructing the Conscious Brain*, Vintage
- Darwin, C., 1859/1996, *On the Origin of Species*, Oxford University Press
- De Beauvoir, S., 1949/2011, *The Second Sex*, trans. C. Borde and S. Malovany-Chevallier, Vintage
- Defoe, D., 1719, *Robinson Crusoe*, Wordsworth Classics
- De Haan, S., 2010, 'Comment: the minimal self is a social self', in Fuchs, T., Sattel, H. and Henningsen, P. (eds.), *The Embodied Self*, Schattauer, 12-17
- De Jaegher, H., 2013, 'How we affect each other. Michel Henry's 'pathos-with' and the enactive approach to intersubjectivity', *Journal of Consciousness Studies*, 22 (1-2): 112-132
- De Jaegher, H. and Di Paolo, E., 2007, 'Participatory sense-making: an enactive approach to social cognition', *Phenomenology and the Cognitive Sciences*, 6: 485-507
- De Jaegher, H. and Di Paolo, E., 2008, 'Making sense in participation. An enactive approach to social cognition', in Morganti, F., Carassa, A. and Riva, G. (eds.), *Enacting Intersubjectivity: A Cognitive and Social Perspective to the Study of Interactions*, IOS Press
- De Jaegher, H., Di Paolo, E. and Gallagher, S., 2010, 'Can Social Interaction Constitute Social Cognition?' *Trends in Cognitive Sciences*, 14(10), 441-447
- De Jaegher, H., Peräkylä, A. and Stevanovic, M., 2016, 'The co-creation of meaningful action: Bridging enaction and interactional sociology', *Philosophical Transactions of the Royal Society B*, 371: 20150378, 1-10
- Deleuze, G. and Guattari, F., 1980/1987, *A Thousand Plateaus: Capitalism and Schizophrenia*, trans. B. Massumi, University of Minnesota Press
- Dennett, D., 1992, 'The self as a center of narrative gravity', in Kessel, F., Coleand, P. and Johnson, D. (eds.) *Self and Consciousness*, Erlbaum
- Descartes, R., 1641/1996, *Meditations on First Philosophy*, trans. J. Cottingham, Cambridge University Press
- Descartes, R., 1644/1983, *Principles of Philosophy*, trans. V. and R. Miller, Dordrecht, Reidel
- Di Paolo, E. and De Jaegher, H., 2012, 'The interactive brain hypothesis', *Frontiers in Human Neuroscience*, 6 (163)
- Di Paolo, E., De Jaegher, H. and Gallagher, S., 2013, 'One step forwards, two steps back – not the Tango: comment on Gallotti and Frith', *Trends in Cognitive Sciences*, 17 (7): 303-304
- Di Paolo, E., 2005, 'Autopoiesis, adaptivity, teleology, agency', *Phenomenology and the Cognitive Sciences*, 4: 97-125
- Di Paolo, E., 2009, 'Extended Life', *Topoi*, 28, 9-21
- Di Paolo, E., Rohde, M. and De Jaegher, H., 2010, 'Horizons for the enactive mind: Values, social interaction, and play', in Stewart, J., Gapenne, O. and Di Paolo, E. (eds.), *Enaction: Towards a New Paradigm for Cognitive Science*, MIT Press
- Dittrich, W., Troscianko, T., Lea, S. and Morgan, D., 1996, 'Perception of emotion from dynamic point-light displays represented in dance', *Perception*, 25 (6): 727-738
- Douglas, A., 1994, *Symbiotic Interactions*, Oxford University Press
- Drayson, Z., 2010, 'Extended Cognition and the Metaphysics of Mind', *Cognitive Systems Research*, 11(4), 367-377
- Dreyfus, G., 2011, 'Self and Subjectivity: a middle way approach', in Sidertis, M., Thompson, E. and Zahavi, D. (eds.), *Self, No Self? Perspectives from Analytical, Phenomenological, and Indian Traditions*, Oxford University Press, 114-156
- Dreyfus, H., 1991, *Being-in-the-World: a Commentary on Heidegger's Being and Time, Division I*, MIT Press
- Dreyfus, H., 1992, *What Computers Still Can't Do*, MIT Press
- Dreyfus, H., 2000, 'Responses', in Malpas, J. and Wrathall, M. (eds.) *Heidegger, Coping, and Cognitive Science: Essays in Honor of Hubert L. Dreyfus*, MIT Press, 337-344
- Dreyfus, H., 2002, 'Intelligence without Representation – Merleau-Ponty's Critique of Mental Representation', *Phenomenology and the Cognitive Sciences*, 1: 367-383
- Dreyfus, H., 2007a, 'Why Heideggerian AI failed and why fixing it would require making it more Heideggerian', *Artificial Intelligence*, 171 (18): 1137-1160
- Dreyfus, H., 2007b, 'The Return of the Myth of the Mental', *Inquiry*, 50/4: 352-365
- Dreyfus, H., 2007c, 'Response to McDowell', *Inquiry*, 50/4: 371-377
- Dreyfus, H., 2008, *On the Internet (Thinking in Action)*, Routledge
- Dreyfus, H., 2012, 'The Mystery of the Background qua Background', in Radman, Z. (ed.), *Knowing Without Thinking: mind, action, cognition, and the phenomenon of the background*, Palgrave MacMillan, 1-12
- Dreyfus, H. and Kelly, S., 2007, 'Heterophenomenology: Heavy-Handed Sleight-of-Hand', *Phenomenology and the Cognitive Sciences*, 6/1-2: 42-55
- Durkheim, E., 1893/1997, *The Division of Labour in Society*, trans. L. Coser, Free Press
- Engelen, E-M. and Röttger-Rössler, B., 'Current Disciplinary and Interdisciplinary Debates on Empathy', *Emotion Review*, 4(1), 3-8
- Engert, V., Plessow, F., Miller, R., Kirschbaum, C. and Singer, T., 2014, 'Cortisol increase in empathic stress is modulated by emotional closeness and observation modality', *Psychoneuroendocrinology*, 45, 192-201

- Eriksson, K. and Coultas, J., 2009, 'Are people really conformist-biased? An empirical test and a new mathematical model', *Journal of Evolutionary Psychology*, 7(1), 5-21
- Fantasia, V., De Jaegher, H. and Fasulo, A., 2014 'We can work it out: An enactive look at cooperation', *Frontiers in Psychology*, 5: 874-885
- Farmer, H., Maister, L. and Tsakiris, M., 2014, 'Change my body, change my mind: the effects of illusory ownership of an outgroup hand on implicit attitudes toward that outgroup', *Frontiers in Psychology*, 4, 1016
- Farroni, T., Csibra, G., Simion, F. and Johnson, M., 2002, 'Eye Contact Detection in Humans from Birth', *Proceedings of the National Academy of Sciences of the United States of America*, 99/14: 9602-9605
- Farroni, T., Menon, E. and Johnson, M., 2006, 'Factors influencing newborns' preference for faces with eye contact', *Journal of Experimental Child Psychology*, 95, 298-308
- Field, T., Woodson, R., Cohen, D., Greenberg, R., Garcia, R. and Collins, E., 1983, 'Discrimination and imitation of facial expressions by term and preterm neonates', *Infant Behaviour and Development*, 6, 485-489
- Feingold, A. and Mazzella, R., 1998, 'Gender Differences in Body Image are increasing', *Psychological Science*, 9(3), 190-195
- Fleischaker, G., 1992, 'Questions concerning the ontology of autopoiesis and the limits of its utility', *International Journal of General Systems*, 21(2), 131-141
- Frankfurt, H., 1988, *The Importance of What We Care About: Philosophical Essays*, Cambridge University Press
- Fridland, E., forthcoming, 'Longer, Smaller, Faster Stronger: On skill and intelligence', [online] <http://mindsonline.philosophyofbrains.com/wp-content/uploads/2016/09/Fridland-Longer-shorter-faster-stronger-draft-2.pdf>, cited 2016
- Frith, C. and Frith, U., 1999, 'Interacting Minds – a Biological Basis', *Science*, 286(5445): 1692-1695
- Frith, U. and Happé, F., 1994, 'Autism: beyond "theory of mind"', *Cognition*, 50(1-3): 115-132
- Fuchs, T. and De Jaegher, H., 2009, 'Enactive Intersubjectivity: Participatory Sense-Making and Mutual Incorporation', *Phenomenology and the Cognitive Sciences*, 8/4: 465-486
- Gadamer, H-G., 1960/1992, *Truth and Method*, trans. J. Weinsheimer and D. Marshall, Crossroad
- Gallagher, S. and Crisafi, A., 2009, 'Mental Institutions', *Topoi*, 28(1), 45-51
- Gallagher, S., 2000, 'Philosophical conceptions of the self: Implications for Cognitive Science', *Trends in Cognitive Sciences*, 4 (1): 14-21
- Gallagher, S., 2001, 'The practice of mind: Theory, simulation or primary interaction?', *Journal of Consciousness Studies*, 8, 83–108
- Gallagher, S., 2005, *How the Body Shapes the Mind*, Oxford University Press
- Gallagher, S., 2008, 'Direct perception in the intersubjective context', *Consciousness and Cognition*, 17, 535-543
- Gallagher, S., 2009, 'Two problems of intersubjectivity', *Journal of Consciousness Studies*, 16(6-8), 289-308
- Gallagher, S., 2010, 'Joint attention, joint action, and participatory sense making', *Alter: Revue de Phénoménologie*, 18, 111-124
- Gallagher, S., 2012, 'In defense of phenomenological approaches to social cognition: Interacting with the critics', *Review of Philosophy and Psychology*, 3(2), 187-212
- Gallagher, S., 2013a, 'The Socially Extended Mind', *Cognitive Systems Research*, 25, 4-12
- Gallagher, S., 2013b, 'A Pattern Theory of Self', *Frontiers in Human Neuroscience*, 7(443), 1-7
- Gallagher, S. and Meltzoff, A., 1996, 'The earliest sense of self and others: Merleau-Ponty and recent developmental studies', *Philosophical Psychology*, 9, 213-236
- Gallagher, S. and Varga, S., 2014, 'Social constraints on the direct perception of emotions and intentions', *Topoi*, 33(1), 185-199
- Gallagher, S. and Zahavi, D., 2008, *The Phenomenological Mind: an Introduction to Philosophy of Mind and Cognitive Science*, Routledge
- Gallese, V. and Cuccio, V., 2015, 'The Paradigmatic Body: Embodied Simulation, Intersubjectivity, the Bodily Self, and Language', in Metzinger, T. and Windt, J. (eds.), *Open MIND*, [online] doi: 10.15502/9783958570269, cited 2016
- Gallese, V., Fadiga, L., Fogassi, L. and Rizzolatti, G., 1996, 'Action recognition in the premotor cortex', *Brain: A Journal of Neurology*, 593-609
- Gallese, V. and Sinigaglia, C., 2011, 'How the Body in Action Shapes the Self', *Journal of Consciousness Studies*, 18 (7-8): 117-143
- Gallotti, M. and Frith, C., 2013a, 'Social Cognition in the We-mode', *Trends in Cognitive Sciences*, 17: 160-165
- Gallotti, M. and Frith, C., 2013b, 'Response to Di Paolo et al.: How, Exactly, Does It 'Just Happen'? Interaction by Magic', *Trends in Cognitive Sciences*, 17: 304-305
- Ganeri, J., 2012, *The Self: Naturalism, Consciousness, and the First-Person Stance*, Oxford University Press
- Gangopadhyay, N. and Schilbach, L., 2012, 'Seeing minds: A neurophilosophical investigation of the role of perception-action coupling in social perception', *Social Neuroscience*, 7(4), 410-423
- Gergely, G., Bekkering, H. and Kiraly, I., 2002, 'Rational imitation in preverbal infants', *Nature*, 415(6873), 755
- Gergely, G. and Csibra, G., 2014, 'Natural Pedagogy', in Banaji, M. and Gelman, S. (eds.), *Navigating the Social World: What Infants, Children, and Other Species Can Teach Us*, Oxford University Press, 127-133
- Gergely, G., 2010, 'Kinds of agents: The origins of understanding instrumental and communicative agency', in: Goshwami, U. (ed.), *Blackwell Handbook of Childhood Cognitive Development*, Blackwell Publishers, 76-105
- Gibson, J., 1979, *The Ecological Approach to Visual Perception*, Routledge
- Giddens, A., 1976, *New Rules of Sociological Method: a positive critique of interpretive sociologies*, Basic Books
- Goffman, E., 1956, *The Presentation of Self in Everyday Life*, University of Edinburgh
- Goffman, E., 1961, *Encounters: Two Studies in the Sociology of Interaction*, Bobbs-Merrill Company
- Goffman, E., 1964, 'The Neglected Situation', *American Anthropologist*, 66:6, 133-136
- Goh, J. and Park, D., 2009, 'Culture sculpts the perceptual brain', *Progress in Brain Research*, 178, 95-111
- Goldman, A., 1989, 'Interpretation Psychologized', *Mind & Language*, 4: 161-185
- Goldman, A., 2005, 'Imitation, Mind Reading, and Simulation', in Hurley, S. and Chater, N. (eds.), *Perspectives on Imitation: From Neuroscience to Social Science*, 2: 79-93, MIT Press
- Gordon, D., 2010, *Ant Encounters: interaction networks and colony behaviour*, Princeton University Press
- Gosseries, O., Bruno, M., Vanhaudenhuyse, A., Laureys, S. and Schnakers, C., 2009, 'Consciousness in the locked-in syndrome', in Laureys, S. and Tononi, G. (eds.), *The Neurology of Consciousness*, Elsevier, 191–203

- Gould, S. and Lewontin, R., 1979, 'The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme', *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 205(1161), 581-598
- Gross, A., 1968, 'Estimated versus actual physical strength in three ethnic groups', *Child Development*, 39, 283-290
- Grosz, E., 1993, 'Bodies and Knowledge: feminism and the crisis of reason', in Alcoff, L. and Potter, E., *Feminist Epistemologies*, Routledge, 187-215
- Grosz, E., 1994, *Volatile Bodies: Toward a Corporeal Feminism*, Indiana University Press
- Guenther, L., 2013, *Solitary Confinement: Social Death and its Afterlives*, Minnesota University Press
- Hall, J., Coats, E. and Smith LeBeau, L., 2005, 'Nonverbal Behavior and the Vertical Dimension of Social Relations: A Meta-Analysis', *Psychological Bulletin*, 131(6), 898-924
- Harrison, J. and Baron-Cohen, S. (eds.), *Synaesthesia: Classic and Contemporary Readings*, Blackwell Publishing
- Harvard Law Review, 1995, 'Feasibility and Admissibility of Mob Mentality Defenses', *Harvard Law Review*, 108 (5): 1111-1126, The Harvard Law Review Association
- Haugeland, J., 1985, *Artificial Intelligence: the very idea*, MIT Press
- Haugeland, J., 1998, *Having Thought: essays in the metaphysics of mind*, Harvard University Press
- Haugeland, J., 2013, *Dasein Disclosed: John Haugeland's Heidegger*, in Rouse, J. (ed.), Harvard University Press
- Heal, J., 1986, 'Replication and Functionalism', in Butterfield, J. (ed.), *Language, Mind, and Logic*, Cambridge University Press
- Heidegger, M., 1927/1962, *Being and Time*, trans. J Macquarrie and E. Robinson, Basil Blackwell
- Heidegger, M., 1976/2010, *Logic: The Question of Truth*, trans. T Sheehan, Indiana University Press
- Hendin, H., 1982, *Suicide in America*, Norton
- Henry, M., 1963/1973, *The Essence of Manifestation*, trans. G. Etzkorn, Nijhoff
- Henry, M., 1965/1975, *Philosophy and Phenomenology of the Body*, trans. G. Etzkorn, Nijhoff
- Henry, M., 1990/2008, *Material Phenomenology*, trans. S. Davidson, Fordham University Press
- Honneth, A., 2001, 'Invisibility: On the Epistemology of "Recognition"', *Proceedings of the Aristotelian Society, Supplementary Volumes*, 75/1: 111-139
- Hood, B., 2012, *The Self Illusion: Why There is No 'You' Inside Your Head*, Constable
- Hoschild, A., 1983, *The managed heart: commercialization of human feeling*, University of California Press
- Hurley, S., 1998, *Consciousness in Action*, Harvard University Press
- Husserl, E., 1952/1989, *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy*, book 2, *Studies in the Phenomenology of Constitution*, trans. R Rojcewicz and A. Schuwer, Kluwer Academic Publishers
- Husserl, E., 1973, *Husserliana 13-15: Zur Phänomenologie der Intersubjektivität*, I. Kern (ed.), Martinus Nijhoff, cited in Zahavi, D., 2014, *Self and Other: exploring subjectivity, empathy, and shame*, Oxford University Press
- Hutto, D., 2009, 'Interacting? Yes. But, of what kind and on what basis?', *Consciousness and Cognition*, 18, 543-546
- Hutto, D., 2012, 'Knowing without Thinking: the background in philosophy of mind', in Radman, Z. (ed.), *Knowing without Thinking: Mind, Action, Cognition and the Phenomenon of the Background*, New Directions in Philosophy and Cognitive Science, Palgrave-Macmillan, 37-56
- Ikäheimo, H., 2009, 'A vital human need - recognition as inclusion in personhood', *European Journal of Political theory*, 8, 31-45
- Janzen, G., 2007, 'Review of Dan Zahavi's Subjectivity and Selfhood', *Psyche*, 13(1)
- Johnson, M., 1987, *The Body in the Mind: The Bodily Basis Of Meaning, Imagination, And Reason*, University of Chicago Press
- Johnson, S.C., 2000, 'The Recognition of Mentalistic Agents in Infancy', *Trends in Cognitive Sciences*, 4, 22-28
- Johnson, S., 2001, *Emergence: the connected lives of ants, brains, cities, and software*, Scribner
- Jonas, H., 1966/2001, *The Phenomenon of Life: Toward a Philosophical Biology*, University of Chicago Press
- Judge, T. and Cable, D., 2004, 'The Effect of Physical Height on Workplace Success and Income: Preliminary Test of a Theoretical Model', *Journal of Applied Psychology*, 89(3), 428-441
- Jurek, S. and Friedman, S., 2012, *Eat and Run: my unlikely journey to ultramarathon greatness*, Bloomsbury
- Kant, I., 1790/2000, *Critique of the Power of Judgement*, trans. P. Guyer (ed.) and E. Matthews, Cambridge University Press
- Kaplan, D., 2003, *Ricoeur's Critical Theory*, SUNY Press
- Kaspar, K., Jurisch, A. and Schneider, M., 2015, 'Embodied Cognition and Humor: The Impact of Weight Sensations on Humor Experience and the Moderating Role of Gender', *Current Psychology*, 35(3), 377-385
- Katiz, M., Meschulach-Sarfarty, O., Auerbach, J. and Eidelman, A., 1988, 'A re-examination of newborns' ability to imitate facial expressions', *Developmental Psychology*, 24, 3-7
- Kershaw, S., Newton, J. and Williams, D., 2008, 'The influence of tooth colour on the perceptions of personal characteristics among female dental patients: comparisons of unmodified, decayed and 'whitened' teeth', *British Dental Journal*, 204(5), 256-257
- Kierkegaard, S., 1844/1980, *The Concept of Anxiety*, trans. R Thomte, Princeton University Press
- Kierkegaard, S., 1847/2009, *Works of Love*, trans. G. Pattison, HarperCollins
- Kierkegaard, S., 1849/1983, *The Sickness Unto Death*, trans. H. Hong and E. Hong, Princeton University Press
- Kirby, S., Cornish, H. and Smith, K., 2008, 'Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language', *Proceedings of the National Academy of Sciences of the United States of America*, 105(31), 10681-10686
- Kirchoff, M., 2015, 'Extended Cognition & the Causal-Constitutive Fallacy: In Search for a Diachronic and Dynamical Conception of Constitution', *Philosophy and Phenomenological Research*, 90(2), 320-360
- Kitayama, S. and Park, J., 2010, 'Cultural neuroscience of the self: understanding the social grounding of the brain', *Social Cognitive and Affective Neuroscience*, 5(2-3), 111-129
- Kiverstein, J. and Clark, A., 2009, 'Introduction: Mind Embodied, Embedded, Enacted: One Church or Many?', *Topoi*, 28 (1): 1-7
- Kiverstein, J., 2012, 'What is Heideggerian Cognitive Science?', in Kiverstein, J. and Wheeler, M. (eds.), *Heidegger and Cognitive Science*, Palgrave Macmillan, 1-61
- Kiverstein, J. and Wheeler, M., 2012, *Heidegger and Cognitive Science*, Palgrave Macmillan
- Kosfeld, M., Heinrichs, M., Zak, P., Fischbacher, U. and Fehr, E., 2005, 'Oxytocin Increases Trust in Humans', *Nature*, 435 (7042): 673-676

- Koubová, A., 2014, 'Invisible excess of sense in social interaction', *Frontiers in Psychology*, 5, 1081
- Kozlowski, L. and Cutting, J., 1977, 'Recognizing the sex of a walker from a dynamic point-light display', *Perception and Psychophysics*, 21 (6): 575-580
- Krueger, J., 2010, 'The Who and How of Experience', in Sidertis, M., Thompson, E. and Zahavi, D. (eds.), *Self, No Self? Perspectives from Analytical, Phenomenological, and Indian Traditions*, Oxford University Press, 27-55
- Kugiumutzakis, G., 1985, *The Origin, Development and Function of Early Infant Imitation*, Uppsala University
- Kyselo, M. and Di Paolo, E., 2013, 'Locked-in Syndrome: a Challenge for Embodied Cognitive Science', *Phenomenology and the Cognitive Sciences*, 14(3), 517-542
- Kyselo, M., 2014, 'The Body Social: An Enactive Approach to the Self' *Frontiers in Psychology*, 5(986), doi: 10.3389/fpsyg.2014.00986, cited 2016
- Laurie, N., Dwyer, C., Holloway, S. and Smith, F., 1999, *Geographies of New Femininities*, Routledge
- Leach, C., Spears, R., Branscombe, N. and Doosje, B., 2003, 'Malicious pleasure: schadenfreude at the suffering of another group', *Journal of Personality and Social Psychology*, 84(5), 932-943
- Leary, M. and Baumeister, F., 1995, 'The Nature and Function of Self-Esteem: Sociometer Theory', in Zanna, M. (ed.) *Advances in Experimental Social Psychology*, 32, 1-62
- Lieberman, P., 2013, *The Unpredictable Species: What Makes Humans Unique*, Princeton University Press
- Lipps, T., 1909, *Leitfaden der Psychologie*, Verlag von Wilhelm Englemann
- Lovelock, J., 1979, *Gaia: a new look at life on Earth*, Oxford University Press
- Lovelock, J. and Margulis, L., 1974, 'Atmospheric homeostasis by and for the biosphere: the Gaia Hypothesis', *Tellus*, XXVI, 1-10
- Luhmann, N., 1986/1989, *Ecological Communication*, trans. J. Bednarz Jr, University of Chicago Press
- Luhmann, N., 1984/1995, *Social Systems*, trans. J. Bednarz Jr and D. Baecker, Stanford University Press
- Luhmann, N., 2002/2012, *Introduction to Systems Theory*, trans. P. Gilgen, Polity Press
- Lulé, D., Zickler, C., Häcker, S., Bruno, M., Demertzi, A., Pellás, F., Laureys, S. and Kübler, A., 2009, 'Life can be worth living in locked-in syndrome', *Progress in Brain Research*, 177, 339-351
- Lutz, A., Lachaux, J., Martinerie, J. and Varela, F., 2002, 'Guiding the study of brain dynamics by using first-person data: synchrony patterns correlate with ongoing conscious states during a simple visual task', *Proceedings of the National Academy of Sciences of the United States of America*, 99(3), 1586-1591
- Lutz, A. and Thompson, E., 2003, 'Neurophenomenology', *Journal of Consciousness Studies*, 10(9-10), 31-52
- Lynch, J., 1979, *The Broken Heart: the medical consequences of loneliness*, Basic Books
- Malpass, R. and Kravitz, J., 1969, 'Recognition for faces of own and other race', *Journal of Personality and Social Psychology*, 13(4), 330-334
- Margolis, E., Samuels, R., and Stich, S. (eds.), 2012, *The Oxford Handbook of Philosophy of Cognitive Science*, Oxford University Press
- Margulis, L., Sagan, D. and Elredge, N., 1995, *What is Life?*, Simon and Schuster
- Maritain, J., 1947/1985, *The Person and the Common Good*, trans. J. Fitzgerald, University of Notre Dame Press
- Marsh, L. and Onof, C., 2007, 'Introduction to the special issue "Perspectives on Social Cognition"', *Cognitive Systems Research*, 9(1-2), 1-4
- Marx, V. and Nagy, E., 2015, 'Fetal Behavioural Responses to Maternal Voice and Touch', *PLoS ONE*, 10(6), [online] <http://dx.doi.org/10.1371/journal.pone.0129118>, cited 2016
- Matheson, R., *I Am Legend*, Gold Medal Books
- Matsumoto, D., 2006, 'Culture and Nonverbal Behaviour', in Manusov, V. and Patterson, M., *The SAGE Handbook of Nonverbal Communication*, Sage Publications Inc, 219-235
- Matsumoto, D. and Kudoh, T., 1987, 'Cultural Similarities and Differences in the Semantic Dimensions of Body Posture', *Journal of Nonverbal Behaviour*, 11(3), 166-179
- Maturana, H., 1975, 'The Organization of the living: A theory of the living organization', *The International Journal of Man-Machine Studies*, 7, 313-332
- Maturana, H., 1978, 'Biology of language: The epistemology of reality', in Miller, G. and Lenneberg, E. (eds.) *Psychology and biology of language and thought: Essays in honor of Eric Lenneberg*, Academic Press, 27-63
- Maturana, H., 1990, 'Ontology of observing: the biological foundations of self consciousness and the physical domain of existence', in Luhmann, N. (ed.), *Beobachter: Konvergenz der Erkenntnistheorien?*, Wilhelm Fink Verlag, 47-117
- Maturana, H., 1995, 'Biology of self-consciousness' in Tratteur, G. (ed.) *Consciousness: Distinction and reflection*, Bibliopolis, 145-175
- Maturana, H., 2002, 'Autopoiesis, structural coupling and cognition: a history of these and other notions in the biology of cognition', *Cybernetics and Human Knowing*, 9 (3-4): 5-34
- Maturana, H., Mpodozis, J. and Letelier, J., 1995, 'Brain, language and the origin of human mental functions', *Biological Research*, 28(1), 15-26
- Maturana, H. and Varela, F., 1980, *Autopoiesis and Cognition: the Realization of the Living*, 43, D Reidel Publishing Co
- Maturana, H. and Varela, F., 1987, *The Tree of Knowledge: the biological roots of human understanding*, Shambala
- McDowell, J., 1994, *Mind and World*, Harvard University Press
- McGann, M. and De Jaegher, H., 2009, 'Self-other contingencies: enacting social perception', *Phenomenology and the Cognitive Sciences*, 8(4), 417-437
- Melamed, T., 1994, 'Correlates of physical features: Some gender differences', *Personality and Individual Differences*, 17, 689-691
- Meltzoff, A. and Brooks, R., 2001, 'Like me" as a building block for understanding other minds: Bodily acts, attention, and intention', in Malle, B., Moses, L. and Baldwin, D. (eds.), *Intentions and Intentionality: Foundations of Social Cognition*, MIT Press 171-191
- Meltzoff, A., 1995, 'Understanding the Intentions of Others: re-enactment of intended acts by 18-month-old children', *Developmental Psychology*, 31, 838-850
- Meltzoff, A. and Moore, M., 1977, 'Imitation of Facial and Manual Gestures by Human Neonates', *Science*, 198, 75-78
- Meltzoff, A. and Moore, M., 1983, 'Newborn Infants Imitate Adult Facial Gestures', *Child Development*, 54, 702-709
- Meltzoff, A. and Moore, M., 1995, 'A Theory of the Role of Imitation in the Emergence of the Self', in Rochat, P. (ed.), *The Self in Infancy: Theory and Research*, Elsevier, 73-93
- Menary, R., 2007, *Cognitive Integration: mind and cognition unbounded*, Palgrave Macmillan
- Menary, R., 2010, 'Introduction to the special issue on 4E cognition', *Phenomenology and the Cognitive Sciences*, 9(4), 459-463

- Merleau-Ponty, M., 1945/2012, *Phenomenology of Perception*, trans. D. Landes, Routledge
- Merleau-Ponty, M., 1964, *The Primacy of Perception and Other Essays on Phenomenological Psychology, the Philosophy of Art, History and Politics*, Eadie, J. (ed.), Northwestern University Press
- Mingers, J., 1994, *Self-Producing Systems: Implications and Applications of Autopoiesis (Contemporary Systems Thinking)*, Springer
- Molenberghs, P., Halász, V., Mattingley, J., Vanman E. and Cunnington, R., 2012, 'Seeing is believing: Neural mechanisms of action–perception are biased by team membership', *Human Brain Mapping*, 34(9), 2055-2068
- Moriguchi, Y., Touroutoglou, A., Dickerson, B. and Feldman Barrett, L., 2013, 'Sex differences in the neural correlates of affective experience', *Social Cognitive and Affective Neuroscience*, 9(5), 591-600
- Nagy, E., Kompagne, H., Orvos, H., Pal, A., Molnar, P., Jansky, I., Loveland, C. and Bardos, G., 2005, 'Index finger movement imitation by human neonates: motivation, learning, and left-hand preference', *Pediatric Research*, 58(4), 749-753
- Nagy, E., Pilling, K., Orvos, H. and Molnar, P., 2013, 'Imitation of tongue protrusion in human neonates: specificity of the response in a large sample', *Developmental Psychology*, 49(9), 1628-1638
- Neisser, U., 1993, 'The Self Perceived', in Neisser, U. (ed.), *The Perceived Self: Ecological and Interpersonal Sources of Self Knowledge*, Cambridge University Press, 3-24
- Nichols, S. and Stich, S., 2003, *Mindreading: an Integrated Account of Pretence, Self-Awareness, and Understanding of Other Minds*, Oxford University Press
- Noë, A., 2004, *Action in Perception*, MIT Press
- Noë, A., 2009, *Out of Our Heads: Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness*, Hill & Wang
- Nielsen, M., 2006, 'Copying actions and copying outcomes: social learning through the second year', *Developmental Psychology*, 42(3), 555-565
- Nietzsche, F., 1881/1997, *Daybreak: Thoughts on the Prejudices of Morality*, Clark, M. and Leiter, B. (eds.), trans. R. Hollingdale, Cambridge University Press
- Northoff, G. and Bermpohl, F., 2004, 'Cortical midline structures and the self', *Trends in the Cognitive Sciences*, 8, 102–107
- O'Regan, J. and Noë, A., 2001, 'A Sensorimotor Account of Vision and Visual Consciousness', *Behavioural and Brain Sciences*, 24, 939-1031
- Overgaard, S. and Michael, J., 2013, 'The interactive turn in social cognition research: A critique', *Journal of Philosophical Psychology*, 28(2), 160-183
- Pagel, M., 2012, *Wired for Culture: Origins of the Human Social Mind*, WW Norton & Company
- Palermos, S., 2014, 'Loops, Constitution, and Cognitive Extension', *Cognitive Systems Research*, 27, 25-41
- Park, J. and Kitayama, S., 2010, 'Interdependent selves show face-induced facilitation of error processing: cultural neuroscience of self-threat', *Social Cognitive and Affective Neuroscience*, 9(2), 201-208
- Parnas, J. and Sass, L., 2011, 'The Structure of Self-Consciousness in Schizophrenia', in Gallagher, S. (ed.), *The Oxford Handbook of the Self*, Oxford University Press, 521-546
- Petitot, J., Varela, F., Pachoud, B. and Roy, J-M. (eds.), 1999, *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science*, Stanford University Press
- Phillips, W., Baron-Cohen, S. and Rutter, M., 1992, 'The role of eye contact in goal detection: Evidence from normal infants and children with autism or mental handicap', *Development and Psychopathology*, 4(3), 375-383
- Polanyi, M., 1964, *The Study of Man*, University of Chicago Press
- Popova, Y., 2014, 'Narrativity and enaction: the social nature of literary narrative', *Frontiers in Psychology*, 5(895), 1-14
- Protevi, J., 2009, *Political Affect: Connecting the Social and the Somatic*, University of Minnesota Press
- Przyrembel, M., Smallwood, J., Pauen, M. and Singer, T., 2012, 'Illuminating the dark matter of social neuroscience: Considering the problem of social interaction from philosophical, psychological, and neuroscientific perspectives', *Frontiers in Human Neuroscience*, 6(190), 1-15
- Rakova, M., 2006, *Philosophy of Mind A-Z*, Edinburgh University Press
- Rameson, L., Morelli, S. and Lieberman, M., 2012, 'The neural correlates of empathy: experience, automaticity, and prosocial behavior', *Journal of Cognitive Neuroscience*, 24(1), 235-245
- Rasmussen, J. (ed.), 1973, *Man in isolation and confinement*, Transaction Publishers
- Ratcliffe, M., 2002, 'Heidegger's attunement and the neuropsychology of emotion', *Phenomenology and the Cognitive Sciences*, 1, 287-312
- Ratcliffe, M., 2004, 'Realism, biologism and 'the background'', *Philosophical Explorations*, 7(2), 149-166
- Ratcliffe, M., 2009, 'Existential Feeling and Psychopathology', *Philosophy, Psychiatry and Psychology* 16/2: 179-211
- Ratcliffe, M., 2013a, 'Why Mood Matters', in Wrathall, M. (ed.), *Cambridge Companion to Heidegger's Being and Time*, Cambridge University Press, 157-176
- Ratcliffe, M., 2013b, 'The Structure of Interpersonal Experience', in Moran, D. and Thybo Jensen, R., (eds.), *The Phenomenology of Embodied Subjectivity*, Springer, 221-238
- Ratcliffe, M., 2014, *Experiences of Depression: a study in phenomenology*, Oxford University Press
- Ravenscroft, I., 1997/2016, *Folk Psychology as a Theory*, [online] <http://plato.stanford.edu/entries/folkpsych-theory/#TheoTheo> cited 2016
- Re, D., Dzhelyova, M., Holzleitner, I., Tigue, C., Feinberg, D. and Perret, D., 2012, 'Apparent height and body mass index influence perceived leadership ability in three-dimensional faces', *Perception*, 41(12), 1477-1485
- Rietveld, E. and Kiverstein, J., 2014, 'A Rich Landscape of Affordances', *Ecological Psychology*, 26 (4): 325-352
- Rietveld, E., 2008a, 'The skillful body as a concerned system of possible actions: Phenomena and neurodynamics', *Theory and Psychology*, 18 (3): 341-363
- Rietveld, E., 2008b, 'Situated normativity: The normative aspect of embodied cognition in unreflective action', *Mind*, 117 (468): 973-1001
- Rietveld, E., 2010, 'McDowell and Dreyfus on unreflective action', *Inquiry*, 53 (2), 183-207
- Rietveld, E., 2012a, 'Bodily intentionality and social affordances in context', in Paglieri, F. (ed.) *Consciousness in Interaction: the Role of the Natural and Social Context in Shaping Consciousness*, J. Benjamins, 207-226
- Rietveld, E., 2012b, 'Context-switching and responsiveness to real relevance', in Kiverstein, J. & Wheeler, M. (eds.), *Heidegger and Cognitive Science: New Directions in Cognitive Science and Philosophy*, Palgrave Macmillan, 105-135
- Rietveld, E., 2013, 'Affordances and unreflective freedom', in Moran, D. and Thybo Jensen, R., (eds.), *The Phenomenology of Embodied Subjectivity*, Springer, 21-42

- Rizzolatti, G., Fadiga, L., Fogassi, L. and Gallese, V., 1999, 'Resonance behaviors and mirror neurons', *Archives Italiennes de Biologie*, 137(2-3), 85-100
- Roth, S., 2010, *Shared Agency*, [online] <http://plato.stanford.edu/entries/shared-agency/> cited 2016
- Rowlands, M., 2010, *The New Science of the Mind: from Extended Mind to Embodied Phenomenology*, MIT Press
- Rupert, R., 2009, *Cognitive Systems and the Extended Mind*, Oxford University Press
- Samson, D., Apperly, I., Braithwaite, J., Andrews, B. and Bodley Scott, S., 2010, 'Seeing it their way: evidence for rapid and involuntary computation of what other people see', *Journal of Experimental Psychology*, 36 (5): 1255-1266
- Sartre, J., 1943/1984, *Being and Nothingness: an Essay on Phenomenological Ontology*, trans. H. Barnes, Washington Square Press
- Scheler, M., 1913/2008, *The Nature of Sympathy*, Transaction
- Schilbach, L., 2014, 'On the relationship of *online* and *offline* social cognition', *Frontiers in Human Neuroscience*, 8, 278
- Scholl, B. and Tremoulet, P., 2000, 'Perceptual Causality and Animacy', *Trends in Cognitive Science*, 4(8): 299-309
- Schütz, A., 1932/1967, *The Phenomenology of the Social World*, trans. Walsh, G. and Lehnert, F., Northwestern University Press
- Seidl, D., 2004, 'Luhmann's theory of autopoietic social systems', [online] [http://www.zfog.bwl.uni-muenchen.de/files/mitarbeiter/paper2004\\_2.pdf](http://www.zfog.bwl.uni-muenchen.de/files/mitarbeiter/paper2004_2.pdf) cited 2016
- Seigel, J., 2005, *The Idea of the Self: Thought and Experience in Western Europe since the Seventeenth Century*, Cambridge University Press
- Senju, A. and Csibra, G., 2008, 'Gaze following in human infants depends on communicative signals', *Current Biology*, 18, 668-671
- Senju, A., Csibra, G. and Johnson, M., 2008, 'Understanding the referential nature of looking: Infants' preference for object-directed gaze', *Cognition*, 108(2), 303-319
- Shapiro, L., 2008, 'Functionalism and the Boundaries of the Mind', *Cognitive Systems Research*, 9, 5-14
- Shapiro, L. (ed.), 2014, *The Routledge Handbook of Embodied Cognition*, Routledge
- Sheets-Johnstone, M., 1999, *The Primacy of Movement*, John Benjamins Publishing
- Short, W., forthcoming, 'Embodied, extended, and distributed cognition in Roman technical practice', in Anderson, M., Cairns, D. and Sprevak, M. (eds.), *A History of Distributed Cognition Volume 1: from Early Greece to Late Antiquity*, Publisher under discussion
- Smith, E., 2002, *Playing Hard Ball*, Little, Brown
- Smith, E., 2004, *On and Off the Field*, Viking Press
- Smith, K. and Kirby, S., 2008, 'Cultural evolution: implications for understanding the human language faculty and its evolution', *Philosophical Transactions of the Royal Society B*, 363(1509), 3591-3603
- Sorokowski, P., 2009, 'Politicians' estimated height as an indicator of their popularity', *European Journal of Social Psychology*, 40(7), 1302-1309
- Soussignan, R., Courtial, A., Canet, P., Danon-Apter, G. and Nadel, J., 2011, 'Human newborns match tongue protrusion of disembodied human and robotic mouths', *Developmental Science*, 14(2), 385-394
- Sparrow, B., Liu, J. and Wegner, D., 2011, 'Google effects on memory: Cognitive consequences of having information at our fingertips', *Science*, 333, 776-778
- Sporer, S., 2001, 'Recognizing Faces of Other Ethnic Groups: An Integration of Theories', *Psychology, Public Policy and Law*, 7(1), 36-97
- Sprevak, M., 2009, 'Extended Cognition and Functionalism', *Journal of Philosophy*, 106(9), 503-527
- Stapleton, M., 2013, 'Steps to a "Properly Embodied" Cognitive Science', *Cognitive Systems Research*, 22, 1-11
- Stein, E., 1917/2008, *On the Problem of Empathy*, trans. W. Stein, Institute of Carmelite Studies
- Steiner, E. and Stewart, J., 2009, 'From autonomy to heteronomy (and back): the enaction of social life', *Phenomenology and the Cognitive Sciences*, 8(4), 527-550
- Stich, S. and Nichols, S., 1992, 'Folk Psychology: Simulation or Tacit Theory', *Mind & Language*, 7 (1-2): 35-71
- Stillman, T., Baumeister, R., Lambert, N., Crescioni, A., DeWall, C. and Fincham, F., 2009, 'Alone and Without Purpose: Life Loses Meaning Following Social Exclusion', *Journal of Experimental Social Psychology*, 45(4), 686-694
- Straus, E., 1952, 'The Upright Posture', *Psychiatric Quarterly*, 26(1), 529-561
- Sullivan, P. and Rickers, K., 2013, 'The effect of behavioral synchrony in groups of teammates and strangers', *International Journal of Sport and Exercise Psychology*, 3, 286-291
- Sutton, J., McIlwain, D., Christensen, W. and Geeves, A., 2011, 'Applying intelligence to the reflexes: Embodied skills and habits between Dreyfus and Descartes', *Journal of the British Society for Phenomenology*, 42 (1): 78-103
- Sutton, J., 2007, 'Batting, Habit, and Memory: the embodied mind and the nature of skill', *Sport in Society*, 10 (5): 767-786
- Tarr, B., Launay, J. and Dunbar, R., 2014, 'Music and social bonding: "self-other" merging and neurohormonal mechanisms', *Frontiers in Psychology*, 5, 1096
- Thelen, E., 2004, 'Motor development as foundation and future of developmental psychology', *International Journal of Behavioural Development*, 24, 385-397
- Theiner, G., Allen, C. and Goldstone, R., 2010, 'Recognizing Group Cognition', *Cognitive Systems Research*, 11 (4):378-395
- Theodoridou, A., Rowe, A., Penton-Voak, I. and Rogers, P., 2009, 'Oxytocin and social perception: oxytocin increases perceived facial trustworthiness and attractiveness', *Hormones and Behavior*, 56 (1): 128-132
- Thompson, E. and Stapleton, S., 2009, 'Making sense of sense-making: Reflections on enactive and extended mind theories', *Topoi*, 28 (1): 23-30
- Thompson, E., 2007, *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*, Harvard University Press
- Torrance, S. and Froese, T., 2011, 'An inter-enactive approach to agency: Participatory sense-making, dynamics, and sociality', *Humana Mente*, 15, 21-53
- Tylén, K., Allen, M., Hunter, B. and Roepstorff, A., 2012, 'Interaction vs. observation: distinctive modes of social cognition in human brain and behavior? A combined fMRI and eye-tracking study', *Frontiers in Human Neuroscience*, 6, 331
- Umberson, T. and Hughes, M., 1987, 'The Impact of Physical Attractiveness on Achievement and Psychological Well-Being', *Social Psychology Quarterly*, 50(3), 227-236
- Valentine, G., 2001, *Social Geographies: Society and Space*, Longman
- Van Inwagen, P., 1983, *An Essay on Free Will*, Oxford University Press
- Varela, F., Thompson, E. and Rosch, E., 1991, *The Embodied Mind: Cognitive Science and Human Experience*, MIT Press

- Varela, F., 1979, *Principles of Biological Autonomy*, Elsevier North Holland
- Varela, F., 1996, 'Neurophenomenology: A methodological remedy for the hard problem', *Journal of Consciousness Studies*, 3, 330–349
- Varela, F., 1997, 'Patterns of Life: Intertwining identity and cognition', *Brain Cognition*, 34, 72-87
- Villalobos, M. and Ward, D., 2016, 'Lived Experience and Cognitive Science Reappraising Enactivism's Jonsonian Turn', *Constructivist Foundations*, 11(2), 204-212
- Walter, H., 2012, 'Social Cognitive Neuroscience of Empathy: Concepts, Circuits, and Genes', *Emotion Review*, 4(1), 9-17
- Walter, S., 2010, 'Cognitive extension: the parity argument, functionalism, and the mark of the cognitive', *Synthese*, 177(2), 285-300
- Ward, D. and Stapleton, M., 2012, 'Es are good. Cognition as enacted, embodied, embedded, affective and extended', in Paglieri, F. (ed.), *Consciousness in Interaction: The role of the natural and social context in shaping consciousness*, John Benjamins Publishing
- Wheeler, M. and Di Paolo, E., 2011, 'Existentialism and cognitive science', in Reynolds, J., Woodward, A. and Joseph, F. (eds.), *The Continuum Companion to Existentialism*, Continuum, 241-259
- Wheeler, M., 1995, 'Escaping from the Cartesian mind-set: Heidegger and artificial life', in Moran, F., Moreno, A., Merelo, J. and Chacon, P. (eds.), *Advances in Artificial Life: Proceedings of the Third European Conference on Artificial Life*, Springer-Verlag, 65-76
- Wheeler, M., 1996, 'From Robots to Rothko: the Bringing forth of Worlds', in Boden, M. (ed.), *The Philosophy of Artificial Life*, Oxford University Press
- Wheeler, M., 2005, *Reconstructing the Cognitive World*, MIT Press
- Wheeler, M., 2010a, 'Minds, Things and Materiality', in Malafouris, L. and Renfrew, C. (eds.), *The Cognitive Life of Things: Recasting the Boundaries of the Mind*, McDonald Institute for Archaeological Research Publications, 29-37
- Wheeler, M., 2010b, 'In Defense of Extended Functionalism', in Menary, R. (ed.), *The Extended Mind*, MIT Press, 245-270
- Wheeler, M., 2011c, 'Embodied Cognition and the Extended Mind', in Garvey, J. (ed.), *The Continuum Companion to Philosophy of Mind*, Continuum
- Wheeler, M., 2011a, *Martin Heidegger*, [online] <http://plato.stanford.edu/entries/heidegger/> cited 2016
- Wheeler, M., 2011b, 'Mind in Life or Life in Mind? Making Sense of Deep Continuity', *Journal of Consciousness Studies*, 72, 148-167
- Wheeler, M., 2013, 'Science Friction: Phenomenology, Naturalism and Cognitive Science', *Royal Institute of Philosophy Supplement*, 72, 135-167
- Wiesenfeld-Hallin, Z., 2005, 'Sex differences in pain perception', *Gender Medicine*, 2(3), 137-145
- Williams, K., 2007, 'Ostracism: the kiss of social death', *Social and Personality Psychology Compass*, 1, 236–247
- Wilson, R., 2004, *Boundaries of the Mind: The Individual in the Fragile Sciences*, Cambridge University Press
- Wilson, R., 2005, 'Collective Memory, Group Minds, & the Extended Mind Thesis', *Cognitive Processing*, 6: 227-236
- Wojtyła, K., 1969/1979, *The Acting Person*, trans. A Potocki, Reidel Publishing Company
- Wolf, S., 1990, *Freedom Within Reason*, Oxford University Press
- Wrathall, M., 2000, 'Background practices, capacities, and Heideggerian disclosure', in Malpas, J. and Wrathall, M. (eds.) *Heidegger, Coping, and Cognitive Science: Essays in Honor of Hubert L. Dreyfus*, MIT Press, 93-114
- Xu, X., Zuo, X., Wang, X. and Han, S., 2009, 'Do You Feel My Pain? Racial Group Membership Modulates Empathic Neural Responses', *Journal of Neuroscience*, 29(26), 8525-8529
- Young, I., 1980, 'Throwing Like a Girl: a Phenomenology of Feminine Body Comportment Motility and Spatiality', *Human Studies*, 3(2), 137-156
- Zahavi, D., 1999, *Self-Awareness and Alterity: a Phenomenological Investigation*, North-western University Press
- Zahavi, D., 2001, 'Beyond Enpathy: Phenomenological Approaches to Intersubjectivity', *Journal of Consciousness Studies*, 8/5-7: 151-167 North-western University Press
- Zahavi, D., 2005, *Subjectivity and Selfhood: Investigating the First-Person Perspective*, MIT Press
- Zahavi, D., 2014, *Self and Other: exploring subjectivity, empathy, and shame*, Oxford University Press

