HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

"IT AIN'T ALL IN THE HEAD!"

Situating Cognition to the Body and the Surrounding World

Lauri Guy Vilhelm Calonius
013867722
11 April 2013
Master's Thesis
Supervisors: Tarja Knuuttila and Sami Paavola
Theoretical Philosophy
Department of Philosophy, History, Culture and Art Studies
Faculty of Arts
University of Helsinki

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Tiivistelmä

Tiedekunta/Osasto Fakultet/Sektion – Faculty	Laitos/Institution– Department
Humanistinen tiedekunta	**Filosofian, historian, kulttuurin ja taiteiden tutkimuksen laitos**

Tekijä/Författare – Author

Lauri Guy Vilhelm Calonius

Työn nimi / Arbetets titel – Title

”IT AIN'T ALL IN THE HEAD!” – Situating Cognition to the Body and the Surrounding World

Oppiaine /Läroämne – Subject

Teoreettinenfilosofia

Työn laji/Arbetets art – Level	Aika/Datum – Month and year	Sivumäärä/ Sidoantal – Number of pages
Pro gradu	**Huhtikuu 2013**	79

Tiivistelmä/Referat – Abstract

Tutkielmassa tarkastellaan suuntausta, joka eroaa perinteisestä käsityksestä mielestä ja kognitiosta. Perinteisen käsityksen mukaan mieli ja kognitio ovat yksinomaan selitettävissä päänsisäisinä prosesseina. Tämän tutkimuksen tarkoituksena on erityisesti tarkastella neljää uudempaa kognition teoriaa. Käsiteltävät teoriat ovat kehollinen ja ympäristöön upotettu kognitio, enaktiivinen kognitio, laajennettu kognitio ja hajautettu kognitio. Neljä tutkimuskysymystä ohjaavat tutkielmaa: I) Mihin teoreettisiin käsityksiin nämä neljä eri teoriat ovat sitoutuneet ja miten ne suhtautuvat toisiinsa? II) Miksi hajautetun kognition teoria tulisi liittää muiden teorioiden joukkoon? III) Mitä kritiikkiä nämä teoriat kohtaavat ja miten ns. kognition tunnusmerkin hakeminen vaikuttaa niin teorioiden puolustajiin kuin vastustajiinkin? IV) Miten kognitiivisen agenttiuden käsite näkyy näissä eri teorioissa?

Tutkielma on metodiltaan lähdekirjallisuuden vertaileva ja kriittinen tutkimus. Lähdekirjallisuus on laaja-alainen ja kattava. Kirjallisuus käsittää muun muassa filosofian, kognitiotieteen, psykologian sekä tekoälytutkimuksen keskeisiä tutkimuksia. Pääasialliset henkilöt, joiden tutkimuksia tarkastellaan ovat: Andy Clark, Francisco Varela, Ezequiel Di Paolo, Alva Noë, Edwin Hutchins, John Sutton, Richard Menary, Fred Adams, Kenneth Aizawa, ja Robert Rupert. Edellä mainittujen lisäksi tutkimuksessa analysoidaan muita käsiteltävien teorioiden perusteita.

Tutkimus koostuu kahdeksasta luvusta. Johdantoluku määrittää tutkimuksen teoreettisen taustan, esittelee tutkimuksen rakenteen ja asettaa tutkimuksen tavoitteet. Luvuissa 2–5 käsitellään kutakin neljästä kognition teoriasta. Luku 6 taas keskittyy näiden teorioiden kriittiseen analyysiin sekä niille esitettyjen haasteiden ja vasta-argumenttien tarkasteluun. Luku 7 käsittelee sitä, kuinka näissä eri teorioissa hahmotetaan kognitiivisen agenttiuden käsite. Luku 8 on tutkielman yhteenveto ja siinä vastataan kootusti tutkimuskysymyksiin. Tämä luku myös viittaa mahdollisiin jatkotutkimuskysymyksiin työssä heränneiden avointen kysymysten kautta.

Tutkielma havainnollistaa uudempien kognition teorioiden yhtäläisyyksiä ja eroja, jotka nousevat esiin avaamalla näiden teoreettisia taustaoletuksia. Työ osoittaa kuinka perinteinen näkökulma kognitioon joutuu ottamaan osaa keskusteluun kognition luonteesta yhdenvertaiselta perustalta uudempien teorioiden kanssa. Johtopäätöksenä todetaan, että näiden uudempien teorioiden parempi ymmärtäminen on tärkeä askel kognition luonteen ymmärtämisessä.

Avainsanat – Nyckelord – Keywords

mieli, kognitio, kehollisuus, upotettu kognitio, enaktivismi, laajennettu mieli, laajennettu kognitio, hajautettu kognitio, filosofia, mielen filosofia, agenttius, kognitiivinen ekologia, umwelt, representaatio,

Säilytyspaikka – Förvaringställe – Where deposited

Keskustakampuksen kirjasto

Muita tietoja – Övriga uppgifter – Additional information

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Abstract

Tiedekunta/Osasto Fakultet/Sektion – Faculty	Laitos/Institution– Department
Faculty of Arts	**Department of Philosophy, History, Culture and Art Studies**

Tekijä/Författare – Author

Lauri Guy Vilhelm Calonius

Työn nimi / Arbetets titel – Title

"IT AIN'T ALL IN THE HEAD!" – Situating Cognition to the Body and the Surrounding World

Oppiaine /Läroämne – Subject

Theoretical Philosophy

Työn laji/Arbetets art – Level	Aika/Datum – Month and year	Sivumäärä/ Sidoantal – Number of pages
Master's Thesis	**April 2013**	79

Tiivistelmä/Referat – Abstract

The thesis examines positions that do not restrict cognition and the mind solely to the brain and its neuronal processes. The examination is framed in terms of what it means to say that cognition is embodied, embedded, enactive, extended and distributed (4ED). Four research questions guide this exploration: I) What are the theoretical commitments the different accounts of cognition in 4ED hold on to and how do these relate to each other? II) (Why) should distributed cognition be added with the rest of the 4E accounts? III) What critique does 4ED face and how does providing a mark of the cognitive affect both the proponents as well as the opponents of 4ED? IV) How does the notion of cognitive agency figure in the 4ED approach(es) to cognition?

The methodology of the thesis is a comparative and critical study of the source literature. The literature is wide-ranging and substantive. It expands the fields of philosophy, cognitive science, psychology and artificial intelligence among others. The main authors under study include Andy Clark, Francisco Varela, Ezequiel Di Paolo, Alva Noë, Edwin Hutchins, David Kirsh, John Sutton, Richard Menary, Michael Wheeler, Fred Adams, Kenneth Aizawa, ja Robert Rupert. In addition to the aforementioned authors other sources relevant for the development and refinement of 4ED are examined.

The examination is systematic and runs through eight chapters. Chapter one is the introduction that sets out the theoretical background, structure and the aim of the thesis. Chapters 2–5 are dedicated to the presentation of the different approaches to cognition under study. Chapter six is dedicated to the critical analysis of the approaches and the examination of some of the challenges that have been raised against the different approaches to cognition. Chapter seven looks into a more detailed question and aims to clarify the way the approaches conceive cognitive agency. Chapter eight is the concluding chapter that ties the thesis together, clarifies the main issues and gestures at possible future research in the form of open questions that have risen.

The thesis uncovers the theoretical commitments the different unorthodox approaches to cognition have. In doing this it also illustrates their important similarities and differences. The study also shows how the orthodox conception of cognition that conceives it essentially as a neuronal process bounded by the head is pushed to take part in the debate on equal ground with the unorthodox positions. In general then the thesis points to the importance of a better understanding of the unorthodox approaches to cognition as a means to a greater overall understanding of the nature of cognition and its place in the world.

Avainsanat – Nyckelord – Keywords

embodiment, embeddedness, enaction, enactivism, extended cognition, extended mind, distributed cognition, philosophy, philosophy of mind, agency, scaffolding, cognitive ecology, umwelt, niche, representation, coupling-constitution fallacy, motley crew argument, mark of the cognitive, nonderived content.

Säilytyspaikka – Förvaringställe – Where deposited

City Centre Campus: Main Library

Muita tietoja – Övriga uppgifter – Additional information

Acknowledgements

First and foremost I would like to express my sincerest gratitude to my mother, father and my beloved fiancé whose unfaltering support and encouragement during my studies and the completion of the thesis was an indispensable source of strength.

My deepest appreciation goes to my supervisors Tarja Knuuttila and Sami Paavola whose genuine interest in my work, their insightful comments, fair and precise critique, and timely assistance helped structure the thesis and continuously motivate my research. I owe an immense debt of gratitude to them and could not have hoped for better thesis supervisors.

I would like to acknowledge the Academy of Philosophy Ltd. (Filosofian Akatemia Oy) for the financial support they have offered me. Yet, more importantly I would like to thank the people in the Academy, especially Lauri Järvilehto, Timo Tiuraniemi and Anne Salminen, for the excited exchange of ideas that has been a true source of inspiration. They have helped guide me into new avenues of discovery and ways of conceiving matters in the thesis. I am very grateful for the Academy for providing me with a truly inspirational research environment.

Last but not least I would like to extended a sincere thank you to my examiners professor Gabriel Sandu and professor Matti Sintonen for taking the time out of their busy schedule to read, comment and assess the thesis.

Table of Contents

1 Introduction

Thinking, or knowledge getting, is far from being the armchair thing it is often supposed to be. The reason it is not an armchair thing is that it is not an event going on exclusively within the cortex [...] Hands and feet, apparatus and appliances of all kinds are as much a part of it as changes within the brain.[1]

Traditionally in the modern western thought the mind has been conceived as *isolated* and detached from the body and the rest of the world. Descartes[2] for instance put forward a metaphysical dichotomy between the mind and the body to the conclusion that the mind is something ontologically distinct from physical substance. Specifically, mind is concerned with thinking and other such mental phenomena, but is not divisible and extendable in space. Descartes also advanced the view that the brain (the pineal gland to be specific) was the location where the mind causally interacted with the body so that mindful action would be possible.[3] This interaction was conceived to be quite rudimentary, as Michael Wheeler puts it: '[there was a] bidirectional channel of causation that ran from body to mind in perception, and from mind to body in action.'[4] This metaphysical dichotomy has been dubbed *Cartesian substance dualism.*

Now, even though there has been divergence from this sort of Cartesianism and everything is for instance argued to be of one physical substance, i.e. mental phenomena are reducible or identical to, realized or determined by, or supervene on some physical processes, for the most part the notion of an isolated mind has persisted both in philosophy and in the multidisciplinary field of cognitive science.[5] One clear example of the sort of isolationism present in the early work of cognitive science is a computer analogy of cognition, where the mind is seen as the *software* run by the *hardware* of the brain. In this picture cognition is characterised as a process of symbol-manipulation that relies on the syntactic (rule-like) and semantic (meaningful) properties of the symbols. This symbol-manipulation is an input-output process, where the symbolic inputs received through the senses get translated in the mind/brain to other symbols that result as an output, e.g. overt physical behaviour. Hence, because the focus is purely on the internal manipulation of the symbolic inputs and outputs, cognition is essentially seen as an *isolated* process separated from the world.[6] Susan Hurley aptly characterises this sort of view as a *sandwich model of cognition.* Perception and action are separated from each other and the higher cognitive processes; they are the buns and cognition is the filling in between.[7]

---

1 John Dewey 1916: 13–14

2 Descartes 1641; There are modern philosophers that also adhere to the tradition of conceiving the mind as essentially isolated, e.g. Locke, Hume and Kant (see Gallagher 2009).

3 Descartes 1641: 51–64, 241

4 Wheeler 2005: 21

5 Wheeler 2005: 21–23; Rowlands 2010: 2; For philosophers who hold on to an isolated mind see e.g. Fodor (1983, 2000) and Adams & Aizawa (2008). In this thesis I will examine more fully Adams & Aizawa's position.

6 Cowart 2005: sec.1; Shapiro 2007: 338–339; Varela *et al.* 1991: 8; Rowlands 2010: 2–6

7 Hurley 2001: 3-4, see also her 1998.

This sort of traditional approach to cognition will henceforth be referred to as *cognitivism*.[8] Even if proponents of cognitivism have moved away from the early simplistic picture of the computer analogy, they still hold on to the notion whereby the brain with its neuronal processes houses and ultimately explains all cognitive processes and mental phenomena. Hence, the mind is still isolated to the brain and the boundary of cognition has been drawn at the skin and skull of the head. As Mark Rowlands succinctly summarises the cognitivist position: 'whatever else is true of mental processes, whether they are abstract formal processes or patterns of activity in a neural network (or both)—they are processes that occur inside the head of the thinking organism.'[9]

Lately in the philosophy of mind and the general field of cognitive sciences, there has risen an interest in the possibility of departing from this cognitivist position. These views react against the sandwich model of cognition because it neglects the role our bodies, the surrounding world, tools and artefacts, and our active participation in this (cultural) environment have in structuring cognitive processing. Some of the authors have started to talk about *4E Cognition*, originally due to Shaun Gallagher[10] and alluding to the nature of cognition being *embodied*, *embedded*, *enacted*, and *extended*. The move from traditional theories of cognition to 4E theories may be characterised as moving from a restricted, or isolated, view of cognition to a more systemic view that encompasses the brain, body and the world as equal partners.

The current debate surrounding these issues is vivid. This is exemplified for instance by journal themes and issues dedicated to exploring the 4E nature of cognition. See for instance *Phenomenology and the Cognitive Sciences (2010) 9(4): Special issue: 4E Cognition: Embodied, Embedded, Enacted, Extended*; *Topoi (2009) 28(1): The enacted mind and the extended mind*; *Teorema (2011) 30(2): Special Section: The Extended Mind;* and *Cognitive Systems Research (2010) 11(4): Special Issue on Extended Mind*.

Yet, the beginnings of 4E may be traced back to the early 1990's. Early seminal works include the following: in embodied and embedded cognition John Haugeland's 1998 paper 'Mind Embodied and Embedded' and Andy Clark's 1997 book *Being There*; in enactive cognition Varela, Thompson, and Rosch's 1991 book *The Embodied Mind* (where ideas of embodiment and embeddedness were also present); in extended cognition Andy Clark and David Chalmers' 1998 paper 'The Extended Mind' (although similar ideas were already present in Clark's 1997 book and Rob Wilson's 1994 paper 'Wide Computationalism').[11]

In addition to the four Es there could be rightly added a D; that of *distributed* cognition developed most notably by researchers in the University of San Diego (California), such as Edwin

8 Varela *et.al.* 1991; Other terms for this type of approach to cognition include *individualism* (Wilson & Clark 2009; Wilson 2004; Menary 2010b), *classicism* (Cowart 2005), *isolationism* (Clark 1998a), and *brainbound* (Clark 2008a).

9 Rowlands 2010: 2–3

10 See Gallagher (2008a) and Rowlands (2009b: 3, 219n4).

11 There are many other early theorists that share a commitment to 4E type of approach to cognition. See for instance Lakoff & Johnson (1980), Hurley (1998), Van Gelder (1995), Rumelhart et al. (1986), Beer (1990), Ballard (1991) and Järvilehto (1998a, b; 1999; 2000; 2009)

Hutchins, Jim Hollan and David Kirsh in the late 1980's and 1990's. Hutchins' 1995 book *Cognition in the Wild*, may be regarded as the seminal work in this context. The relevance of distributed cognition comes from its similar and pertinent theoretical commitments and the growing interest in it in the current discussion about the nature of cognition.

The most significant and fiercest critics of 4ED are Fred Adams and Kenneth Aizawa with their numerous individual and co-authored articles and their 2008 book *The Bounds of Cognition*, and Robert Rupert with his numerous articles and his 2009 book *Cognitive Systems and the Extended Mind*. It is important to note that in addition to the 1990's works the idea of cognition not being restricted to the head is already present in the phenomenological and pragmatic traditions. The ideas of philosopher's such as Husserl, Heidegger, Merleau-Ponty, Peirce, and Dewey bear close resemblance to the contemporary accounts. In fact these ideas have re-emerged in the contemporary 4ED literature and are cited and referenced, if not explicitly made use of, by some of the 4ED theorists.[12]

The aim of this thesis then is to examine, compare and contrast the different 4ED approaches to cognition in relation to each other and the criticisms that have been raised against them.[13] There are four specific research questions that guide this thesis. I) *What are the theoretical commitments the different accounts of cognition in 4ED hold on to and how do these relate to each other?* I deem this question crucial because even if the views share similar commitments they are often unwarrantedly confounded by simply lumping them together without paying attention to their distinct theoretical features. II) *(Why) should distributed cognition be added with the rest of the 4E accounts?* This second question stems from the increasing interest and relevance of distributed cognition in the current debate about the boundaries of cognition. This is most evident in the so-called "second-wave extended cognition" and questions pertaining to cognitive agency. Yet distributed cognition has not been accredited a detailed analysis in the context of 4ED. Hence, I deem it important to provide one, and in doing so to see how it might offer a unique point of view and insight to the nature and study of cognition. III) *What critique does 4ED face and how does providing a mark of the cognitive affect both the proponents as well as the opponents of 4ED?* The relevance of a mark of the cognitive stems from the challenges that have been raised against 4ED. Yet importantly, it seems that it levels the ground between 4ED and the orthodox cognitivism. This again highlights the possible viability of the 4ED approach. There lies also the independent (meta-level) issue of the need of providing this kind of mark in the first place in order to study cognitive phenomena and draw the boundaries of cognition. IV) *How does the notion of cognitive agency*

12 For accounts of the historical roots of 4ED type approach to cognition and mind see e.g. Clark (1997: ch.8 sec. 8), Thompson (2007: Part one), Gallagher (2009), Anderson (2003), and Clancey (2009).

13 In this thesis there will be a conceptual distinction between mind and cognition. Mind and mental states in general are distinguished from the broader notions of cognition and cognitive processes in that they essentially involve experiential states (e.g. sensation of pain) and propositional attitudes (beliefs, desires, intentions, etc.) The focus in this thesis will be on cognition. (Clark & Chalmers 1998: 12; Clark 2008a: xxviii, 233n4; Drayson 2010: 374–375; Rupert 2004). Also cognition and cognitive processes need/should not be understood as *conscious* phenomena. (Clark & Chalmers 1998: 10; Adams & Aizawa 2001: 48)

figure in the 4ED approach(es) to cognition? If one is persuaded about the viability of 4ED, then interesting issues are brought to light with regards to how the different views conceive cognitive agency. Each of the 4ED theories have a somewhat different take on agency, and hence it is interesting to see what kind of similarities and differences between the approaches are highlighted.

The first research question concerning the theoretical commitments of the 4ED approaches runs through the whole thesis. Yet specifically, chapters 2–5 focus exclusively on looking into the specific theoretical commitments of each individual 4ED approach. The second research question to do with the status of distributed cognition as part of 4E is covered in chapter 5. After examining the various notions of 4ED I will inspect in chapter 6 the most notable criticisms and defences of 4ED, as well as the interesting issues this debate brings forth concerning the mark of the cognitive (the third research question). The exploration into the notion of cognitive agency follows in chapter 7 (the fourth research question). There I will look into how the different 4ED approaches conceive cognitive agency, and how this highlights both the theoretical overlap and differences between the approaches. Finally chapter 8 concludes the thesis with a summary of the main aspects that are covered under the four research questions. I will ponder the significance of this sort of discussion in general, and in particular whether a comprehensive framework for cognition is seen to emerge. I will also look at some open questions and further possible directions for the applicability of 4ED.

The philosophical relevance of the issues examined in this thesis vary from the principal and long-lasting question concerning the nature of mind and cognition to the importance of providing an elucidatory examination of the unorthodox approaches to cognition. Due to the current vivid debate surrounding these 4ED approaches it is important to clarify, compare and contrast the various positions since this sort of examination is lacking in the literature. Indeed as will become clear in the course of this thesis there is overlap between the approaches exemplified by the continuing refinement, development and evolution of the conceptions. Thus, by first establishing a solid grounding on the various conceptions one is then later able to take on other interesting philosophical questions. For instance there is a host of normative questions that arise from allowing cognition to include cultural artefacts and other people. Also it would be very interesting to see how the role of representation varies across the 4ED approaches and what this would mean for the orthodox conception of representation. Since the 4ED approaches have a multi- and interdisciplinary backgrounds, with the authors taking inspiration from various empirical fields such as cognitive science, artificial intelligence, neuroscience, biology and psychology, the current investigation also pushes the boundaries of traditional philosophical inquiry. Hence, I ultimately hope this thesis helps to articulate and provide nouvelle points of view to the understanding and study of cognitive phenomena and this way provide a basis for further philosophical investigation that might precipitate from the sort of approaches to cognition under discussion here.

2 Embodied-Embedded Cognition

> 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.[14]

This chapter explores the notion of embodied-embedded cognition, which comprises the first two Es of 4E cognition. I have grouped together the notions of embodiment and embeddedness for four reasons. 1) They have risen quite simultaneously and interdependently. 2) Embodiment has often been used in the sense of embeddedness, or interchangeably with It. I.e. emphasis has often in the literature been placed both on the role of the environment (which is the focus of embeddedness) and the body (which is the focus of embodiment) in structuring cognition.[15] 3) Conceptually they are quite close, since it is not a huge leap to allow the environment to have a significant influence on cognition once the boundary of the skull is broken and the body is allowed to have a more significant role in structuring cognition. 4) Yet, there is a conceptual difference between the two notions,[16] hence I have not merely stuck with talking about either embodiment or embeddedness. So, the dual notion aims to capture the theoretical proximity of embodiment and embeddedness as well as their distinct character.

Embodied-embedded cognition has risen from various fields and the following is only a sampling of the early authors introducing the ideas of embodied-embedded cognition. Linguistics: Lakoff & Johnson (1980); dynamical systems theory: Van Gelder (1995); developmental psychology: Thelen & Smith (1994); systems psychology: Järvilehto (1998a, b, 1999, 2000); artificial intelligence: Brooks (1990, 1991a, b, c), Beer (1990); cognitive science: Varela et al. (1991), Kirsh & Maglio (1995) philosophy: Wilson (1994), Clark (1997), Haugeland (1998), Hurley (1998). Furthermore, in addition to the phenomenologists and pragmatists mentioned in chapter 1 the inspiration and roots of embodied-embedded cognition may also be found from the work of Vygotsky (1979, 1986), Von Uexküll (1934) and Gibson (1979). In what follows I will examine further some of these authors.[17]

The point of reaction embodied-embedded theorists take is most often with cognitivism. The cognitivist commitments that people react against are well described by Michael Wheeler, who distinguishes eight principles in cognitivism that advance a picture of an isolated mind:

---

14 Haugeland 1998: 236–237 original emphasis
15 See e.g. Cowart (2005) and Shapiro (2007)
16 See.g. Haugeland (1998), Wilson (2004), Gallagher (2005), Kiverstein & Clark (2009) and Rupert (2009a)
17 For overviews on the subject see Anderson (2003), Wilson & Foglia (2011), Cowart (2005).

1) In cognitive activity there persists a *subject-object dichotomy* between the cognitive organism and the environment.
2) Mind and cognition are to be explained in terms of *representational states* and the way these states are structured, transformed and manipulated.
3) Intelligent action is the outcome of *general purpose reasoning* that works by retrieving the relevant representations for the behavioural context and then manipulating and transforming those representations appropriately so as to determine the corresponding correct behavioural output.
4) The nature of human perception is essentially inferential
5) The connection between perception and action takes the sequential form of *sensing, representing, planning* and finally *acting* (or moving).
6) The cognitive agent is essentially *disembedded* from the environment due to the three-fold nature of the environment: a) the environment merely holds problems for the agent to solve, b) it merely provides informational inputs through the senses, and c) it is merely the stage where reflected actions are acted out.
7) The cognitive agent is also *disembodied* in the sense that in cognitive scientific explanations no conceptual or theoretical room is made for the physical embodiment of the agent.
8) Psychological explanations do not take into account the *temporal process* of mental phenomena.[18]

I take these eight aspects to be an illuminating description of the general cognitivist position. I do not maintain that all of these are *necessary* aspects of cognitivism. Rather they are here to shed light on some of the characteristics of isolated cognition the cognitivists hold on to as well as to provide a theoretical background from which embodied-embedded theorists depart. I will return to these at relevant places later in this chapter.[19]

Embodied-embedded cognition refers to the way cognition is significantly structured by the type of body an organism has, as well as the type of environment it is situated in. It might seem unsurprising to claim that our actions, perceptions and thoughts are intertwined, or that our bodies and the environment affect our cognition. One often hears descriptions such as "humans are essentially psycho-physico-social beings". But the issue at stake here is on the nature of this intertwinement and the principled separability of mind, body and world. The aim of this chapter is to critically examine the notions present in Wheeler's eight aspects.[20] This chapter takes inspiration from Andy Clark's argumentation in his 1997 book *Being There* because in it Clark presents a concise and comprehensive picture of what it means to say that cognition is embodied and embedded. A picture that synthesises various research programmes in various fields that all put into question the orthodox conception of cognition. Nevertheless, I will largely rely on the original expositions of embodied-embedded cognition that are cited by Clark.

18 Wheeler 2005: 23–53

19 I chose Wheeler's eight aspects as a characterisation of cognitivism due to its comprehensive nature. Other characterisations were also available and valid, but which in my opinion too readily identified cognitivism simply as emphasising internal representations and their manipulation see e.g. Rowlands (2010: 51–52) or Menary (2010c: 605–606). Yet, to be fair contemporary cognitivists such as Adams & Aizawa (2001, 2008, 2010b) explicitly hold on to a view of cognition as essentially involving non-derived representations. If one is interested in Wheeler's full argumentation see his (2005: ch.2–4)

20 Clark 2008a: xxvi–xxviii.

This chapter is structured around various concepts and ideas that put into question the isolated cognitivist approach. These concepts importantly shape the current debate concerning the 4ED nature of cognition. The chapter has two main sections. I will start by showing how cognitive processing should be conceived as decentralised and dependent on the surrounding environment. I will then (re-)emphasise the way cognition is embodied. I conclude the chapter by raising the most pertinent notions of embodiment and embeddedness, as well as elucidate its departure from the cognitivist alternative.

2.1 Decentralisation, Ecology and Scaffolding

Clark argues for a systemic view of cognitive processing that is essentially *decentralised* in the sense that mind, body, and world *all* play an essential role in structuring cognition.[21] There lies important theoretical commitments and insights within this decentralised approach. First, Clark is wary of the excessive need for centralised world-modelling and emphasises a more direct connection between the organism and the environment, as well as between perception and action. Here he relies, among others, on Rodney Brooks' research in artificial intelligence (AI) and robotics.[22] Second, Clark is wary of the idea that the development of cognitive capacities follows a single-cause view of brain maturation and the unfolding of some genetically imprinted *blue-print*. Here he follows developmental psychologists Esther Thelen and Linda Smith.[23] In general decentralisation of cognition emphasises a bottom-up view of operational and developmental *emergence* of higher-cognitive states from lower ones over a more centralised top-down view.

Brooks rejects the orthodox idea that AI systems should be modelled according to a conception of intelligence that essentially consists of the manipulation of information bearing symbols by a central control system. One of the reasons being that the central control system, creates a *representational bottleneck* that hinders and obstructs real-time responses.[24] This is because the incoming sensory information needs to be converted to appropriate symbolic code for the control system to handle it and the output code needs also to be transformed into various motor responses. These transformations make the central control system approach computationally expensive and slow. Brooks dubs this cumbersome framework sense-model (represent)-plan-act, or SRPA for short.[25]

Brooks' alternative approach to avert the looming difficulties is to ground AI systems in the

21 See for instance Clark 1997 and 2008a.
22 See Brooks 1990, 1991a, b, c; Brooks et al. 1988.
23 See Thelen & Smith 1994 and Thelen 1995.
24 Other reasons relate specifically to the way AI systems are designed; often resulting in a detachment from the real-world dynamics and interactions due to a reliance on a "picture" of the world that is fed in by the system's designers. This also leads to a *frame problem* since the AI system may not assume anything besides the explicitly stated matters. (Brooks 1991a: 2–5 and 1990: 1–3; Shanahan 2009)
25 Brooks 1991a: 2–5; 1990: 1–3; 1991b; Clark 1997: 13–15, 21–22 Note also Wheeler's fifth aspect of cognitivism.

physical world and to 'use the world as its own best model'.[26] The driving idea behind this approach is the recognition that all the relevant information may be found in the world, one just needs to sense it often enough to get an up-to-date picture of it. Brooks and his colleagues have implemented this idea by building real-life robots with a computational *subsumption architecture* that avoids the need for a central control system. The architecture consists of multiple complete activity producing systems, or *layers*. These layers work in parallel, are in a direct interaction with the surrounding environment through sensors and actuators, and hence directly connect sensing to action. An aspect of control comes from the layers being in a simple interaction with each other; either turning off (subsuming) the activity of another or joining in on it. This aspect of control brings the layers an *implicit* goal or purpose in their activity, one which they match to the conditions in the surrounding environment. But note that since the systems do not have a central control structure, the apparent central control is ascribed to the system only by its outside observers. The pattern of purposeful behaviour emerges to the observer from the competing behaviours of the layers.[27] An example of a three-layered subsumption architecture is the following: The first layer makes the robot avoid hitting both static and moving objects in the environment, if no avoidance is needed it stays put. The second layer makes the robot wander about in the environment by giving it random headings every 10 seconds, if it is not currently avoiding obstacles. The third layer may suppress the second layer and make the robot explore distant places. These three layers exhibit the three key basic intelligent operations that Brooks argues form the basis for higher-level intelligence: survival related tasks, vision and mobility.[28] An example of a robot with a 14-layer subsumption architecture is a robot called “Herbert” who roams the floors of MIT's robotics lab collecting soda cans, all the time reacting to the ever-changing dynamic environment without a central control system or a plan of the premises.[29] The significance here lies not in the nature of this seemingly mundane and trivial task, but on the way the task is completed. Brooks' research undermines the traditional sandwich model of cognition, and even though his field is artificial intelligence the findings could be applied to mirror a picture of human cognition.

Similarly to Brooks, Thelen & Smith depart from the traditional views of development that approach it from a top-down single-cause view, according to which the development of locomotory abilities for instance are explained in terms of brain maturation, neural growth, or converting reflexes into instrumental actions.[30] Thelen & Smith take the development of walking as a sort of case study through which they argue for their dynamical systems approach of the development of behaviour and cognition.[31] For instance, relying on empirical evidence Thelen & Smith manage to

26 Brooks 1990: 3
27 Brooks 1991a: 5–8; 1990: 1–4
28 Brooks 1991a: 7–10
29 Brooks et al. 1988; Connell 1989
30 Examples of these maturationist and cognitivist approaches include Forssberg (1985), Konner (1991), Zelazo (1984).
31 Thelen & Smith 1994: 3–9. For another good and often cited dynamical systems approach to cognition see van Gelder (1995).

show how learning to walk consists of many subcomponents that do not develop linearly, and how this nonlinear development is structured both by bodily and environmental factors. There are two important notions here that I take to be most relevant and illuminating for the present discussion.

Self-organisation is the idea of how the sub-components of an open system (in this case the system generally composed of the brain, body and the world) may relate to each other in various unexpected ways that give rise to an emergent organisation.[32] This emergent pattern is something different from the elements that constitute it. So, the insight from self-organisation is that high-level patterns emerge from the interaction of distributed simple components that follow local rules and principles without the need of a central controller, or a developmental blue-print.[33] For instance Thelen & Smith discovered that the development of walking and reaching behaviours are dependent on multiple context-dependent local factors that do not arise linearly, such as bodily growth, environmental factors, alertness, brain maturation and learning. Development of these abilities could not be accounted for in terms of some high-order single-cause hypothesis.[34] *Soft-assembly* relates to the interaction of the different structures (neuronal, bodily and environmental) that influence behaviour and problem-solving activity. Soft-assembly allows the problem-solving activity to be dynamic and responsive to changes in the system by selecting and adapting new solutions found in the relevant structures. For instance, when one finds oneself suddenly walking on an icy surface one needs to readjust to these changes and recruit new aspects of upright locomotion *on the go*.[35] Similarly, consider how Herbert, in contrast to the centralised approach, is soft-assembled since it responds fluidly to the changing environment. The solutions arise and are solved within the context the problem is posed. Thelen aptly illustrates the decentralised approach to cognition the following way:

> A fundamental assumption in a dynamical approach to development is that behavior and cognition, and their changes during ontogeny, are not represented anywhere in the system beforehand either as dedicated structures or symbols in the brain or as codes in the genes. Rather, thought and behaviour are "softly assembled" as dynamical patterns of activity that arise as a function of the intended task at hand and an individual's "intrinsic dynamics" or the preferred states of the system given its current architecture and previous history of activity.[36]

Related notion to decentralisation is, what Clark dubs, *ecological control*. A system exhibiting ecological control achieves its goals by delegating the problem-solving responsibility to

32 Thelen & Smith (1994: 45–54) focus on Belousov–Zhabotinsky reactions in chemistry, and Haken's discoveries of laser. Examples of self-organizing systems are ubiquitous also in biology, paradigm cases being how ants forage and birds flock by following few simple effector cues or rules (see e.g. Clark 1997: 40; Wheeler 2005: 94–95).
33 Thelen & Smith 1994: 54–56, 78–79, 82–83; Wheeler 2005: 94–96; Clark 1997: 40, 73
34 Thelen & Smith 1994. See also Thelen et al. 1982, Thelen et al. 1984 and Thelen & Ulrich 1991
35 Thelen & Smith 1994: 60, 84, 311, 321; Clark 1997: 42–45
36 Thelen 1995: 76

reliable sources found in the body, brain and/or the surrounding environment, not by micromanaging every detail of the desired action or response.[37] For instance the information-processing load required in problem-solving activity may be reduced by 'sensitizing the system to particular aspects of the world–aspects that have special significance because of the *environmental niche* the system inhabits.'[38]

This is what Clark dubs *niche-dependent sensing*.[39] He draws inspiration for this notion from the biologist Jacob Von Uexküll's emphasis on how each animal has its own specific way it perceives and effects the environment it is situated in. This perceptual world relative to a given animal is what Von Uexküll dubs an animal's *Umwelt*. Each animal has its own umwelt, its own specific relation to the surrounding environment that is *carved* out of the general physical world. Von Uexküll insists that the relation between the animal and the objects in its umwelt should be conceived as complete and interdependent, consisting of functional cycles that constitute a systematic whole. For instance the perceptual world of Herbert, its umwelt, due to the nature of its various sensors essentially consists of table surfaces, obstacles to avoid and coke cans that attract.[40] Now, Von Uexküll's insights seem to give a more substantial role for the relation between a given animal's way of life and the specific features that it picks up, attends to, or focuses on in the environment. Clark applies this to human perceptual worlds as well. Instead of building a full picture of the world, humans attend only to parts of the environment. Relying on Dana Ballard's[41] *animate vision* research Clark, paraphrasing Brooks, maintains that we humans can also *use the world as its own best model* and visit and re-visit the surrounding environment as and when needed without the need of building a complete three-dimensional representation, or model, of it. Perception here is deemed more as an exploration of the environment rather than passive reception of stimuli which is then modelled in the head.[42] In John Haugeland's words emphasis is on 'perceiving instead of representing'[43] because the former is cheap and the latter expensive, as became clear in Herbert's case.

Both Haugeland and Clark maintain that James Gibson's theory of *ecological perception*[44] brings forth further insights into the relationship between the perceiver and the perceived, as well as between perception and action.[45] Similarly to Von Uexküll, Gibson places emphasis on the specific nature of the organism, its situation in the environment, and the interdependent relationship between these two structures. Gibson argues for a notion of *affordances* as intimately structuring the way we perceive the world. An affordance is a potentiality for interaction with an

37 Clark 2006: 4–5; 2008a: 5–6
38 Clark 1997: 24 emphasis mine
39 Clark 1997: 23–25
40 Von Uexküll 1934: 6, 10–13; Clark 1997: 24
41 Ballard 1991
42 Clark 1997: 25–31; 1999: 345–346
43 Haugeland 1998: 219
44 Gibson 1979
45 Haugeland 1998: 221–223; Clark 1997: 50–51; 1999: 346–347

object or a state of affairs in the environment, relative to a given animal. That is, a chair for instance offers sitting to a normal height adult, but not necessarily for a child. For a beaver a chair might afford something completely different; if made out of wood it affords nutrition.[46]

I take there to be three pertinent aspects that may be drawn from Gibson's argumentation. First, Gibson takes up the idea of a *niche* but rather than seeing it as *where* animals live, Gibson conceives a niche as *how* animals live.[47] He defines a niche as a set of affordances. Niches understood this way emphasise the complementarity of animals and the environment. A niche implies a certain type of animal as much as an animal implies a certain type of niche. Hence it seems that Gibson's (as well as Von Uexküll's) conception of a niche cuts across the dichotomy of subjective and objective present in cognitivism. Second, Gibson maintains that the environment holds different affordances, or niches, for specific actions (e.g. caves for living or hiding, woods for hunting, water for swimming or drinking, etc.) Moreover, the environment maybe transformed in order to change what it affords us. Affordances are therefore as much to do with the environment as they are to do with actions. Third, Gibson maintains that the *value* or *meaning* of objects and states of affairs may be perceived directly. Here he departs from orthodox psychology, in that objects are not to be understood as being composed of their properties or qualities, and perception of these objects being the discrimination of these qualities. Rather Gibson maintains that 'what we perceive when we look at objects are their affordances, not their qualities.'[48] So, *direct perception* means that one does not perceive value-free objects and then somehow add meaning and value to them, rather perceiving an affordance is perceiving a value-rich ecological object. Gibson calls the perceivable features that designate some affordance *high-order invariants.*[49]

This third aspect is what Haugeland takes to be Gibson's most significant and controversial insight. What is pertinent here is not the idea or possibility of affordances, but that they could be *perceived* rather than *inferred.*[50] Gibson himself notices this when he says that 'the central question for the theory of affordances is not whether they exist and are real but whether information is available in ambient light for perceiving them.'[51] Haugeland points out that if it indeed is the case that affordances can be a feature of the ambient light itself and we may pick these up through our perceptual system, then our perceptual capacities are more complex than previously thought and it is this complexity in the perceptual system that ties us *intimately* with the complexities of the world.[52] Clark makes a similar point and emphasises how the possibility of action-relevant information being available for our perceptual capacities lends credence to a more intimate

46 Gibson 1979 (especially chapter 8)
47 The notion of a *niche* is present in ecology in general. Yet, note how Gibson's conception of niche closely resembles Von Uexküll's notion of *Umwelt* since they both place emphasis on the specific lifestyle of the animal and not merely its location in the environment.
48 Gibson 1979: 134
49 Gibson 1979: 127–130, 133–141
50 Note the departure again from cognitivism exemplified by Wheeler's fourth aspect.
51 Gibson 1979: 140
52 Haugeland 1998: 221–223

connection between perception and action. This again undermines the need for the cumbersome SRPA routine and emphasis the close coupling of the cogniser and the world. Clark illustrates this by a finding[53] that shows how in order to catch a fly ball no complex calculations of trajectories, accelerations and distances are required. Instead one merely needs to run so that one does not allow the ball to look like it's curving towards the ground, rather one keeps the ball so that it appears to move in a straight line in the visual field. So, instead of perceiving, representing, planning and acting, one picks up a relevant higher-order invariant through perception and coordinates/adapts one's actions in real time with the environment so as to fulfil the goal.[54]

The final aspect of this section relates to the reliance on environmental structures as *external scaffolding* that ease cognitive processing. The notion of scaffolding originates from the developmental psychologist Lev Vygotsky, who argued for the way external structures such as public language influences the development of cognitive capacities. Vygotsky's related notion is the *zone of proximal development*, which pinpoints an optimal point of learning where a task is not too easy to be accomplished alone, nor is it too hard that one simply cannot do, but is one that may be accomplished with the help of others. It refers to the way development is *diachronic* in nature. Primarily the development of different capabilities such as language happens in the *intermental* level where it is effected and facilitated by relying on the help of parents for instance, before the development continues internally, in the *intramental* level. I will return to this in chapter 5 where Richard Menary makes use of these Vygotskian ideas in his account of extended cognition.[55]

In the context of embodied-embedded cognition external scaffolding refers to external resources that may include cultural ones such as language as well as specific inanimate artefacts such as pen and paper. For example, the use of pen and paper facilitates and expands our mathematical skills. It is far easier to calculate long multiplications with the help of pen and paper rather than solely in the head.[56] In a way one can *off-load* some of the cognitive burden on to the external tools, whereby one can directly perceive the numbers and operations needed to be done. The recognition that environment may be changed and external scaffolding built is not new, but the notion pertinent for present discussion is a process called *cognitive niche construction*.[57] Specifically it is a

> process by which animals build physical structures that transform problem spaces in ways that aid (or sometimes impede) thinking and reasoning about some target domain or domains.[58]

53 McBeath et al. 1995
54 Clark 1999: 346–347
55 Vygotsky 1934; Clark 1997: 45, see also ch. 10
56 This pen-and-paper example has become one of the standard examples in the 4ED literature, originally the analysis is due to Rumelhart et al. (1986).
57 Clark 1997: 45–47; 2008a: 61–63; 2005a: 255–257; Wilson 2004: 218–219; Wilson & Clark 2009: 58–61
58 Clark 2008a: 62

This transformation of the relevant physical problem-solving space is illuminated by a distinction between *pragmatic* and *epistemic actions* introduced by Kirsh & Maglio.[59] Pragmatic action is one taken so as to alter the world so that one's goal is achieved. For example, taking a glass from the cupboard in order to have a glass of milk. Epistemic action in contrast is one taken so as to alter the world so that one's cognitive, or problem-solving, load is eased. For instance, the grouping of jigsaw puzzle pieces according to shape and colour, or the physical manipulation of scrabble tiles, aid the problem-solving tasks involved by facilitating searching, and pattern and word recognition.[60]

Generally the reliance on the surrounding environment in problem-solving and information-processing tasks is captured by two of Clark's principles:

> *007-Principle*: In general, evolved creatures will neither store nor process information in costly ways when they can use the structure of the environment and their operations upon it as a convenient stand-in for the information-processing operations concerned. That is, *know only as much as you need to know to get the job done*.[61]

> *Principle of Ecological Assembly*: [...] the canny cognizer tends to recruit, on the spot, whatever mix of problem-solving resources will yield an acceptable result with a minimum of effort.[62]

The essence of these principles is that the cognitive cost may be shared amongst all the components and structures of the hybrid cognitive system, and that the cognitive agent is not biased towards any of the available resources, whether internal or external. I will return to Clark's conception of agency in chapter 7.

2.2 Taking Embodiment Seriously

In the sections above I have talked about some of the main concepts and notions that I take to be crucial for an understanding of cognition as embodied and embedded. Yet, it might seem that the emphasis has been more on the embedded nature of cognition, how it is intimately tied to the surrounding environment. Hence, I will now briefly raise some of the embodied aspects of cognition that might have been neglected or obscured in the preceding discussion.

Firstly, from their infant studies Thelen & Smith found that the structural properties of the infants' legs greatly affect infant stepping. For instance it stops due to gain in the body mass of legs,

59 See Kirsh & Maglio 1994 and Kirsh 1995
60 Clark 1997: 62–64; Kirsh & Maglio 1994: 513–515; Kirsh 1995
61 Clark 1989: 64 emphasis mine
62 Clark 2008a: 13 italics removed.

yet reappears when the mass is reduced (e.g. due to buoyancy by placing the legs submerged in water).[63] Similarly the springlike properties of tendons and muscles in legs allow coping with the changes in the dynamical environment so that normal walking and gait is possible.[64]

Yet, more importantly in his 2005 book *How the body shapes the mind* Shaun Gallagher takes up the issue of our embodiment and its relevance for our cognition. Specifically, he re-analyses the distinction between a *body image* and a *body schema*, that according to him has historically been confusing.

> A body image consists of a system of perceptions, attitudes, and beliefs pertaining to one's own body. In contrast, a body schema is a system of sensory-motor capacities that function without awareness or the necessity of perceptual monitoring [...] So the difference between body image and body schema is like the difference between a *perception* (or conscious monitoring) of movement and the actual *accomplishment* of movement, respectively.[65]

Gallagher supports this distinction with empirical evidence of two types of pathologies where either the body image is intact and the body schema dysfunctional or vice versa. In cases of *unilateral neglect*, due often to brain damage from a stroke, the patient's body image is impaired while the body schema remains unaffected. The patient for instance might not be able to perceive her left side of the body, it remains unattended in the sense that it might not be dressed or the left side of the head is not combed. Yet, there is no effect on the motor functions; the subject is capable of normal gait (yet she would not notice if her left shoe were to drop while walking) and the performance of movements that require both hands such as buttoning her shirt. In contrast, in rare cases of *deafferentation* patients who have lost all tactile and proprioceptive input from below the neck the body image remains intact while the body schema is impaired. These patients do not have a sense of body posture or limb position, and yet they may still control their limbs and movements (only) through visually guiding them.[66]

In what follows I will mainly focus on the body schema since I deem it more crucial for helping to understand the relevance of the body in the embodied-embedded approach to cognition. Yet, it is crucial to recognise the importance of drawing this distinction in the first place, since it undermines the (cognitivist) idea of the body's trivial embodiment; as merely a lump of matter that simply houses the brain and is directed by the brain. Hence, Gallagher's exposition aims to show that the role of the body is more intricate and offers more to our cognitive performances than traditionally thought.

---

63 Thelen & Smith 1994: 12–13, 94–121; See also Thelen & Ulrich 1991

64 Thelen & Smith 1994: 60; Clark 1997: 42–43

65 Gallagher 2005: 24 original emphasis.

66 Gallagher 2005: 24–25 and ch.2 for a more detailed discussion of the pathologies and the neuroscientific aspect of the body image–body schema distinction.

A body schema is a system of sensory-motor functions and processes that are subpersonal. They involve *close to automatic* tacit performances that dynamically govern posture and movement. A collection, or repertoire, of motor schemas Gallagher calls *motor programs*. These are, on the behavioural level, flexible and corrigible patterns, that are either entirely learned, or if innate nevertheless enhanced through practice and habituation. Examples of these sorts of behaviours include, swallowing, reaching, grasping, walking and so forth. Similarly, when we are learning to ride a bicycle or catch a fly ball we first closely attend to every (motor) detail within the task, but once the ability is mastered none of these facts are longer attended to; the attention is more directed at the more general task in hand. Hence, Gallagher, is strict about saying that these body schematic operations are not reflex-like or completely automatic. Rather, they are close to automatic in the sense that they are often part of voluntary, goal-directed, intentional acts. One final aspect I want to raise about body schemas is the way they allow a close integration or immersion with the environment. For instance body schemas may incorporate objects into it such as a blind man's cane[67] or a carpenter's hammer. The basis for this often lies in their use in some goal-directed or intentional action.[68] Moreover, this integration has a neuronal basis.

> This extension of the body schema into its surrounding environment is reflected in its neural representations. Not only do bimodal premotor, parietal, and putaminal neuronal areas that represent a given limb or body area also respond to visual stimulation in the environmental space nearby, for some of these neurons the visual receptive field remains "anchored" to the body part when it moves (Fogassi *et al.* 1996; Graziano and Gross 1998; Graziano *et al.* 1994).[69]

Yet, Gallagher is adamant that the neuronal representations do not explain everything, since the past history, practice and habituation of the agent, as well as the environmental situation and the intentionality inherent in the action all influence the body schematic performance.

> The body schema, however, is not reducible to a purely neurophysiological explanation of motor control, since the way my body moves is in support of my pragmatic intentions and in response to environmental features that either afford or prevent my action, or make it difficult.[70]

67 Gregory Bateson (1972: especially 6.4: 325) questioned where the boundaries of the unit of cognitive analysis should be drawn. He used the example of a blind man exploring the world with his cane to argue that one should not delimit or define the boundaries *a priori*. Rather one necessarily needs to incorporate the given situation and environment where the cognitive processes develop and operate in within the unit of analysis. Bateson's argumentation could be seen as a predecessor to extended and distributed theories of cognition (see e.g. Hutchins 2010).

68 Gallagher 2005: 26, 32–34, 36–39, 47–48

69 Gallagher 2005: 37. Also Maravita & Iriki's (2004) empirical studies with Macaques' tool use strengthen the case for body schematic integration of tools through their intentional use.

70 Gallagher 2005: 246, see also 138–146 for empirical support of this.

This echoes Gallagher's notion of a *prenoetic* performance of the body, which is a performance 'that helps to structure consciousness but does not explicitly show itself in the contents of consciousness. In just such performances the body acquires a certain organization in its relations with its environment.'[71] I take this to be a crucial aspect for the way our bodies may be said to implicitly structure our actions, movements and conscious processes.

2.3 Summary

In this chapter I examined embodied-embedded cognition through various concepts stemming from various fields of study. In summing up three themes may be raised, that speak for the impact of embodied-embedded cognition. Namely, when conceiving cognitive processing from a systemic point of view that allows a more substantial role for the world outside the brain to influence cognition, the body and the environment may be said to *coordinate*,[72] *constrain*, and/or *facilitate* organism's cognitive capabilities. This becomes evident when the above discussion is situated with regards to Wheeler's eight aspects of cognitivism.

Coordination arises when cognition is decentralised and perception and action are connected more dynamically and directly with the surrounding environment. There is no longer need for a general purpose reasoning engine (aspect 3) that executes a cumbersome, cognitively slow and expensive sense-represent-plan-act -cycle (aspect 5). Moreover, perception is no longer thought to be inferential, but more directly in tune with the dynamic environment (aspects 4 and 8). Here are already present some facilitating aspects, namely the reducing of the cognitive processing load by getting rid of the sense-represent-plan-act -cycle. Yet, the facilitating and constraining aspects are more evident when the relationship between the organism and the environment is no longer conceived to be dichotomous (aspect 1) but to exist synchronously. This is exemplified by *environmental niches* and *Umwelten.* Moreover this allows the organism's body and the environment to actively structure cognitive processing (aspects 6 and 7): both constraining the possible processes (due to for instance our perceptual capabilities) and facilitating them (for instance by off-loading some of the work to external scaffolding structures).

71 Gallagher 2005: 32. See also p.2–3.

72 The term *coordinate* could be replaced with e.g. *regulate* or *guide.* The point being not to fall back onto a cognitivist reading that postulates some kind of central control system. See e.g. Wilson & Foglia (2011) and Port and van Gelder (1995).

3 Enactive Cognition

> Upon analysis, we find that we begin not with a sensory stimulus, but with a sensori-motor coordination, the optical-ocular, and that in a certain sense it is the movement which is primary, and the sensation which is secondary, the movement of body, head and eye muscles determining the quality of what is experienced. In other words, the real beginning is with the act of seeing; it is looking, and not a sensation of light. The sensory quale gives the value of the act, just as the movement furnishes its mechanism and control, but both sensation and movement lie inside, not outside the act.[73]

In this chapter I will examine a related view to the embodied-embedded approach, that of *enactive cognition*, or *enactivism*. I will distinguish two projects or strands within the enactive approach, one more concerned with the nature of cognition and the other partly building upon this and putting forward a (sensorimotor) theory of perception. My aim is to provide a general account of enactive cognition prevalent in contemporary discussion. There are other varieties of enactivism that might not completely fit this picture, such as Daniel Hutto's *Radical Enactivism*, which, for the sake of brevity will not be examined in detail here. I have also excluded any discussion on enactivism's take on consciousness and the explanatory gap even though they feature prominently in it. This is because, if included, it would take us to far afield from the purview of this thesis.[74]

3.1 Enactivism

Originally the terms *enactivism*, *enaction* and *enactive cognition* are due to Varela, Thompson and Rosch in their 1991 book the *Embodied Mind*. There they set out an approach to cognition that departs from, and puts into question, classical cognitivism with its emphasis on the centrality of representation in cognition. Hence they make the following claim:

> We propose as a name enactive to emphasise the growing conviction that cognition is not the representation of a pregiven world by a pregiven mind but is rather the enactment of a world and a mind on the basis of a history of the variety of actions that a being in the world performs.[75]

Broadly stated then, enactive cognition emphasises the way cognition depends on the activity of the cognitive agent. This is in contrast to the cognitivists' sense-represent-plan-act -cycles, which

73 Dewey 1896: 358–359

74 If one is interested in Hutto's work see e.g. a collection of papers edited by Richard Menary (2006b); on issues to do with consciousness see e.g. Thompson (2007), Hurley & Noë (2003), and O'Regan & Noë (2001b).

75 Varela et al. 1991: 9

according to enactivism neglect the dynamic nature of the environment and the mutual dependence of the cognitive agent and the environment.[76] More specifically, there may be distinguished five essential, mutually interdependent, ideas within the enactive approach to cognition.[77]

1) *Autonomy*: living organisms should be understood organisationally in terms of autonomous systems that generate and sustain their own activity so as to bring forth, or *enact*, their own cognition. Hence, systems capable of cognition need to be autonomous in the sense of being capable of self-government. Di Paolo et al. illustrate this the following way: 'if a system "has no say" in defining its own organization, then it is condemned to follow an externally given design like a laid down rail track.'[78] Another aspect of autonomous systems is that they are *operationally closed.* That is, the processes constituting a given system constitute it as a unity, so that if one of these processes is examined then necessarily other processes relevant for the system's constitution will also be examined at some point. This is not to say that autonomous systems are *materially closed*, i.e. closed from outside influences. On the contrary, autonomous systems need to be *thermodynamically open* so that the environment and the system may interact with each other. It is a necessary condition for an autonomous system that it is influenced by processes not belonging to it. In other words cognitive systems must be coupled to the environment. In contrast to cognitivism's sandwich model Di Paolo et al. characterise enactivism's commitment to autonomy the following way:

> Viewing cognitive systems as autonomous is to reject the traditional poles of seeing cognition as responding to an environmental stimulus on the one hand, and as satisfying internal demands on the other – both of which subordinate the agent to a role of obedience. It is also to recognize the "ongoingness" of sensorimotor couplings that lead to patterns of perception and action twinned to the point that the distinction is often dissolved.[79]

2) *Sense-making*: the interactive and relational aspect of autonomy is brought to light by the notion of sense-making. Organisms regulate their interactions with the world in order to transform it into a salient and meaningful *environment*. Into one that is suited for the organism's needs, into an

76 Ward & Stapleton to appear: 3–4

77 I take this division into five key concepts from Di Paolo et al. (2007) and Thompson (2005, 2007). Yet, in giving an account of the five concepts I will be relying on Thompson (2005: 407–409; 2007: parts one and two), Thompson & Varela (2001), Di Paolo et al (2007: 6–14), Thompson & Stapleton (2009: 23–25) and Ward & Stapleton (to appear: 3–12) since their treatment of enactivism is uniform enough and flows from the original treatment present in Varela et al. (1991). I will only make more detailed references either as clarifications or when directly quoting passages.

78 Di Paolo et al. 2007: 8. Note also the resemblance with Brooks' design requirements for AI systems in section 2.2 especially footnote 26.

79 Di Paolo et al. 2007: 8–9. Here they cite, among others, evidence from the evolutionary roboticist Randal Beer's findings of how in simple agents the decision making process (between two actions) is extended in time and does not happen in an instant (Di Paolo et al. 2007: 56n2).

umwelt that possesses value and meaning. Organisms are structurally coupled to the world through their sensorimotor capacities. By autonomously constructing, regulating, organizing and maintaining these exchanges with the environment the organism picks up relevant features for itself and hence creates meaning. The normative status of the interactions arises from the recognition that some of the interactions constrain the autonomy of the organism and some facilitate it. Hence, this autonomous sense-making activity may be characterised as a kind of autonomous adaptivity. Now, enactivism's emphasis on an organism's *activity* may be highlighted by contrasting this notion of autonomous sense-making activity with cognitivism on the one hand, and Gibsonian ecological psychology on the other.[80] First, it is not the case that organisms passively receive information from the environment that they then convert in to an internal representational format. Rather, in their actions they are already directly part of generating meaning; '*they enact* [i.e. bring forth] *a world*.'[81] Second, high-order invariances should not be conceived as something that are out there in the environment, that are then retrieved through our perceptual capacities in direct or indirect means. Information and meaning is not first present in the environment and then extracted from it. Rather, invariances and meaning are the outcome of the *formative* interrelational activity of the organism and the environment.[82]

3) *Emergence*: implicitly in 1) and 2) there is a notion of emergence present. Autonomy arises from the operational closure and is therefore not a property of the components; in sense-making meaning and value are not properties found either in the organism or in the environment but are the properties found in the relational sphere between the two. Di Paolo et al. qualify emergence the following way:

> [...] in order to distinguish an emergent process from simply an aggregate of dynamical elements, two things must hold: 1) the emergent process must have its own identity and 2) the sustaining of this identity and the interaction between the emergent process and its context must lead to constraints and modulation to the operation of the underlying levels.[83]

As I understand it, the first requirement aims at guaranteeing the identifiability of the emergent property that is enabled (but not completely determined) by the component processes. The second requirement alludes to the mutual interdependence between the emerging and enabling processes. Di Paolo et al. cite life itself as evidence for emergence, specifically how for instance cells contain

80 This contrast is largely due to Di Paolo et al. 2007: 9

81 Di Paolo et al. 2007: 9 original emphasis

82 I have left out discussion concerning the notion of *autopoiesis* (developed by Maturana & Varela 1980; and later Weber & Varela 2002) that is often referred to in this context since its relevance as a necessary condition for autonomy, sense-making, and cognition in general is contested (see e.g. Thompson & Stapleton 2009). If one is interested in the notion of autopoiesis I direct you to Thompson (2004 and 2007: ch.5), Di Paolo (2005 and 2009) and the aforementioned literature.

83 Di Paolo et al. 2007: 10

properties that no constitutive part at a lower molecular level contain. Yet, they do recognise that the notion of emergence is problematic and is in need of further analysis along the lines of dynamical systems theory, due mainly to the need of explaining the interactions of different levels.[84]

4) *Embodiment*: enactivism emphasises the way cognition emerges from temporally and spatially embedded *embodied actions*. Specifically enactivism takes embodiment to mean that:

> mind is inherent in the active, worldful body, that the body is not a puppet controlled by the brain but a whole animate system with many autonomous layers of self-coordination and self-organization and various degrees of openness to the world that create its sense-making activity.[85]

Hence it seems that enactivism does not depart much from the theoretical commitments of the embodied-embedded approach to cognition. Where it does depart, I take it, is in the emphasis on the active participation and bodily activity of the agent in generating cognition.

5) *Experience*: according to enactivism (conscious) experiences should be brought into the focus of scientific investigation and not simply left as a side issue. This is accomplished by bringing in phenomenological analysis to the study. This fifth aspect is mainly a thematic and methodological one, and hence will not be investigated further here.

Taking these five aspects together one may say that for enactivism cognition is essentially relational. It is not fully internal, in terms of being located solely within the head and explained in terms of neural processes alone, since as Thompson & Stapleton make clear: 'Cognition is not an event happening inside the system; it is the relational process of sense-making that takes place between the system and its environment.'[86] Yet, the internal processes are crucial in terms of taking part in enabling the sense-making activity. Di Paolo goes so far as to claim that since cognition is best understood as a process in a relational domain it has no location.

> It simply makes no sense to point to chunks of matter and space and speak of containment within a cognitive system. Inspect a baby all you want and you'll never find out whether she's a twin.[87]

Hence enactivism departs from the isolated cognitivist view of cognition towards a more systemic one, in the sense that it accredits the body, brain and the world all an essential and necessary role in constituting cognition.

---

84 Di Paolo et al. 2007: 10–11
85 Di Paolo et al. 2007: 12
86 Thompson & Stapleton 2009: 26
87 Di Paolo 2009: 19

3.2 Enactive Perception

The other strand or project within enaction is a more focused one and relates to a theory of perception. Specifically I will focus on Alva Noë's *enactive theory of perception* presented in his 2004 book *Action in perception*. My focus on Noë's account is due to three reasons: 1) it offers a good example of implementing an enactive account to the issue of perception, even though it is not a full-fledged derivation of enactivism presented above.[88] 2) It offers a nice illustration of another type of account of perception (this is in addition to Gibson's ecological account presented in the previous chapter) that keeps with the theme of 4ED.[89] 3) It is the account of enactivism that is (most) often contested by the critics. Usually on similar grounds as 4ED in general is challenged.[90]

Noë starts his account by the following description:

> Perceptual experience acquires content thanks to our possession of bodily skills. *What we perceive* is determined by *what we do* (or what ever we know how to do); it is determined by what we are *ready* to do [...] we *enact* our perceptual experience; we act it out.[91]

So for Noë perception should be thought of as action, a type of skilful bodily activity. He motivates this with an image of a blind man perceiving the world through touch; skilfully exploring and probing the environment with a cane. Noë takes this sort of perception as paradigmatic; we do not perceive the world all at once but gradually in time build a picture of it through interactions with it. This account is opposed to the cognitivist approach which sees perceptual systems as essentially located in the brain that build internal representations of the world. Yet, Noë is not as anti-representationalist as for instance Gibson is, and allows representations to have a limited role in cognition and perception. Hence he claims the following:

> No doubt perception depends on what takes place in the brain, and very likely there are internal representations in the brain (e.g. content-bearing internal states). What perception is, however, is not a process in the brain, but a kind of skillful activity on the part of the animal as a whole.[92]

88 Noë 2004: 233n1

89 I do not take issue here whether Noë's account is the most plausible one in the field. For similar approaches and discussion on (sensorimotor) theories of perception see: O'Regan & Noë (2001a, 2001b); Hurley (1998, 2001); Hurley & Noë (2003); Gangopadhyay & Kiverstein (2009); Ward, Roberts & Clark (2011). See also the collection of papers in Gangopadhyay et al. (2010)

90 See Adams & Aizawa (2008), Adams (2007), Prinz (2006), Block (2005), and to an extent Rupert (2009). The argument usually is that Noë's account confounds external causal coupling relations as constitutive ones of cognition. I will return to this sort of challenges in chapter 6.

91 Noë 2004: 1

92 Noë 2004: 2

There are two aspects in Noë's approach: in addition to the claim that perception should be understood as skilful activity, the *content* of perceptual experience is also due to the possession of bodily skills. Specifically Noë's central claim is that the ability to perceive is both dependent on, and constitutive (in part) of, a possession of, what he calls, *sensorimotor knowledge*, which is built up of *sensorimotor dependencies*.[93] Sensorimotor knowledge is an implicit understanding of how our experiences and perceptions are effected by movement. Within this idea lies a, now familiar, (enactive) commitment to the sensorimotor relationship between the perceiver and the perceived. Noë maintains that there are two dimensions of how perception is effected by movement: 1) movement-dependence: our bodily movements adjust the sensory stimuli and 2) object-dependence: when perceiving an object, it is its movements that also change the sensory stimuli.[94]

In addition to these two dimensions there are specific *perspectival properties*, or *p-properties* for short, that effect perception. P-properties are objective environmental properties. Specifically they instantiate how things *look* to a perceiver and they are relational in the sense that 'they are not relations between objects and the interior, sensational effects in us. Rather, they are relations among objects, the location of the perceiver's body, and illumination.'[95] The two dimensions that effect perception mentioned above effect specifically the perception of p-properties and thus allow the perception of how things actually are in the world. For instance we visually perceive tomatoes as spherical, we experience them as round, three-dimensional objects. This is so even though we are only able to see their facing sides and hence are unable to see them completely.[96] Noë explains this phenomena the following way:

> Our perceptual sense of the tomato's wholeness–of its volume and backside, and so forth–consists in our implicit understanding (our expectation) that movements of our body to the left or right, say, will bring further bits of the tomato into view. Our relation to the unseen bits of the tomato is mediated by patterns of sensorimotor contingency.[97]

So, a tomato is perceived as a whole due to our capability of understanding how our potential movements, and the object's movement, would reveal further aspects of it. Here Noë distances himself from cognitivist inferential perception (recall Wheeler's aspect 4) by maintaining a similar commitment to direct perception as Gibson does.

93 Noë 2004: 1–2, 12
94 Noë 2004: 64
95 Noë 2004: 85
96 Noë 2004: 75–79
97 Noë 2004: 63

> We don't conjecture or infer how things are from how they look. In actively encountering the way in which how things look varies with movement, we *directly* encounter how things are.[98]

Finally Noë[99] illustrates the plausibility of his approach to perception by giving an account of a form of blindness that is not due to damage to visual sensory system (such as cataracts, retinal disease, damage to eyes etc.) and hence not due to the lack of sensation or sensitivity. Noë dubs this sort of blindness *experiential blindness* and characterises it as the 'inability to integrate sensory stimulation with patterns of movement and thought.'[100] In order to convince of the existence of experiential blindness and hence the tenability of an enactive approach to perception he cites studies[101] of congenitally blind patients who have had their cataracts (which block incoming light to the retina) removed in order to restore sight. What these cases show is that even though the surgeons restore visual sensations, by removing the cataracts, sight is not restored. The following is a description of Oliver Sacks' patient Virgil's experience after the surgery.

> Virgil told me later that in this first moment he had no idea what he was seeing. There was light, there was movement, there was color, all mixed up, all meaningless, a blur. Then out of the blur came a voice that said, "Well?" Then, and only then, he said, did he finally realize that this chaos of light and shadow was a face—and, indeed, the face of his surgeon.[102]

Noë maintains that since it is the case that patients such as Virgil clearly only have visual sensations of light, colour, movement and so forth, they still lack the ability to see.

> The visual impressions [the patients] now [after their operations] receive remain confusing and uninformative to them, like utterances in a foreign language. They have sensations, but the sensations don't add up to experiences with representational content.[103]

Moreover, perceptual difficulties seem to persists even some time has passed after the operation. Noë cites here Gregory & Wallace's patient S.B. who seems not to be able to make use of, respond to, and understand visual impressions that he receives. He does not pay any attention to things in his visual field unless prompted. Noë argues that S.B. lacks the ability to grasp the sensorimotor significance of the impressions, how the impressions change as he moves, and hence are without

98 Noë 2004: 85 original emphasis
99 Noë 2004: 3–11
100 Noë 2004: 4
101 Gregory & Wallace 1963, Sacks 1995 and Valvo 1971
102 Sacks 1995: 114, cited in Noë 2004: 5
103 Noë 2004: 5

content. S.B. continues to be experientially blind. Noë maintains that the evidence for experiential blindness is substantial and something that may not be captured by a simple inferential input-output picture.[104] Noë's point, paraphrasing Kant, is that: 'intuitions–patterns of stimulation–without knowledge of the sensorimotor significance of those intuitions, are blind.'[105]

3.3 Summary

In this chapter I examined the enactive approach to cognition, both in general as a theory of cognition and in particular as a sensorimotor theory of perception. Common to these approaches is the emphasis on bodily activity and the relational aspect of cognitive processes. Enactivism departs from cognitivism's heavy reliance on internal representations and its sandwich model of cognition. It resembles the embodied-embedded approach in that it accredits a more significant role for the environment and the body to structure cognition. So much so that the brain, body and the world are all mutually interdependent in giving rise to cognition. Here it departs from the embodied-embedded approach since the body and the environment are not thought of merely as coordinating, facilitating or constraining the inner processes. For enactivism cognition has no strict location. *It is a process in the relational domain* between the enabling neuronal processes and the sensorimotor couplings of the organism and the environment.

104 Noë 2004: 5–6

105 Noë 2004: 11. Cf. Kant (1881: 51 [62]): 'Thoughts without contents are empty, intuitions without concepts are blind.'

4 Extended Cognition

If, 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 (for that time) part of the cognitive process. Cognitive processes ain't (all) in the head![106]

The examination of the extended cognition hypothesis[107] in this chapter is divided in to two general parts. I will start by examining what has recently been dubbed *first-wave extended cognition*,[108] understood essentially as parity-based extended functionalism. Here I will rely on Andy Clark's and David Chalmers' 1998 paper 'The Extended Mind', which is regarded as the seminal and most contested paper of extended cognition. I will also show how Clark, and Michael Wheeler in particular, conceive extended cognition essentially in coarse-grained functionalist terms. After this I will move on to the next part, where I will examine *second-wave extended cognition*, which emphasises complementarity and cognitive integration arguments over the reliance on functionalism and parity considerations of the first-wave extended cognition. The main authors here will be Richard Menary, John Sutton and Mark Rowlands. The chapter is thus divided because I deem the current debate around extended cognition to revolve around making this distinction.

4.1 First-Wave Extended Cognition

Andy Clark and David Chalmers[109] (C&C hereafter) are motivated to challenge the cognitivist prejudice of isolated *headbound* cognition and put forward an argument for the literal extension of cognition and mind beyond the organism. They take a step further from the embodied-embedded nature of cognition, and demand that the *epistemic credit* involved in epistemic actions should be spread so that the external parts also count as proper parts of the cognitive process. In order to convince us about the plausibility of this C&C put forward a kind of *veil of metabolic ignorance* for identifying cognitive systems. That is, they question what it would take for us to identify an element or a process as (part of the) cognitive. This is the essence of their *parity principle*, quoted in the beginning of this chapter.[110] The parity principle urges one to ignore the apparent importance of

106 This is the so-called *Parity principle* found originally in Clark & Chalmers (1998: 8). Here it is in an amended form with the qualification 'for that time' found in Clark (2005b: 2) with original emphasis.

107 Similar approaches to cognitive extension under discussion here go by various different names such as *vehicle externalism* (Hurley 1998; Rowlands 2006), *locational externalism* (Wilson 2004), *wide computationalism* (Wilson 1994), *active externalism* (Clark & Chalmers 1998), *environmentalism* (Rowlands 1999), *cognitive integration* Menary 2007, 2010a), and the *amalgamated mind* (Rowlands 2010). In what follows I will examine further some of these approaches.

108 The distinction between first- and second-wave extended cognition is originally due to John Sutton (2010).

109 Clark & Chalmers 1998

110 Clark & Chalmers 1998: 8; Clark 2007: 166-167

the head and to focus rather on the functional role played by the different parts of a cognitive system. Hence contrary to some critics[111] it does not claim that the external part of the cognitive system works (or needs to work) exactly like the inner part, nor does it say anything about the necessary or sufficient conditions for an external *vehicle of cognition*.[112] In what follows I will return to the points about functional role and vehicles of cognition.

C&C argue for *active externalism*, which they distinguish from Putnam and Burge[113] type *passive* externalism of content where the relevant external features do not drive cognition in the present situation, in the hear-and-now. In contrast active externalism grants the external features an ineliminable role in affecting cognitive processing. This is exemplified by combining active externalism with the notion of a *coupled system,* to which has already been alluded to in the previous chapters. C&C maintain that human's may link with external scaffolding in a two-way interaction that allows the creation of a closely coupled cognitive system in its own right. The essence of parity, active externalism and closely coupled systems is captured by the following quote:

> All the components of the system play an active causal role, and they jointly govern behaviour in the same sort of way that cognition usually does. If we remove the external component the system's behavioural competence will drop, just as it would if we removed part of its brain. Our thesis is that this sort of coupled process counts equally well as a cognitive process, whether or not it is wholly in the head.[114]

So for instance the formation of words by rearranging scrabble tiles, the completion of a jig saw puzzle by grouping the pieces, or calculating with pen and paper may be seen as extended cognitive processes in themselves, and not just as epistemic actions easing the cognitive processing load. According to C&C they are best characterised as part of thought, not as individual actions.[115]

One could raise the question here[116] whether there is a possible way of conceiving epistemic actions as *not* extended within the extended cognition framework? I think a simple answer would be: no, they should always be thought of as extended cognitive processes. Yet this implies another more interesting question for extended cognition: why should we prefer extended cognition over the embodied-embedded way of conceiving epistemic actions? One possible answer here would be that it is simply wrong to stick with embodied-embedded approach and that extended cognition conceives the nature of cognition correctly, i.e. it is explanatorily more plausible. Perhaps it brings forth new and exciting ways of conceiving the use of artefacts and tools in cognitive activity that the embodied-embedded way neglects. Another related answer is that

111 See e.g. Adams & Aizawa (2001, 2008, 2009, 2010a) to be discussed more fully in chapter 6.
112 Clark 2007: 166-167; 2005b: 2
113 Putnam 1975; Burge 1979
114 Clark & Chalmers 1998: 8–9
115 Clark & Chalmers 1998: 8–12.
116 I would like to thank Ilmari Hirvonen for pointing this out to me.

extended cognition has more explanatory power for cognitive science in general.[117] Now, interestingly Robert Rupert[118] takes up this sort of argumentation but uses it in favour of the embodied-embedded approach. I will return to Rupert's argument in section 7.2, since it brings to light further interesting and important differences in 4ED with regards to cognitive *agency*. But, note that Mark Sprevak[119] argues against this sort of general argumentative tact for or against extended cognition that is based on "inference to the best explanation" considerations. Relatedly, Samuli Pöyhönen validly recognises that none of the authors that rely on explanatory power argumentation has provided a systematic account of the notion of explanatory power they utilise. Interestingly Pöyhönen aims to provide one and in doing so offer an independent way of demarcating cognitive systems.[120]

Now, let's get back to a few more important theoretical features of first-wave extended cognition. First, Clark maintains that the possession of a *contentful* cognitive state is best thought of as a property of the whole active system. Hence, he insists that what is specifically at issue in extended cognition is that the *cognitive vehicles*, that are capable of bearing cognitive content and thus enable the systemic-level contentful states, can be distributed across the brain, body and the environment.[121] Second, for an external element to be considered a cognitive vehicle it needs to be poised to fulfil three criteria: *reliability*, *accessibility* and *automatic endorsement*. A cognitive vehicle is reliable if it is portable and available for use as and when required. It is also easily accessible when it is always at hand and routinely appealed to. Finally, any information retrieved from the vehicle needs to be deemed trustworthy and not routinely put into question, i.e. more or less automatically endorsed. By fulfilling these three criteria an external cognitive vehicle may be said to become *transparent* for the system's cognitive processing. This transparency is purported to be akin to the functioning of working biological cognitive vehicles.[122]

One could raise the question here whether these conditions are too individualistic and presupposes a model of an already formed cognitive agent that might pose problems for the tenability of extended cognition.[123] For instance, one could argue that the point of view about cognition inherent in extended cognition is wrong-headed; it is not that cognition is first in the head and then possibly gets extended, but that cognition should from the start be conceived as a *distributed* cultural phenomenon for instance. Rupert recognises a similar point and asks whether these conditions presuppose some kind of persisting systemic integrity, which might bring forth

117 See e.g. Clark & Chalmers (1998), Clark (2007, 2008a) and Wilson (2004) for this sort of argument.
118 Rupert 2004; 2009a, b; 2010a, b
119 Sprevak 2010
120 Pöyhönen 2012
121 Clark 2005b: 1 (footnote 1). Clark takes vehicles of cognition to carry for instance representations and propositional attitudes. Menary's and Rowlands' take on cognitive vehicles is somewhat different as they understand them as information-bearing structures (Menary 2010a; Rowlands 1999, 2003) I will not adjudicate the possible differences here and will simply take cognitive vehicles to be information-bearing structures and/or to carry representations, propositional attitudes etc.
122 Clark 2008a: 79–80; 2010a: 46–47; 2005b: 2–3
123 I would like to thank Sami Paavola for pointing this out to me.

problems for the notion agency in extended cognition.[124] I will not delve into this now, but will take up these sort of worries in chapter 7 where a more thorough analysis of the different 4ED approaches is made through the notion of agency.

4.1.1 First-Wave Extended Cognition as Extended Functionalism

There may be drawn a close connection between the parity based first-wave extended cognition and *functionalism*.[125] Now, functionalism takes the architecture of the mind to be a complex cognitive system, which is characterised by a network of causal input-output relations. These input-output relations spread out across the brain, body and the world. Individual mental states/cognitive processes, are characterised by their function, or the role, they play within this system of inputs and outputs. So, what matters is not the internal constitution of the mental state or cognitive process, but its causal functional role within the whole cognitive system. Specifically, the role between systemic inputs, systemic outputs, and other systemic states. So, for instance pain can be characterised as being caused by tissue damage and causing wincing and groaning behaviour. One important implication of this view is that cognitive states are *multiply realisable*. That is, functional analysis of the mind does not restrict the type of system that is capable of implementing, or realising, cognitive states. Cognitive states can be realised in physically different systems by various different mechanisms.[126]

When the idea of multiple realisability is compared with the parity principle, the connection between functionalism and first-wave extended cognition becomes clear. Remember that the parity principle urges one to ignore the arbitrary boundary of skin and skull and rather focus on the nature of the processes within the whole system. In purported cases of extension what is important is the functional role the external element contributes to the overall cognitive system or a given process and not the physical structure of the element or the implementing system. Moreover, first-wave extended cognition theorists argue that there is *functional isomorphism* between the inner and outer states and processes and that this functional similarity should be conceived essentially as *coarse-grained*. That is, it is the similarity of the broader common-sense role of the states and processes in driving and explaining behaviour that matters. In contrast, *fine-grained* functionalism requires the realizing processes (of the coarse-grained roles) to be similar in all the resources (inner and outer).[127] So, extended functionalism does not purport to claim that the

124 Rupert 2010a: 330n6

125 For clear arguments for the connection between extended cognition and functionalism see Wheeler (2010) and Clark (2008a, b)

126 Kim 2006: 117–119, 151, 159; Levin 2010: 1, 3.1, 3.2, 4.5; Wheeler 2010: 247–249

127 Clark 2008a: 88–89, 99, 114–115; 1998: 98–100; Wheeler 2011: 419; to appear: 4, 9. Wheeler (2010: 258–261) gestures at the possibility of fine-grained *microfunctionalism* present in extended cognition, and that fine-grained functional roles could also be extended and not solely realised internally. Since for time being this remains merely a gesture, and not a fully developed argument, I will not examine it further.

external resources function exactly like the inner ones. Hence, the possible counter-argument[128] that requires the functional similarity to be fine-grained and essentially modelled according to brain processes might not apply. As Wheeler puts it: the wrong way to understand parity is to

> first […] fix the benchmarks for what it is to count as a proper part of a cognitive system by identifying all the details of the causal contribution made by (say) the brain. Then we look to see if any external elements meet those benchmarks.[129]

This is just the sort of *bio-chauvinistic prejudice* the parity principle was made to undermine in the first place.[130] So, for the first-wave extended cognition it is the coarse-grained functional similarity between the internal and external resources in driving behaviour that ultimately matters in the formation of a hybrid (extended) cognitive system.

To be fair there remains tension between extended functionalism and for instance Rupert's argument that the internal and the external cannot be functionally equivalent since for instance the (fine-grained) processes and properties of biological memory are so vastly different and ones that are not captured by some extended memory systems.[131] Yet, it has been recognised that a *stalemate* is bound to arise between the proponents and opponents of extended functionalism.[132] Rowlands for instance notes that the stalemate arises because Wheeler claims that Rupert's criticisms are question-begging because they require a chauvinistic form of functionalism (that excludes extended cognition), yet it could be argued alike that the reliance on coarse-grained functionalist arguments is also question begging against the requirement for a fine-grained analysis.[133] This problem could then be located at a more general problem with functionalism, raised e.g. by Block,[134] in defining the nature of the inputs and outputs.

4.1.2 Extended *Mind*

As a final aspect concerning first-wave extended cognition I would like to go through C&C's now infamous thought-experiment and argument for the extension of mind and mental states. C&C move on from extended cognitive processes, such as may be present when calculating with pen and paper, to the possibility of extending mind and mental states. They argue that even our

128 Made for example by Rupert (2004, 2009a, b, 2010a).
129 Wheeler to appear: 3–4. See also Wheeler (2010: 249–255) and Clark (2008a: 114–116) for this sort of argumentation.
130 Clark 2005b: 2; 2007: 167
131 Rupert 2004: 405–415; 2009a
132 See e.g. Kiverstein & Farina 2011; Wheeler 2010; Rowlands 2009b; Sprevak 2009; Walter & Kästner 2012
133 Rowlands 2009b: 633–634; Wheeler 2010
134 Block 1978

propositional attitudes; beliefs, intentions, etc., may be extended. They illustrate this with a thought-experiment and ask us to imagine two individuals Inga and Otto, who both are eager to go to the Museum of Modern Art (MoMA). Inga thinks for a moment and recalls the correct address of the museum and then goes to MoMA. In doing this it seems clear that she had a (dispositional) belief of the location of MoMA and that she had that belief even before attending to her memory (before the belief became occurrent). Now consider Otto who suffers from Alzheimer's disease and whose biological memory is degrading. In order to facilitate the structuring of his life Otto relies on a notebook. He carries it with him all the time, writes down new information when encountered and looks up old information when needed. So, when Otto decides to go to MoMA he consults his notebook and looks up the correct address and arrives at the museum. C&C maintain that it seems that Otto arrived at MoMA because he had the desire to go there and the correct belief of the location of MoMA. Importantly, it also seems plausible to say that Otto had his belief even before consulting his notebook, much in the same way as when Inga had her belief before consulting her memory. So, the *role* the notebook plays to Otto is akin to Inga's working biological memory. Hence, the thought-experiment urges us to consider Inga (with her appropriately functioning biological memory) and Otto (with his notebook) on a par and therefore the plausibility of Otto's mental states extending to the notebook he is using.[135]

Moreover, Otto's notebook may be said to be a legitimate physical vehicle of mental content since it fulfils the three criteria C&C set up for the candidates of external cognitive vehicles. The notebook is reliable since Otto carries it with him all the time and uses it as and when required. Hence the information in the notebook is also easily accessible, since it is always at hand and routinely appealed to. Finally, the retrieved information is more or less automatically endorsed. That is, Otto does not routinely put into question the information but deems it trustworthy. Moreover, by fulfilling the criteria of external vehicle the notebook becomes in a way *transparent* to Otto's cognitive processing. This transparency is again purported to be akin to the functioning of a working biological memory. For instance, consider how Inga's incorrect information about or misremembering of the location of the museum is similar to the notebook having incorrect information.[136]

So, it is the close coupling between Otto and the notebook and the active role the notebook plays in the system that makes it the case that Otto and the notebook count as an extended cognitive system, with all the parts of the system enjoying full cognitive status. That is, the notebook contributes to the behaviour of the entire system, in that were it be lost Otto's cognitive capabilities would diminish significantly. Similarly, the notebook plays an active role in belief-formation, for if the information about the location of MoMA were incorrect then Otto would not be able to get there. Therefore, in the belief-formation process the information in the notebook

135 Clark & Chalmers 1998: 12–13
136 Clark 2008a: 79–80; Clark 2010a: 46–47; Clark 2005b: 2–3

functions equivalently to the information in biological memory, i.e. they have similar *functional poise*.[137] Hence, what ultimately matters for making the case for extended mind is the similar functional poise of the information in the two cognitive systems (Inga and Otto + notebook) and the transparency of the notebook in Otto's cognitive processing.

4.2 Second-Wave Extended Cognition

Relating to the importance and the role of the parity principle there is divergence among the first-wave theorists. Whereas Wheeler maintains that 'the parity principle provides the *only* viable basis for [extended cognition]',[138] Clark deems it merely a sufficient condition for cognitive extension. For Clark the parity principle by itself is not meant as an argument for extended cognition but merely as a device to question the bio-chauvinistic prejudice. According to Clark there are two components in extended cognition argumentation: one placing emphasis on functional isomorphism and parity, the other on the distinct yet *complementary* nature of the internal and external resources. As early as 1998 Clark maintained the primacy of the second component when he argued the following way:

> Given this second line of argument (the one stressing complementarity), it is best to see functional isomorphism as at most part of a sufficient condition for cognitive extension [...] The more interesting and plausible argument, I feel, is the one which describes the seepage of mind into the world by stressing that "the brain's brief is to provide complementary facilities that will support the repeated exploitation of operations upon the world [and] to provide computational processes (such as powerful pattern completion) that the world, even as, manipulated by us, does not usually afford" [Clark 1997: 58].[139]

Now, the second-wave extended cognition theorists such as Richard Menary, John Sutton, and Mark Rowlands have recognised this disparity in first-wave extended cognition. They have also recognised the critics' neglect of Clark's point about the two distinct components of the extended cognition argumentation.[140] Sutton for instance validly recognises that most of the criticisms towards extended cognition have been directed at, or are stemming from, the parity principle and the commitment to coarse-grained functionalism. Recall for instance the aforementioned stalemate.[141] Yet, the second-wave theorists have mixed feelings with regards to the first-wave: on the one hand they contend that their approach is not necessarily incompatible with the first-wave, but rather

137 Clark & Chalmers 1998: 13–16; Clark 2005b: 2; 2008a: 79
138 Wheeler manuscript: ch.5 p.2; See also Wheeler 2011.
139 Clark 1998: 99
140 Sutton 2010: 204–205; Sutton et al. 2010: 524–525; See also Kirchhoff 2011 and Walter 2010.
141 For the most adamant critics see Adams & Aizawa (2001, 2008, 2009) and Rupert (2004, 2009a, b, 2010a).

"flows from it", as Sutton puts it. Sutton and Menary for instance do clearly recognise that the payoff from first-wave extended cognition is that it establishes the hybridity of cognitive processing; that cognition involves both internal and external aspects, which are to be integrated into a systemic whole.[142] But on the other hand, the second-wave theorists deem there to be a tension between parity based argumentation and their own. They go so far as to claim that the parity principle is either wrong or incomplete as a motivation for extended cognition since it at least neglects the value in external resources possibly having different character than the internal ones.[143] Therefore, the second-wave theorists emphasise an approach to cognitive extension that is more akin to the second aspect of extension that Clark distinguished (i.e. complementarity).

There are roughly four key notions in play within second-wave extended cognition.[144] The *complementarity*, *integration* and *manipulation* of internal and external resources, and the *transformation* these bring about. Let's start with Sutton's *Complementarity principle*:

> in extended cognitive systems, external states and processes need not mimic or replicate the formats, dynamics, or functions of inner states and processes. Rather, different components of the overall (enduring or temporary) system can play quite different roles and have different properties while coupling in collective and complementary contributions to flexible thinking and acting.[145]

Sutton makes two further qualifications. a) In the idea of complementarity one should resist the tendency of conceiving the internal and external as distinct realms with *fixed properties*. The point being that one should not from the outset, for instance, exclude the possibility of the external resource being as dynamic and fluid as the internal one. The characteristics of internal and external might well turn out to be *interchangeable* for cognitive practice. At this stage Sutton merely warns about not making this mistake and leaves it to further discovery whether the fixed properties view should be abandoned.[146] b) The other aspect relates to the way the *integration* to external complementary contributions may alter, or *transform*, the inner parts of the cognitive system. I will now turn to this second aspect and examine it through Menary's extensive analysis.[147]

Menary adhere's to Sutton's complementarity argumentation,[148] but conceives extended cognition essentially in terms of *cognitive integration*, which may be broadly construed as involving

142 Sutton 2010: 189–190, 193–196, 204–206; Menary 2010a: 227–229
143 Sutton 2010: 198–200; Menary 2006a: 333
144 These notions stem from all three second-wave theorists under examination here; Sutton (2010), Sutton et al. (2010), Rowlands (1999, 2006, 2010), Menary (2006a, 2007, 2009, 2010a, b, c). Each theorist places varying emphasis on these notions, yet I deem the general approach to be similar enough to warrant a comprehensive picture of second-wave extended cognition to be presented.
145 Sutton 2010: 194
146 Sutton 2010: 189, 206–207, 213, 216n.2. Specifically Sutton gestures here at the possibility of developing a new and distinct *third-wave extended mind*. Interestingly, Kirchhoff (2011) has taken up this type of project along the lines of a *dynamical properties view*.
147 Sutton 2010: 206–208
148 Menary 2010b: 571

some cognitive processes as *coordinated processes*. Whereby,

> cognition is the coordination of bodily processes of the organism with salient features of the environment, often created or maintained by the organism. A coordinated process allows the organism to perform cognitive tasks that it otherwise would be unable to; or allows it to perform tasks in a way that is distinctively different and is an improvement upon the way that the organism performs those tasks via neural processes alone.[149]

In cognitive integration, and specifically in those coordinated processes, are two important regulatory aspects. 1) the *manipulation* of external structures as a constitutive part of cognition, and 2) the *cognitive transformation* this manipulation brings about.[150] I will next treat each of these aspects accordingly.

4.2.1 Cognitive Integration: the Manipulation of External Structures

Menary holds a *manipulation thesis*, which he credits originally to Rowlands:

> cognitive processes are not located exclusively in the skin of cognising organisms because such processes are, in part, made up of physical or bodily *manipulation* of structures in the environments of such organisms.[151]

Menary takes the aim of cognitive manipulation to be the restructuring of informational and physical structure of the environment. It is important to note that this thesis seems to be very similar to the theoretical commitments of embodied-embedded cognition. Specifically inherent in the idea of manipulation are the, now familiar, ideas of ecological control, (cognitive) niche construction, and the self-built or natural external scaffolding constructions that are relied upon when completing cognitive tasks. What distinguishes Menary from the embodied-embedded approach is that for him the manipulation thesis is insufficient on its own, due mainly to the purely *causal* nature of the manipulations. Hence he argues that the manipulations need to be *normative* as well. These normative manipulations should be understood as embodied practices stemming from learning, training, and enculturation.[152] Menary distinguishes four broad classes of these normative manipulations.

The first two classes of normative manipulations are nothing new by now. Menary takes

149 Menary 2010b: 563
150 Menary 2010b
151 Rowlands 1999: 23 original emphasis; in Menary 2010a: 228
152 Menary 2007: 84; 2010a: 233; 2010b: 563–564.

epistemic actions, that were already introduced in section 2.2, as instances on normative manipulations. Thus Menary emphasises the fact that not all the relevant states and processes (for the completion of a cognitive task) need to be in the head, and that much of the explanatory power of cognitive abilities is also due to the manipulations of the environment (recall the examples of grouping puzzle pieces and scrabble tiles).[153] Also, what Menary dubs, *biological couplings* are familiar from the embodied-embedded view and enactivism in that they emphasise the close relationship between organism and environment. Menary does not investigate this in detail but merely gives Webb's analysis of phonotaxis in crickets, Ballard's animate vision, Dawkins' extended phenotypes, and O'Regan and Noë's sensory-motor contingencies as examples of biological couplings.[154] One could speculate here whether enactive cognition in general (and not just the perceptual aspect of it exemplified by O'Regan and Noë) would also fit into this picture of biological couplings, since it places a large emphasis on the relational aspect of cognition, approaches it from a biological perspective, and emphasises the structural coupling of the organism and the environment.[155] As a side note, this might also be one of the reasons why Menary and Rowlands say that the second-wave arguments take a more enactive approach to cognition by emphasising the bodily activities in the world (or actions on external information-bearing structures) as a part (in addition to neural processes and vehicles) of the constituting aspects of cognition.[156] Be as it may, with these first two normative manipulations Menary clearly aims to reduce the cognizer's reliance on internal representations of the environment as well as form a more intimate relationship between the organism and the world.

Now, the following two normative manipulations are unique to Menary's exposition. *Self-correcting actions* are similar to epistemic actions in that they direct and structure practical actions when completing tasks. Yet, they are distinct because self-correcting actions do not involve a direct physical manipulation of the environment. Examples of self-correcting actions include the use of props, language and gestures. When describing this Menary focuses on the role of linguistic structures in influencing cognition and future actions. He relies here largely on Lev Vygotsky's ideas and studies of the role of language in infant development and completion of tasks.[157] Menary adhere's to Vygotsky's diachronic account of development. Higher cognitive capacities, such as reasoning, are primarily (developmentally) a social phenomenon that happen in the *intermental*

153 Menary 2010b: 564–568

154 Webb 1994; Ballard 1991; Dawkins 1982; O'Regan & Noë 2001a, b.

155 Menary 2010a: 237 and Menary 2010b: 564.

156 Menary 2010a: 227; Rowlands 2010: 71. See Menary (2009) for a more detailed account of the relationship between extended cognition and enactivism. Note that there is an ongoing debate concerning the possible incompatibility of extended cognition and enactivism. Rowlands (2009b, 2010) argues for the incompatibility of Noë's form of enactivism and his extended cognition, still holding on to the possibility of adhering to some other form of enactivism. Wheeler (to appear) also argues for the incompatibility of extended functionalism and what here has been presented as the general approach of enactive cognition. But see Di Paolo (2009) for a response to Wheeler and the rejection of functionalism but the acceptance of some form of extended cognition closer to that of second-wave extended cognition. See also Thompson & Stapleton (2009) for similar type of argumentation.

157 See section 2.2 for Vygotsky's impact on the notion of scaffolding and his notion of the zone of proximal development. Menary relies here on Vygotsky (1978).

plane. A child is guided in her intermental development and problem-solving activity by her own (activity specific) egocentric speech and the linguistic utterances of her parents (this linguistic help of the parents is captured by the notion of the zone of proximal development). Slowly by accomplishing cognitive tasks this way the need for parental scaffolding reduces and the egocentric speech is internalised into an inner monologue. Hence, the *intramental* cognition then arises only through/after the development of intermental capacities. Yet, Menary agues that the child's speech is not merely accompaniment of her (problem-solving) activity. Rather it gives a cognitive structure for the activity; sequencing the required actions to be taken, searching for solutions, recognising failures etc.[158] Hence, Menary argues that '[s]elf-corrective speech, whether private or public, is used to structure, direct and correct actions that lead to the completion of cognitive tasks.'[159]

The second unique aspect of Menary's normative manipulation is *cognitive practice*, which refers to the way external representational systems are manipulated according to *cognitive norms* in order to complete cognitive tasks. According to Menary a norm is cognitive, as opposed to social or moral, when it is tied directly to the completion of a cognitive task, such as solving a problem, planning or making inferences. The manipulation of external vehicles is normative because the correct way to manipulate them is learned and because they have a goal of fulfilling some cognitive task.[160] Menary investigates writing as an example of cognitive practice. Another one would be the classic example of the use of pen and paper when doing maths. Menary argues that writing paraphernalia: pens, papers, typewriters, keyboards etc. enable the creation of external representations, or vehicles,[161] such as words, sentences, paragraphs and essays, that are not cognitive in themselves. Rather it is the coordination of the internal and external vehicles and the creation and manipulation of these external vehicles that is cognitive.[162] Menary maintains this as one of the crucial features of second-wave extended cognition. So, the external process (the physical manipulation of words and sentences on a paper) is different from, but complementary to, the internal processes and nevertheless should be accredited cognitive status in the overall system. Contrary to purely internal processing, these bodily manipulations of external vehicles allow for a more efficient way of completing complex cognitive tasks, such as writing an essay, since the pieces of text are enduring and stable and hence more easily edited. Keeping and editing an essay solely in the head would be very difficult. It is far more easy to manipulate the world directly.[163]

---

158 Menary 2010b: 564–565, 568–570; 2010a:238–241

159 Menary 2010b: 570

160 Note that Menary does not give a straightforward account of what he means by *normativity* but in this context his way of conceiving normativity seems to be in terms of instrumental normativity, with means–end type oughts, along the lines of von Wright (1963). See also Glüer & Wikforss (2010).

161 Recall (from note 134) that Menary takes cognitive vehicles to be information-bearing structures.

162 Hence contrary to some of the critics (most notably Adams & Aizawa: 2001, 2008) it is not the case that the external artefact, vehicle or representation is cognitive by itself. Rather it becomes a *part* of an extended cognitive system when appropriately incorporated. I will return to this point in chapter 6, when the challenges to 4ED are examined.

163 Menary 2010b: 564–565, 570–571

This cognitive practice aspect of manipulation leads Menary to the second more general regulatory aspect of cognitive integration, that of *cognitive transformation*. Here he investigates the developmental underpinnings of cognition, and gives an account of how the manipulative norms just discussed for instance are acquired. Broadly speaking, cognitive transformations are due to the cognitive and social niche the individual inhabits. For instance, the guidance of parents set out certain social and cognitive norms for the child that she then slowly acquires through learning. Menary specifically looks 'how the normative and social structure of the environment, mediated by learning and training histories, has a direct transformatory effect on the body.'[164] Menary puts forward *a dual component process of transformation*[165] whereby one may either manipulate public symbol systems or internal ones, or a mix of the both, when completing cognitive tasks. He arrives at this through examining Gallagher's discussion of the development of body schemas as a way of integrating to the environment, and the internalisation of number systems that transforms internal mathematical capacities.[166]

Menary's discussion of body schemas is in line with that presented in section 2.3 and hence will not be rehearsed here. Just recall that body schemas subpersonally structure our interactions with the environment, are mastered through learning, practice and habituation, and that they may incorporate tools within it. Specifically, Menary argues that our bodies have eventually been transformed to be capable of writing and drawing, i.e. creating *representational structures*. This is accomplished for two reasons. First, the tools for manipulating representations are incorporated into the body schema the way Gallagher put forward. Second, the manipulations of representations are governed by the manipulative norms set out by Menary that were discussed above. [167]

In addition to the development of these public skills Menary argues that our "innate" internal representational capacities are also transformed through learning. Menary relies on Stanislav Dehaene's and his colleagues'[168] studies and subsequent model of the development of mathematical capacities, that suggest we possess an intuitive understanding of numbers, quantities and addition, upon which the system of cultural symbols of words and numbers are then added. Yet, Menary takes the ontogenetic development of mathematical skills, e.g. the internalisation of the public numeral system, to allow the manipulation of new operations and representational formats and hence the transformation of our cognitive capacities.[169] Rowlands maintains a similar view when he argues that when we manipulate external information-bearing structures, the information may

164 Menary 2010b: 572
165 Menary 2010b: 576. I will follow here Kirchhoff (2011: 7) who recognises three key elements within Menary's account of transformation: 1) *extension* of the cognitive architecture by integration, 2) *enhancement* of cognitive capabilities, and 3) the *reformation* of brain's representational capacities.
166 Menary 2010b: 572–577
167 Menary 2010b: 575
168 Dehaene et al. 1999; Dehaene 2007 and Nieder & Dehaene 2009
169 Menary 2010b: 575–577

be transformed from the merely present to the available, which then may be used in subsequent cognitive processing.[170] Nevertheless, Menary specifically maintains that,

> [t]he deeply transformative power of our learning histories in the cognitive niche is one that *reformats the representational capacities of the brain in terms of public symbol systems*. We internalise public symbol systems in the way Dehaene suggests, but we also learn techniques for manipulating those symbol systems in public space.[171]

So the essence of second-wave extended cognition lies in the recognition of the distinct but complementary contributions environmental structures may have on our cognitive processing and how the successful integration of these environmental resources through their appropriate manipulation constitutes an extended, or more aptly hybrid, cognitive systems that spans the brain, body and the world. Otto + notebook is an extended system and instantiates an extended process of remembering just because Otto is capable of manipulating the external vehicle and the information in it according to some cognitive norms that enable him to fulfil the task of getting to MoMA. Hence, what matters is not the notebook and its functional poise as in first-wave extended cognition but Otto's ability to write down sentences and then retrieve the information contained in them and use this information appropriately in order to get to MoMA.[172]

4.3 Summary

In this chapter I looked into the notion of extended cognition. Specifically I examined two waves of it. Both of these waves have a somewhat distinct take on how best to understand cognitive extension. Yet, they both depart from the general embodied-embedded approach in that they move from a view of environmental resources as external props and aids for fulfilling cognitive tasks to a constitutive claim of their incorporation to the overall cognitive system.[173] They also hold unintelligible to conceive the external elements as cognitive in themselves. Rather, the arguments revolve around incorporating them to the overall (extended) cognitive system as equal partners with the rest of the elements. They also endorse a view where the hybridity of the cognitive system enhances cognitive capacities. First-wave extended cognition relies on parity considerations and the functional isomorphism of the internal and external resources in driving cognition. Whereas second-wave extended cognition is weary of the reliance on parity and functionalism and instead emphasises the complementarity, manipulation and the subsequent integration of the structures.

---

170 Rowlands 2010: 84

171 Menary 2010b: 576 emphasis added.

172 Menary 2006a: 333; 2010c: 609

173 Note that this applies only to the general embodied-embedded approach, since e.g. Gallagher seems to be making the constitutive claim with his notion of body schemas.

5 Distributed Cognition

> The emphasis on finding and describing "knowledge structures" that are somewhere "inside" the individual encourages us to overlook the fact that human cognition is always situated in a complex sociocultural world and cannot be unaffected by it [...] I hope to show that human cognition is not just influenced by culture and society, but that it is in a very fundamental sense a cultural and social process. To do this I will move the *boundaries of the cognitive unit of analysis* out beyond the skin of the individual person [...][174]

In this chapter I will turn my attention to the final aspect of 4ED, that of distributed cognition. This chapter is guided by the first research question (I) What are the theoretical commitments the different accounts of cognition in 4ED hold on to and how do these relate to each other? Yet, the chapter will specifically be targeted by the second one: II) (Why) should distributed cognition be added with the rest of the 4E accounts? In giving an account of distributed cognition I will mostly draw from the expositions of Edwin Hutchins, Jim Hollan and David Kirsh, since they have been developing and utilising the framework of distributed cognition from the start and still continue to do so. Hutchins traces the roots of distributed cognition to the mid 1980's having theoretical influences in anthropology, sociology and the cognitive sciences. Yet, most notably it is Hutchins' own 1995 book *Cognition in the Wild* that is cited as one of the key texts in setting out the theoretical foundations of distributed cognition. My focus here is on these particular authors because it is their, especially Hutchins', exposition that participates in the current debate and is most discussed by the opponents and proponents of 4E.[175]

This chapter has three main sections. I will start by giving a general account of distributed cognition in 5.2. I will examine the methodology, principles and types of distributed cognition that its theorists take to be most significant. Then in section 5.3. I will compare and contrast distributed cognition to the other 4E approaches and argue for why distributed cognition should be included with the rest of the 4E. My aim is to draw out both similarities as well as differences between the approaches, and in doing so further clarify the theoretical commitments of distributed cognition.

5.1 "Cognition in the wild"

One of the distinguishing features of distributed cognition with reference to the other 4Es is its relationship to anthropology; specifically its methodological approach of *cognitive ethnography* with its emphasis on the study of *cognition in the wild*. The term "cognition in the wild" was coined

174 Hutchins 1995: xiii–xiv emphasis added.

175 Hutchins 2001: 2068. Note that for instance Norman (1993) and Saloman (1993) are cited as other seminal early texts in distributed cognition. See e.g. Hollan et al. (2000) and Hutchins (2001).

by Hutchins in his 1995 similarly titled book. With the term he refers to the idea of how human cognition should be primarily studied in its natural everyday environment and activity, contrary to the "captivity" of normal laboratory settings. In the book Hutchins' subject of ethnographic study is navigation onboard a US Navy ship.[176] The attention in cognitive ethnography is not only in individual cognition but also on the role material, social and environmental factors have in creating meaning and action.[177] Furthermore, the methodological focus is on events as Hollan et al. describe the event-centredness of cognitive ethnography: 'We are interested not only in what people know, but in how they go about using what they know, to do what they do.'[178]

In addition to the methodology of cognitive ethnography distributed cognition maintains two interrelated principles that according to them distinguishes it from other approaches to cognition. The first principle concerns *the boundaries of the unit of analysis for cognition*.[179] I find the following two quotes to illuminate this first principle.

> The central claim of the distributed cognition framework is that the proper unit of analysis for cognition should not be set a priori, but should be responsive to the nature of the phenomena under study.[180]

> Distributed cognition looks for cognitive processes, wherever they may occur, on the basis of the functional relationships of elements that participate together in the process [...] A cognitive process is delimited by the functional relationships among the elements that participate in it, rather than by the spatial colocation of the elements.[181]

So, the boundary of the cognitive unit/phenomenon/process under study is relative to the explanatory needs; sometimes the correct boundary is drawn at the skin and skull of a person, sometimes this is too large a boundary and a more detailed analysis is required, yet most often the skin and skull are too restrictive and a larger boundary is needed.[182] Hence, the important point here is that cognition is not restricted from the outset to the brain leaving anything outside it non-cognitive. It should be conceived more holistically. Also contrary to the (first wave) extended cognition for instance it is not the case that cognition is primarily in the individual's head and then perhaps extends from it to the surrounding world thus forming a hybrid cognitive system. Rather, the initial point of view is a systemic one that does not start from a predetermined location, instead

176 Hutchins 1995: xiii–xiv, 370–371
177 Hollan et al. 2000: 179; Hollan & Hutchins 2010: 243
178 Hollan et al. 2000: 179
179 Hollan et al. 2000: 175–176; Hutchins 2001: 2068; Hollan & Hutchins 2010: 240–241. Subsequent exposition is based on these three articles due to their similar treatment of the subject and will not be referenced further, other than when directly quoting passages.
180 Hutchins 2011a: 426. Recall also Bateson's (1972) similar point about not setting up boundaries *a priori* noted in footnote 74.
181 Hollan et al. 2000: 175
182 Hollan et al. 2000: 175–176; Hutchins 2001: 2068; Hollan & Hutchins 2010: 240–241

one seeks to 'find a system that can dynamically configure itself to bring subsystems into coordination to accomplish various functions.'[183]

So, distributed cognition is specifically interested in conceiving cognition as a holistic phenomenon. Something that is larger than the individual. Another reason behind this is the second principle, which concerns *the range of mechanisms that participate in cognitive processes*. Here distributed cognition departs from the traditional cognitivist explanations that exclusively rely on internal symbol manipulations or neuronal events, and looks for a broader class of cognitive events that may not be encompassed by the boundary of the skin and skull.[184] For instance, Hutchins' study of navigation onboard a US Naval ship speaks for the way the bridge of the ship could be seen as a unit of cognitive analysis. Hutchins maintains for instance that in the task of marking the ships location a great deal of coordinated behaviour and problem solving among different individuals in the navigational team, parts of the ship and specific navigational instruments is required. But, even though this task involves numerous representational subsystems no single subsystem is solely responsible for the representation of the ship's location. Rather, this representation comes as a whole, and is best explained as being realised by the relevant individuals, navigational instruments and their coordinated actions.[185] Chapter 7.5 discusses in more detail a challenge to this type of distributed conception that has been raised by Rupert. There the issue is discussed in terms of cognitive agency and problems with "group cognitive states". The gist of the debate revolves around the possible redundancy of positing group cognitive states over and above the individual states of the particular individuals taking part in the group. But for now I deter the discussion to chapter 7.5.

In applying these two aforementioned principles to cases of *cognition in the wild*, i.e. actual cases of cognitive activity such as navigation, jury decision making, problem-solving etc. Hollan, Hutchins and Kirsh distinguish three kinds of cognitive distribution: 1) socially distributed cognition, 2) coordination of internal and external structures, and 3) distribution in time.[186] I will next treat each of these accordingly.

1) The main idea in socially distributed cognition is that social organisation is itself a cognitive architecture. Hollan et al. argue that this is so because cognitive processes involve the transmission and transformation of information, the patterns of which indicate some underlying cognitive architecture; and since social organisation determines the way information flows in a group of people it itself may be seen as a cognitive architecture. Moreover, distributed cognition includes emergent phenomena that arise not only from the interactions between people, but also from interactions between individuals and the structure of the environment (recall Hutchins'

183 Hollan et al. 2000: 175
184 Hollan et al. 2000: 175–176; Hutchins 2001: 2068; Hollan & Hutchins 2010: 240–241
185 Hutchins 1995: especially ch.3 and 4
186 Hollan et al. 2000: 176–180; Hutchins 2001: 2068–2071; Hollan & Hutchins 2010: 241–244

navigation study above).[187] This leads nicely to the second kind of cognitive distribution.

2) Hollan et al. take the coordination of internal and external structures along the lines of embodied-embedded cognition discussed above in chapter 2.[188] They emphasise causal coupling, the fact that minds are not merely representational engines that create internal models of external world, and how minds are emergent both developmentally and operationally. Hence, for distributed cognition tools and instruments are not merely stimuli but are integrated elements in the larger cognitive system.[189] In favour of this idea Hutchins notes how '[i]t is essential to distinguish the cognitive properties required to manipulate the artifact from the [cognitive property] that is achieved via the manipulation of the artifact.'[190]

3) Once the boundary of the unit of cognitive analysis is moved beyond the skin and skull the close connection between culture and cognition and with it the aspect of cognitive distribution in time are brought forward.[191] The relationship between cognition and culture has an intertwined dual nature: culture is a cognitive process as much as cognition is a cultural process. Hollan et al. describe this relationship the following way:

> This means, on the one hand, that culture emerges out of the activity of human agents in their historical contexts, as mental, material and social structures interact, and on the other hand, that culture in the form of a history of material artifacts and social practices, shapes cognitive processes, particularly cognitive processes that are distributed over agents, artifacts, and environments.[192]

So, the cognitive distribution in time arises from the way culture safeguards accumulated knowledge and acts as a resource for learning, problem-solving and reasoning. This way culture both enables (by allowing the reliance on previous solutions) and constrains (by blinding to other possible nouvelle solutions) the accomplishment of tasks that otherwise would not have been achieved.[193]

Finally, Hutchins has recently started to emphasise and analyse more thoroughly the notion of *cultural practice* and its relevance in shaping cognition.[194] Cultural practices incorporate, or function within, the three types of distribution (social, coordination of internal and external, and distribution in time) and are practices that are constrained and/or coordinated by the practices of

187 Hollan et al. 2000: 176–177; Hutchins 2001: 2069–2070; Hollan & Hutchins 2010: 241–242
188 Specifically they take it as "embodied cognition", but their discussion and references are more in line with embodied-embedded cognition discussed in chapter 2. (Hollan et al. 2000: 177–178) This aspect was present already in Hutchins (1995) even though it was not dubbed embodied cognition.
189 Hollan et al. 2000: 177–178; Hutchins 2001: 2070; Hollan & Hutchins 2010: 241
190 Hutchins 2001: 2070.
191 Hutchins 1995: 354; Hollan et al. 2000: 178; Hollan & Hutchins 2010: 241–242
192 Hollan et al. 2000: 178
193 Hollan et al. 2000: 178; Hutchins 2001: 2070–2071; Hollan & Hutchins 2010: 241–242
194 Hutchins 2008; 2011b

other people. Cultural practices have a dual function of organising human interactions in and with the world:

> […] first by furnishing the world with the cultural artefacts that comprise most of the structure with which we interact. Second, cultural practices orchestrate our interactions with the natural phenomena and cultural artefacts that produce cognitive outcomes.[195]

Hutchins maintains that effectively all external representations, such as language, reading and writing, are products of cultural practices. But cultural practices are not to be understood as static disembodied mental representations of knowledge. Instead they are *embodied skills* that include particular ways of perceiving the world.[196] An illuminating example would be the way of seeing and picking out constellations. Stars may bee seen regardless of ones cultural background, but constellations 'exist only by virtue of someone *enacting* it via a cultural practice that allocates visual attention in a particular way.'[197] Yet, Hutchins goes so far as to argue that we make material patterns into representations by enacting their meanings through cultural practices. These *enacted representations* form the basis for understanding how higher-level cognitive processes emerge from lower-level sensorimotor ones. For instance a member of a navigation team may represent the speed of the ship on a chart by appropriately manipulating a pair of dividers on successively marked positions of the ship on a map and the scale features embedded in the sides of the chart. The crew member is able to perceive/read from the span of the dividers placed on the scale the speed of the ship rather that the equally visible distance covered just because he is trained to make the distinction (i.e. his accumulated cultural practices coordinate the process).[198] I take the point here to lie in Hutchins' recognition that perception is never passive reception of input but is always enacted with meaning by the organism's cultural practices, and that sensorimotor activity is necessarily part of the reasoning processes; '[…] cultural practices orchestrate the coordination of low-level perceptual and motor processes with cultural materials to produce particular higher-level cognitive processes.'[199]

Importantly with regards to cognitive agency to be discussed in chapter 7 Hutchins argues that one must be careful with attributing cognitive properties/processes to individuals that are involved in cultural practices since '[t]here is a danger of attributing to the individual cognitive properties that belong to the larger distributed system.'[200] These ideas about cultural practices and larger distributed systems as agents also raise an interesting connection to the philosophy of social

195 Hutchins 2008: 2018
196 Hutchins 2008: 2012; 2011b: 441
197 Hutchins 2011b: 441 emphasis added.
198 Hutchins 2011a: 427–435
199 Hutchins 2011a: 434
200 Hutchins 2008: 2011

sciences, and philosophy of action in particular. The work done by Margaret Gilbert, Michael Bratman, Raimo Tuomela and David Velleman for instance could be highly interesting here.[201] Moreover the work of Deborah Tollefsen that approaches these sort of issues from the perspective of philosophy of action along similar lines as the authors just mentioned but with a connection to first-wave extended cognition would also be very interesting. Yet, unfortunately due to length constraints these sort of points of view will not be considered in this thesis. As has already been mentioned, chapter 7 touches upon the issues of cognitive (and collective) agency but from a different perspective to the philosophy of action theorists.

5.2 Why Distributed Cognition is Included With the 4Es[202]

At this point I think it is crucial to say something about why I have included distributed cognition with the rest of the 4Es. This is because no one in the 4E(D) literature has explicitly made this sort of grouping before, even when distributed cognition and namely Hutchins' theorising is alluded to and made use of. As was noted in chapter 1 the focus has largely been only in the 4Es with varying emphasises. In what follows I will draw four reasons for the inclusion of distributed cognition to the wider 4ED approach to cognition. The relevance of this issue stems from the consequences of bringing distributed cognition into the debate. For instance it opens up new ways of conceiving matters to do with agency and dynamic nature of cognitive phenomena that might have been neglected by the other 4E approaches.

First, like the Es discussed above distributed cognition is often equated or mixed with the other views on cognition without drawing much attention to its distinct theoretical features. In particular distributed cognition is often lumped together with extended cognition in its general form.[203] This has the unwarranted consequences of possibly dismissing distributed cognition for reasons that might not be directly applicable to it; as well as overlooking its unique theoretical features that could benefit and contribute to the overall debate concerning the nature of cognition. Thus the debate could inadvertently be impoverished.

Second, distributed cognition does have a lot of similarities with the other E's. This might be the root cause for lumping distributed cognition with them. To start with, the embodied-embedded perspective is evident in the second kind of distribution of cognition discussed in the previous section. Also the enactive perspective is present in the way distributed cognition emphasises the relevance of culture for cognition and its dynamic nature.[204] Interestingly, Hutchins

201 Gilbert 1990, 2000, 2004, 2009; Bratman 1993, 1999; Tuomela 1977, 2003; Velleman 1997

202 I would like to thank my supervisor Sami Paavola for the initial nudge towards looking into the possible relevance of distributed cognition in the context of the other 4Es.

203 See e.g. Adams & Aizawa 2008; Chemero & Silberstein 2008; Menary 2009, 2010a, 2010c; Sutton 2006; Giere 2007.

204 Hollan & Hutchins 2010: 242 referencing Thompson (2007).

has recently analysed and applied both the embodied-embedded and enactive perspectives to his distributed cognition view. The general theme here is seeing cognition essentially as a biological phenomenon rather than as a uniquely logical process. That is, to get rid of the cognitivist sandwich model of cognition that divorces cognition, perception and action in favour of a more unificatory framework. On the one hand Hutchins emphasises the interaction of body and world, the role of motor activity in reasoning processes, and the connection between perception and action. On the other hand Hutchins relies on the enactivist ideas about how experiences, sensations and perceptions are not merely received but are dependent on the particular actions of a given organism.[205] This aspect is most evident in the discussion about cultural practices presented at the end of the previous section. Finally I would like to raise few issues that speak for the similarities between distributed and extended cognition. One clear similarity is the general tendency to brake the boundaries of cognition and credit more role for the external world in influencing cognition, e.g. by giving a more full-fledged cognitive status for the external scaffolding structures, physical tools and instruments in particular, in the overall cognitive system. Also the emphasis on learning history, cultural acclimatisation and complementariness of the different resources in the second-wave theorising bears close resemblance to the theoretical features of distributed cognition.

Third, it is important to note that distributed cognition is indeed a distinct approach to cognition, with its own tradition and theoretical features. Firstly, there is a quite established research methodology that is based on the framework of distributed cognition. This is most evident in for example "human computer interaction" studies.[206] Yet, a key distinctive feature is the reliance on anthropological considerations and the method of cognitive ethnography. This brings about the methodological focus on *events*, which could be contrasted for instance to Clark's static view of the environment and culture.[207] Moreover, I take one of the crucial and unique features of distributed cognition to be its point of view to the study of cognition. With this I mean the tendency in distributed cognition not to "out the mind" like e.g. first-wave extended cognition purports to do, rather the point of view is from the start a more wider and systemic one. When assessing a given cognitive task distributed cognition seeks to encompass the whole relevant cognitive ecology (cultural, artifactual, historical etc.) in the explanation. One simply cannot look at individuals or cognition in the absence of the wider social, environmental, cultural context. So, in a way distributed cognition does not privilege the brain but rather allocates a proper place for it in the overall cognitive ecology. The last distinctive feature of distributed cognition is its take on *agency*. I will not examine this further here but will return to it in chapter 7 since the question of agency draws interesting differences within the 4ED in general.

The fourth reason for why I deem it necessary to included distributed cognition with the

205 Hutchins 2011a: 427–435

206 Hollan et al. 2000; Hollan & Hutchins 2010. But note that enactive cognition has similar ambitions and there is emerging a research methodology for cognitive science based on enactivism, see e.g. Stewart et al. (eds.) (2011).

207 Hollan et al. 2000: 179–180; Hollan & Hutchins 2010: 242–244; Hutchins 2011b; Clark 2008a.

other 4Es is its growing relevance in the debate about cognition. For instance Hutchins[208] has recently participated in the discussion through the aforementioned issue about agency and the apparent static nature of culture and environment. It is worth noting that Hutchins' account of agency is due to his reaction against Clark's account of agency that he saw as inadequate and mistaken. Others have also recognised the appeal to distributed cognition in these issues.[209] Yet, in my opinion the relevance of distributed cognition has risen mainly due to a growing interest in it by the second-wave extended cognition theorists.[210] Sutton with his colleagues for instance maintain that distributed cognition offers a viable methodological take on the study of cognition; one that is compatible with extended cognition and should be played more attention to. Sutton even says that his second-wave approach moves closer to distributed cognition.[211]

5.3 Summary

In this chapter I examined distributed cognition. I started by offering a general account of its theoretical commitments. Distributed cognition emphasises the need to study cognitive phenomena in the natural everyday environment (cognitive ethnography). It holds that the boundaries of the unit of cognitive analysis should not be set a priori but should be responsive to the phenomena under study and their relevant social, cultural, environmental, and historical context. It stresses that the mechanisms involved in cognition should also be sensitive to the wider socio-cultural world. And finally distributed cognition urges one to recognise that there are three kinds of cognitive distribution: social, coordination of internal and external, and distribution in time. The last point emphasises the fact that it is specifically in cultural practices that we organise our cognitive interactions with the world. I then gave four reasons for why distributed cognition should rightly be added to the rest of the 4E approaches to cognition. Firstly, to avoid its unwarranted identification and mix-up with the other views. Secondly, it shares a similar take on cognition as the others. Thirdly, it is nevertheless a distinct approach to cognition with its own important theoretical commitments. That is, fourthly, it is relevant in current debate due to a growing interest in it and its fruitful theoretical features.

208 Hutchins 2011b.

209 E.g. Kiverstein 2011; Kirchhoff 2011.

210 Sutton et al. 2010; Sutton 2006, 2010; Menary 2009, 2010a, b, c. Note that in and between the expositions of these second-wave extended cognition authors there is no clearly analysed and unified take on distributed cognition and its relationship to the other 4Es.

211 Sutton et al. 2010: 523n.1, 531n16; Sutton 2010: 215–216n1.

So, the way we see it, there are two principal features of intracranial processes – their use of non-derived representations governed by idiosyncratic kinds of processes – that serve to distinguish cognitive from non-cognitive processes. These features constitute a "mark of the cognitive" and they provide some non-question-begging reason to think that cognition is intracranial.[212]

The past five chapters have been exclusively focused on the two first research question: I) What are the theoretical commitments the different accounts of cognition in 4ED hold on to and how do these relate to each other? and II) (Why) should distributed cognition be added with the rest of the 4E accounts? This chapter expands the discussion about the theoretical commitments of 4ED accounts by adding also a third research question to the investigation: III) What critique does 4ED face and how does providing a mark of the cognitive affect both the proponents as well as the opponents of 4ED?

I will raise two key interdependent objections and challenges against 4ED: *the coupling-constitution fallacy* and an argument from *mark of the cognitive* considerations. I consider these particular challenges important because they are the most widely discussed and they affect all of the 4ED approaches. More importantly still, the objections and 4ED authors' responses to them raise further questions about the nature of a proper mark of the cognitive and the general need of providing one in the first place. Interestingly these questions impact cognitivism as well, and in a way level the playing field between cognitivism and 4ED. I will largely follow Fred Adams & Ken Aizawa's (A&A hereafter) argumentation when presenting these criticisms even though there are other authors that have raised similar objections.[213] A&A have been most actively criticising the 4ED type approach, and the debate has significantly revolved around their objections. Even though many of these criticisms are most often directed at extended cognition they apply to 4ED in general (in A&A's texts the 4ED positions are all more or less discussed, even if their focus has been on extended cognition).[214]

6.1 The Coupling-Constitution Fallacy and Replies to It

A&A are *contingent intracranialists* and maintain that although logically and nomologically possible, it is a contingent empirical fact that cognition does not actually extend in the way 4ED approaches

212 Adams & Aizawa 2008: 10
213 See for instance: Rupert 2004, 2009a, b, 2010; Shapiro 2010; Block 2005; Prinz 2006
214 See e.g. Adams & Aizawa (2001 and 2008).

claim, not at this time any way.[215] A&A argue that the 4ED theorists confuse causal influences of external resources on cognition, such as coupling, manipulation or integration type dependencies, as constitutive parts of cognitive processes.[216] A&A argue the following way

> The fallacious pattern is to draw attention to cases, real or imagined, in which some object or process is coupled in some fashion to some cognitive agent. From this, one slides to the conclusion that the object or process constitutes part of the agent's cognitive apparatus or cognitive processing.[217]

For instance, it would be wrong to say that a pen and paper constitute a part of one's cognitive system since the relation one has with the pen and paper is that of causal coupling and not of constitution, which is a different type of relation. The pen and paper are nothing more than external aids that an agent may rely on in her intracranial problem-solving tasks. More generally Adams & Aizawa's argument is that

> the fact that object or process *X* is coupled to object or process *Y* does not entail that *X* is part of Y [...] So, if the fact that an object or process *X* is coupled to a cognitive agent does not entail that *X* is a part of the cognitive agent's cognitive apparatus what does? The nature of *X*, of course.[218]

Hence it seems that A&A's main argumentation has two aspects: first teasing out a coupling-constitution fallacy, and then arguing from the specific nature of *X*, i.e. argue for a specific type of *mark of the cognitive*. This argumentative tact is also one of the reasons I focus on A&A's argumentation since they do not commit, what Susan Hurley calls, the *"causal-constitutive error" error*. That is, '[...] the error of objecting that externalist explanations give a constitutive role to external factors that are "merely causal" while assuming without independent argument or criteria that the causal-constitutive distinction coincides with some external-constitutive boundary.'[219] I will return to the mark of the cognitive aspect in section 6.3. below, since it may be seen as a challenge in its own right.

Now, 4ED theorists have responded to the proposed coupling-constitution fallacy. Clark notes that there are some things that A&A have misunderstood about the extended cognition argumentation for instance.[220] First, it is not the case that extended cognition aims to make some object cognitive through appropriate coupling, by for instance fulfilling the three criteria of

215 Adams & Aizawa 2001: 53
216 Adams & Aizawa 2001: 46–57; 2008: 88–91
217 Adams & Aizawa 2010a: 68
218 Adams & Aizawa 2010a: 68
219 Hurley 2010: 106
220 For the sake of simplicity and brevity I will only rehearse the responses offered by extended cognition theorists, but note that other 4ED theorists could also adopt similar lines of responses.

reliability, accessibility and automatic endorsement. Clark argues that A&A have misunderstood the point about cognitive extension: it is not that some external object is cognitive in itself, rather the appeal to coupling

> is intended to make some object, which in and of itself is not usefully (perhaps not even intelligibly) thought of as *either cognitive* or *non-cognitive*, into a *proper part of some cognitive system*, such as a human agent [...] It is intended [...] to ensure that the putative part is poised to play the kind of role that *itself* ensures its status as part of the agent's cognitive routines.[221]

So, the original point in C&C's hypothesis was to see what it would take for an external resource to be incorporated (as a proper part) into a cognitive system and not what would make an external resource cognitive in and of itself.

Secondly, Clark does recognise that A&A do also understand the incorporative claim. But he argues that this leads to a more complex misunderstanding that feeds directly into their argumentation about the *mark of the cognitive*. Specifically, A&A's last point about the nature of X, quoted above, leads Clark to argue that what in fact A&A are saying is that it is in virtue of the intrinsic nature of some objects and processes that they may be considered '*candidate parts* (for inclusion in a cognitive process), whereas other objects or processes, still in virtue of their own nature, are not.'[222] Hence, it seems that the issue of the mark of the cognitive is the more pressing one here in demarcating the (possibly) cognitive from the (possibly) noncognitive than the alleged coupling-constitution fallacy.[223] I will return shortly to this.

Now, Menary argues quite similarly to Clark and maintains that it is not the case that extended cognition purports to make some external artefact cognitive in itself. Menary argues that A&A's picture is flawed and rests on a misunderstanding of extended cognition:

> If we accept the picture of a cognitive agent as implementing *a discrete cognitive system, before they ever encounter an external vehicle*, then we will have accepted the very picture of cognition we set out to reject.[224]

Instead, for Menary *X* is the agent's manipulation of the external vehicle coupled to the agent's internal brain processes *Y* that *together constitute* a cognitive process *Z*. Moreover, because *Z* is a hybrid process it may not be fully described by either *X* or *Y* alone.[225] In relation to extended cognition Menary argues that:

---

221 Clark 2010b: 83 original emphasis.
222 Clark 2010b: 84–85
223 Clark 2010b: 82–85; Clark 2008a: 86–89
224 Menary 2006a: 333 emphasis added
225 Menary 2006a: 333–334

> The aim is not to show that artifacts get to be part of cognition just because they are causally coupled to a pre-existing cognitive agent, but to explain why *X* and *Y* are so coordinated that they together function as *Z*, which causes further behaviour.[226]

Hence, what matters is not the artefact or the coupling *per se*, but the appropriate coordinated manipulation and integration of the external element.

6.2 Mark of The Cognitive – Motley Crew and Nonderived Content

It seems clear from the responses to the coupling-constitution fallacy that the most stressing issue in A&A's argumentation is in providing an account of the nature of *X* (i.e. the purported cognitive state/process). This aspect of the argument is quoted at the beginning of this chapter and is based on distinguishing a *mark of the cognitive*, i.e. some necessary conditions for a state or process to be considered cognitive.[227] I take there to be especially two reasons for why the mark of the cognitive argument is important and interesting. a) As Rowlands argues the coupling-constitution fallacy (among few other objections) is derivative of distinguishing a mark of the cognitive and hence the fallacy ultimately collapses to the need of providing one.[228] b) Sven Walter and Lena Kästner argue that *cognitive agnosticism*, that is the absence of an account of *what* cognition is, is unsustainable and needs to be addressed before one can question *where* cognition is.[229] In itself the mark of the cognitive for A&A has a dual aspect since for them 'cognition involves *particular kinds of processes* involving *non-derived representations*.'[230] In what follows I will treat each of these elements accordingly: 1) starting with the requirement for "particular kinds of processes" and 2) then following up with the notion of nonderived content.

1) In section 4.2 I alluded to a stalemate that arises between first-wave extended cognition and its opponents that was due to describing the particular kinds of processes underlying cognition. Now, there is another related problem for 4ED in general that pertains to the nature of the processes involved in cognition. A&A insists that 'the cognitive [like other natural domains]

226 Menary 2006a: 334
227 Adams & Aizawa 2001: 48
228 Rowlands 2009a, c, 2010.
229 Walter & Kästner 2012; Walter 2010
230 Adams & Aizawa 2001: 52–53 emphasis added. Other names for non-derived representations include: *original intentionality* (Kiverstein & Clark 2009), *intrinsic content* (Clark 2005b; Menary 2006a), *intrinsic representation* (Adams & Aizawa 2010a) and *non-derived content* (Clark 2010b). I will not take issue here whether there might be some radical conceptual differences between these terms, I will simply take them all to refer to the same thing: henceforth *nonderived content*.

must be discriminated on the basis of underlying causal processes.'[231] A&A argue that the underlying causal processes involved in 4ED approaches to cognition are vastly heterogeneous and feature such a motley that fails to pick out any scientific *natural kind.* And since 4ED approaches fail to exhibit natural kinds they are no science at all and therefore should be abandoned as implausible approaches to the cognitive.[232]

> There are no laws covering humans and their tool use over and above the laws of intracranial human cognition and the laws of the physical tools [...] There just isn't going to be a science covering the motley collection of [for example] "memory" processes found in human tool use.[233] [A] would-be brain-tool science would have to cover too broad a collection of processes.[234]

A&A maintain that as it happens only an *intracranial* science of cognitive processes will be plausible since it has for example already discovered and explained important law-like regularities of primacy, recency, and chunking -effects of human memory.[235]

2) A&A take *nonderived content* to be another essential feature that a cognitive state or process must involve. Specifically A&A are adamant that the claim is not that cognitive processes must *entirely* consists of nonderived content, but that there just needs to be some: 'if you have a process that involves no intrinsic content, then the condition rules that the process is noncognitive.'[236] A&A maintain that something has nonderived content when it means what it does independently of any other representations or contents. That is, its meaning is not *derived* from or dependent upon other 'content-bearing, representational or intentional states.'[237] For instance, numbers and words bear derived content because their meaning is fixed by intentional agents, social practices and/or conventions. A&A take thoughts, experiences and perceptions as paradigm cases that bear nonderived content just because their meaning is not fixed by other things.[238]

Specifically A&A argue that processes get their nonderived content by fulfilling some naturalised conditions. Here they appeal to philosophers such as Dretske, Fodor, Cummins, Searle and Millikan who aim to give a naturalised account of meaning. That is, how thoughts come to have meaning from the non-contentful states and objects in the world. Yet A&A are not committed to any

231 Adams & Aizawa 2001: 52

232 Adams & Aizawa 2001: 51–52, 61–62. Rupert (2004) makes a similar argument. This is why this argument has been dubbed "The Motley Crew Argument". See Shapiro (2010) and Walter & Kästner (2012) originally stemming from the argumentation of Adams & Aizawa (2001, 2010a) and Rupert (2004, 2009a).

233 Adams & Aizawa 2001: 61

234 Adams & Aizawa 2010a: 76

235 Adams & Aizawa 2001: 61–62

236 Adams & Aizawa 2010a: 70. A&A are adamant about making this clarification due to the debate between them and Clark (2005b, 2010a), where they take Clark to misunderstand them as making the stronger claim that *all* cognitive states and processes must involve nonderived content. I will return to Clark's exposition shortly. Note that Fodor (2009) does make the stronger claim.

237 Adams & Aizawa 2001: 50; 2009: 87

238 Adams & Aizawa 2008: 32, 35; Adams & Aizawa 2001: 48

particular view of natural content determination, just that it needs to be natural so as to get nonderived content.[239]

A&A's argument from nonderived content is that since external elements carry only derived content they lack the mark of the cognitive and hence are excluded from being part of cognition. One might worry here that A&A's argument is too loose and that there is still room for external elements in cognition since A&A maintain that not all cognitive processes must bear nonderived content. But this worry should fade away with Aizawa's further clarification: 'The idea is [specifically] that *cognitive vehicles of content* must bear non-derived content.'[240] So, the external elements remain only as aids and tools for internal cognitive processing and cognition does not extend beyond the organism.[241]

6.2.1 Reply to the Motley Crew Argument

The reply to A&A's Motley Crew Argument has two stages. First, the aim is to try and retain the notion of natural kindhood, albeit it is conceived differently, and yet show how it does not necessarily affect the scientific status of 4ED. Second, the aim is to put into question the requirement for *natural* kindhood, and show that kindhood *simpliciter* is enough for scientific enterprises (4ED included).

The first reply to the motley crew is to hold on to the notion of natural kindhood and yet maintain that it does not affect the scientific status of 4ED. This reply[242] is founded on the idea that natural kindhood need not (or should not) be conceived in *essentialist* terms, whereby natural kinds are defined by some intrinsic, necessary and sufficient conditions, that correspond to laws of nature. This is just the sort of conception of natural kindhood A&A champion.[243] One of the reasons for the wariness of this sort of essentialist conception of natural kinds is that it has run into difficulties in describing phenomena in the life sciences; e.g. in contemporary biology and evolutionary theory.[244] Marc Ereshefsky for instance notes how a tenet of essentialism that requires all and only the members of a kind to have a common essence faces difficulties in biology; especially with regards to the notion of species. For it is very difficult to find biological traits that are found in all and only the members of a given species since biological traits are rarely unique to a species and they may disappear from a species in the course of evolution.[245] Thus Samuli Pöyhönen notes how nowadays 'in the special sciences, scientifically interesting kinds of phenomena are often

239 Adams & Aizawa 2008: 35–39; Aizawa & Adams 2005: 662–664; Adams & Aizawa 2010b: 588
240 Aizawa 2010: 337n10 emphasis added.
241 Adams & Aizawa 2001: 53–57
242 This issue was brought to my attention by the work of Samuli Pöyhönen (2012).
243 Bird & Tobin 2010: 1; Walter & Kästner 2012: 18–19; Pöyhönen 2012: 7
244 Pöyhönen 2012: 7–8; Ereshefsky 2010: 2.1
245 Ereshefsky 2010: 2.1. See also Bird & Tobin 2010: 2.1.2.

characterized by relational properties, and do not correspond to laws of nature'.[246]

Pöyhönen presents Richard Boyd's *homeostatic property cluster* theory as an alternative approach of natural kinds, and then applies it to explaining cognitive phenomena and the motley crew argument.[247] Now, Boyd argues that instead of aiming to pinpoint a subset of some necessary and sufficient conditions, natural kinds should be thought of as denoting a larger set of correlated properties. Moreover, the unity of the cluster of these properties is causal rather than conceptual in nature.[248] Specifically these natural *homeostatic property cluster kinds* are determined by 'the members of a cluster of often co-occurring properties and by the ("homeostatic" [causal]) mechanisms that bring about their co-occurrence.'[249] Hence, Pöyhönen applies this definition of homeostatic property cluster kinds to cognitive phenomena whereby cognitive natural kinds could be described as

> consisting of [...] the cluster of observable properties characteristic of the phenomenon and a corresponding [...] cognitive mechanism that guarantees the reliable co-occurrence of properties.[250]

He then goes on to note how more and more psychological phenomena, such as social emotions and psychiatric disorders, are explained in terms of hybrid mechanisms that are not solely intracranial and for instance include social causal factors. Therefore, homeostatic property cluster theory's take on natural kinds seems to be in line with the explanations of cognitive phenomena the 4ED upholds. Moreover, from the point of view of the homeostatic property cluster theory the force of the motley crew seems to fall through. Recall for instance how second-wave extended cognition and distributed cognition do not require cognitive mechanisms (whether internal or external) to be causally similar. Thus, as a cursory analysis of natural kindhood it seems that the scientific status of 4ED may be retained by conceiving natural kindhood along the lines of homeostatic property cluster theory.[251]

Yet, this recourse to the homeostatic property cluster theory might not be necessary as the second reply to the motley crew challenges A&A's reliance on *natural* kindhood argumentation. As a preliminary analysis of the motley crew Sven Walter and Lena Kästner raise a few important issues. First, the motley crew conceives proponents of 4ED to hold "cognition" as a kind of '*umbrella concept*, i.e. a concept subsuming a diversity of processes that share only the fact that they fall

246 Pöyhönen 2012: 7–8. See also Bird & Tobin 2010: 2.1.2 for similar point,
247 Pöyhönen 2012: 7–8
248 Boyd 1991: 141–142; 1999:141–144; Pöyhönen 2012: 8
249 Boyd 1991: 141
250 Pöyhönen 2012: 8
251 Pöyhönen 2012: 8; Boyd 1991: 141–142; 1999:141–144

under this concept'[252] comparable to concepts like "things in this library" or "things that play music". Second, one should distinguish the claim within the motley crew argument that "cognition" is not a *natural* kind concept from the claim that "cognition" is an umbrella concept, since one could hold that cognitive processes are non-natural kinds and yet deny that cognition is an umbrella concept.[253]

Walter & Kästner then object to A&A's insistence that science is only interested in *natural* kindhood, for it could be argued that what is important is kindhood *simpliciter*. They cite, among others, the study of superconductors and oscillators in physics and the study of money in economics as examples of scientific enterprises where the objects of study are not *natural* kinds. Hence, Walter & Kästner argue the following way:

> The mere fact that extended processes involve artefacts and so fail to constitute a *natural* kind does therefore not entail that there could be no science of the cognitive if cognitive processes were extended.[254]

Moreover cognition may involve heterogenous lower-level processes that nevertheless exhibit enough homogeneity at a higher-level so that they could fall under a meaningful scientific study. Hence "cognition" would be more like "money" or "superconductors" than natural kind concepts such as "helium" or umbrella concepts like "things in this library".[255]

Walter & Kästner sketch three available accounts for a concept of cognition, whereby it is not a *full-fledged* natural kind concept, nor an umbrella concept, but nevertheless picks up a class of things with higher-level unity. These are "cognition" as a *cluster concept, nominal kind* and/or *family concept*. Walter & Kästner take Boyd's homeostatic property cluster theory to fit the cluster concept they are advancing and hence I will not rehearse it here.[256] Now, nominal kinds do not provide a scientifically discoverable essence, but a 'definition that specifies form, function, or origin of the concept's referents',[257] i.e. a nominal essence. Cognitive processes form a nominal kind when their 'form, function, or origin that is used to characterize their nominal essence *supplies them with enough higher-level unification*.'[258] Walter & Kästner take Rowlands' account of cognition as an example of a nominal kind view where extended processes count as cognitive. Rowlands for

252 Walter & Kästner 2012: 19
253 Walter & Kästner 2012: 19
254 Walter & Kästner 2012: 20
255 Walter & Kästner 2012: 20
256 Walter & Kästner 2012: 20–21n.26. Note that Boyd does argue for his view to be a form of *natural* kindhood, even if not being of the essentialist type. Hence I take Walter & Kästner's qualification of "not full-fledged natural kindhood" and insistence on "kindhood *simpliciter*" to allude to other than essentialist forms of natural kindhood, or lack of naturalness thereof. Be as it may, Boyd's account nevertheless does offer, as has been argued above, an alternative conception to natural kindhood that fits 4ED and responds to the motley crew.
257 Walter & Kästner 2012: 20
258 Walter & Kästner 2012: 20 emphasis added.

instance maintains that cognition involves information processing, as the manipulation and transformation of information-bearing structures, and that this information processing

> has the *proper function of making available* either to the *subject* or to *subsequent processing operations* information that was (or would have been) prior to (or without) this processing, unavailable.[259]

Finally, the notion of a family concept originates from Wittgenstein[260] who argued that for example the conditions for something to count as a "game" include 'a complicated network of similarities overlapping and criss-crossing: sometimes overall similarities, sometimes similarities of detail.'[261] Yet, there are no common aspects exclusively in all games. Walter & Kästner then apply this analysis to cognitive processes so as to argue that cognitive processes may resemble one another in important ways even though there is no common feature to all and only cognitive processes.[262]

I take the importance of Walter & Kästner's exposition to lie not in any of the particular alternatives, but in the general tact of providing alternatives to natural kind conceptualisation. Moreover, if Walter & Kästner's argumentation holds and cognition is not necessarily a full-fledged natural kind concept that A&A require it to be, then

> [...] a mark of the cognitive would not be spelled out in terms of lower level communalities among all and only cognitive processes but by appeal to higher level causal characteristic [...] The Motley Crew Argument would thus not only have to claim that extended processes are not natural kinds, but that they lack even the sort of 'higher-level unification despite mechanistic dissimilarities' (Clark 2008a: 96) we are familiar with from [superconductors, oscillators and money].[263]

I take this implication from Walter & Kästner's exposition to be important and lend credence to Clark's argumentation.

Clark's reply to the motley crew has two facets. First Clark somewhat diverts the discussion from the requirement of defining cognition in terms of the nature of its underlying causal processes. Instead, he paraphrases Daniel Dennett and insists that 'cognition is as cognition does'.[264] Focus here is on individuating cognition according to its characteristic effects rather than

259 Rowlands 2009a: 8 original emphasis. Note that this is only a part of the conditions of Rowland's mark of the cognitive.
260 Wittgenstein 1953
261 Wittgenstein 1953: §66, partly cited in Walter & Kästner 2012: 21
262 Walter & Kästner 2012: 21
263 Walter & Kästner 2012: 20
264 Clark 2010b: 93

its causes. Hence, what matters is the way cognitive processes are individuated according to the way they support certain kinds of higher-level behaviours.[265] I take this view to fit well with 4ED in general, where the emphasis is often on the interplay of different resources relevant for the fulfilment of a given cognitive task. Clark does not develop this approach much further, but offers it merely as a possibly viable alternative way of conceiving the matter.

The second aspect in Clark's reply also emphasises the focus on larger systemic wholes over the nature of individual processes and parts. He argues that it seems quite possible that even the inner processes that A&A rely on will turn out to involve heterogenous underlying causal processes as well.[266] Moreover, Clark maintains that what ultimately matters is

> (1) the degree of complementarity (between the different contributions) and (2) the degree of integration achieved. Given sufficient complementarity and integration, it becomes plausible (many of us believe) to treat the resultant system as a cognitive whole, with cognitive properties that supervene on more than the biological components alone.[267]

Unfortunately Clark does not develop these ideas much further either, but instead maintains that for a proper systems-level cognitive science one needs to understand both the details of lower-level parts and processes as well as the higher-level coupling of the biological and the artifactual.[268] He looks to the future and says that

> it is the substantive empirical bet of the extended systems theorist that the larger hybrid wholes, comprising biological and nonbiological elements, will *also* [...] prove to be the proper objects of sustained scientific study in their own right.[269]

In support of this Clark notes how even in neuroscience one does not simply study the neural substructures but also look at dynamics of whole processing cycles with soft-assembled neural populations. Clark then asks why we should simply suppose that all the cognitively relevant soft assemblies are intracranial; perhaps we should therefore include our rich environment into the complex cognitive architecture.[270] Moreover, Clark cites a debate between Wayne Gray and his

265 Clark 2010b: 93
266 Clark 2008a: 94–96
267 Clark 2010b: 93
268 Clark 2010b: 92–95. Even if Clark does not develop this further here one could apply the second-wave theorists' analysis of cognitive integration and complementarity in order to support his reply.
269 Clark 2008a: 115 original emphasis.
270 Clark 2008a: 116

colleagues'[271] "time-cost-based" model and Ballard et al.'s[272] "Minimal Memory Model" of cognitive task performance as evidence of standard forms of cognitive scientific investigation that are targeting genuinely (extended) hybrid wholes.[273] I will return to this discussion in section 7.2 where Clark's view on cognitive agency is presented. Recall also that the methodology and framework of distributed cognition are already been used in human-computer interaction research for instance, and that there is being developed cognitive science based on the framework of enactivism.[274] So one might not need to look too far a future to see scientific investigations that target hybrid wholes.

6.2.2 Deflating the Notion of Nonderived Content

Originally Clark argued that nonderived content cannot be a mark of the cognitive since cognition (even if purely intracranial as A&A insist) necessarily involves derived content as well. Clark argued that Venn diagrams for instance, i.e. two overlapping circles that have conventionally determined set-theoretic meaning (i.e. derived content) concerning the overlapped area, may figure into ones "internal" problem-solving activity. And if one can have cognitive processes that bear derived content, then not all cognitive processes bear nonderived content and therefore A&A's condition for the mark of the cognitive fails.[275] But as was noted in 6.3 A&A purported to clarify their argumentation to counter this objection. They maintained that this reading was too strict and that their argument insisted that only *some* processes with nonderived content are necessary and that their nonderived content condition applied only to cognitive vehicles of content and that other processes may well bear derived content.[276]

It is unclear how successful this A&A's proposed "clarification" is since it still seems to be compatible with the view that an extended (or hybrid) process may include nonderived content in its inner aspects and derived content in its external aspects.[277] Clark gets at this point nicely. He first notes that on A&A's revised condition it is acceptable to have an intracranial process that has a part that bears derived content so long as some other part of that process bears nonderived content for that process to be genuinely cognitive. Moreover an extended (or hybrid) process must fail this weakened condition *by involving no nonderived content at all.*[278]

Clark then asks in what sense this holds for the external elements of extended cognitive

271 Gray & Fu 2004; Gray & Veksler 2005; Gray et al. 2006
272 Ballard et al. 1995; Hayhoe 2000
273 Clark 2008a: 118–122
274 See e.g. Stewart et al. (eds.) (2011)
275 Clark 2010a: 48–49; Clark, 2005b, p.5
276 See Aizawa (2010: 337n10), Adams & Aizawa (2010a: 70).
277 Similar point is also made by Walter & Kästner (2012: 19n20).
278 Clark 2010b: 87–88

processes. He argues that the external element seems most clearly to be part of some genuine cognitive process when it is retrieved and used. And at that time there also seems to be plenty of other states and process in play that bear nonderived content. 'At run time, the process is not one that trades solely in representations whose contents are derived or conventional determined.'[279] At other times, it is the functional poise of the external element to be, for instance, part of the supervenience base for some of the cognitive states and processes of the larger system that allows it to be called upon at relevant run-time moments. Hence the external element could be seen as a resource to inform and affect the agent's behaviour at appropriate run-time moments. Clark insists again that the external part is not to be thought of as "intrinsically cognitive" all by itself.[280] He is also ready to accept, this could be just for the sake of the argument, that perhaps it is crucial that cognitive agents in their cognitive activity rely on some states with nonderived content.

Be as this may Clark nevertheless continues by arguing that:

> we have been given no reason at all to accept the further (and crucial) claim that *no proper part* of such properly cognitive system, considered now in splendid isolation from those crucial run-time wholes in which it participates, can afford to contain only representations lacking intrinsic content.[281]

Here it seems that Clark aims at fleshing out the idea that even when the external element is not appealed to, the cognitive system (i.e. the agent + the external resource) cannot be required to bear only states with nonderived content and hence be considered as noncognitive. It also seems to imply that an agent's processes become noncognitive as soon as an external resource is utilised. Even if it is the case that the external part bears only derived content why should this affect the cognitive status of the whole system? Relatedly, Clark argues that A&A fall again into making the mistake of requiring the external part to be cognitive in itself when the crucial point of (first-wave) extended cognition lies in the appropriate coupling and functional poise of the different elements. Similarly, in 4ED in general the focus is on the nature of the relation, interaction, manipulation, integration etc. between the elements and not in the specific intrinsic nature of the elements.[282] Interestingly this is even something A&A themselves recognise as a mistake relating to the proposed coupling-constitution fallacy: 'It does not follow from the fact that one has an X system that every component of the system does X.'[283] Thus, Clark finishes:

279 Clark 2010b: 88
280 Clark 2010b: 88–89
281 Clark 2010b: 89 original emphasis.
282 Clark 2010b: 89–90.
283 Adams & Aizawa 2009: 84, cited in Clark 2010b: 90

> Agreed, the mere fact that the [external element] and [the agent] "form a system" establishes nothing [...] But the way to then proceed is surely *not* by asking, of the candidate part, whether it somehow "possesses" the characteristic that we now want to ascribe to the resultant overall system.[284]

Now, Menary's response to the nonderived content condition is somewhat different to Clark's. Menary's presents a three-fold argument in order to deflate the notion of nonderived content A&A are utilising: 1) he collapses the distinction between derived and nonderived content; 2) he undermines the strong reliance on naturalistic accounts of content determination; 3) he argues that cognitive scientists do not make (or need to make) use of the notion of nonderived content.[285] A&A maintain the following view on content in their recent reply to Clark's and Menary's[286] objections to their nonderived content condition:

> A thought might bear the content that the cat is hungry in virtue of satisfying some conditions on non-derived content, whereas a particular inscription on a piece of paper *might bear that same content* by satisfying some other conditions on derived content. To put the matter another way, there are two questions one might ask of a representation. The first is what content that representation bears; the second is what conditions make it the case that it bears that content.[287]

Even though this passage is not directly quoted in Menary's exposition I take it to be underlying his argumentation. Menary notes that since it seems that for A&A there must be a difference between derived and nonderived content for their argument to work, and that since it seems that there is not a difference in the content itself (see the quoted passage above) the difference must lie in the method of content determination.[288]

But let's start with 1). Menary notes that there looms a contradiction in A&A's original take on content and this more recent one. Recall the point made in section 6.3 that A&A maintain that words whether spoken or written have derived content and that thoughts are paradigm cases of nonderived content. Yet, in this light the more recent claim that derived and nonderived representations may have same contents seems puzzling. On the one hand concerning the more recent claim Menary asks what makes it the case that for instance the thought "the cat is hungry" is cognitive whereas the inscription with the same content "the cat is hungry" is not, barring the mere question begging stipulation that what is inside the head is cognitive and what is outside of it is not?

---

284 Clark 2010b: 90 original emphasis.
285 Menary 2010c
286 Clark (2005b, 2007, 2008a, 2010b) and Menary (2006a).
287 Adams & Aizawa 2010b: 582 emphasis added.
288 Menary 2010c: 612

On the other hand concerning the original claim there is an even more puzzling implication that our conscious linguistic thoughts are not cognitive at all since they necessarily involve only derived content in the form of natural language. Hence, Menary maintains that the claim that derived content cannot be cognitive content is untenable and should be abandoned.[289]

2) Recall that A&A are not committed to any particular naturalistic theory of content determination, but merely the importance lies in the requirement on a *natural determination*, so as to get at the nonderived content. Now, Menary argues that one should be wary of relying on naturalistic accounts of content determination since none of them succeeds in giving an account 'of how richly conceptual cognitive representations can have the richly conceptual contents they do.'[290] Moreover he continues that if 'cognitive semantics is to be explained, in part, in terms of the semantics of public representations then Adams and Aizawa's stipulation fails.'[291] Menary's next step is to gesture at a this sort of explanation. He argues that the propositional contents of thoughts may partly stem from the propositional contents of the natural language sentences they express. For instance, the meaning of π could come from its publicly/conventionally determined meaning: 'π is the ratio of circle's circumference to its diameter (and which is the same as the ratio of a circle's area to the square of its radius).'[292] Hence, if this holds then A&A's strong reliance on natural accounts of content determination is undermined.

3) Finally Menary argues that cognitive scientists do no make use of the notion of nonderived content. Here Menary relies on the work done by Dehaene et al. presented in section 4.3.2. He argues that since Dehaene et al. (who are cognitive (neuro)scientists) are studying phenomena that would not be counted as cognitive if A&A's condition was true, the condition should be abandoned since it does not fit the phenomena Dehaene et al. are discovering.[293]

6.3 Summary

In this chapter I examined the challenges that have been raised against 4ED. I focused on the coupling-constitution fallacy and the argument from mark of the cognitive. I argued that the more significant challenge was the mark of the cognitive objection. I also examined the responses different authors had offered to these challenges.

The abundance of varied and yet valid responses speak for the failure of A&A's arguments. Yet even if it seems that the challenges that A&A have raised fail on multiple accounts this overall

289 Menary 2010c: 612–613; Menary 2009: 40
290 Menary 2010c: 615
291 Menary 2010c: 615
292 Menary 2010c: 615
293 Menary 2010c: 616

debate nevertheless brings forth interesting questions with regards to the nature of cognition. Especially in relation to the third research question concerning the mark of the cognitive. It seems that no universally agreed upon or viable mark of the cognitive exists between the opponents and proponents of 4ED. Since this seems to be a matter that cuts both ways, it in itself speaks for the viability of questioning the more orthodox cognitivist position. It at least levels the ground for the more unorthodox 4ED approach. The debate also raises the more fundamental question pertaining to the relevance of providing a mark of the cognitive. Is it necessary to provide such a mark in the first place? As was noted in section 6.3.1 Walter & Kästner argue that providing one should be crucial since one cannot ask *where* cognition is before knowing *what* it is. Yet, an alternative view would be to continue as it were and draw the boundaries of the cognitive unit of analysis according to the phenomena under study. The point being not to restrict the boundaries *a priori*, but let scientific practice and the phenomena under study dictate them.[294] Alternatively one could also draw the boundaries according to epistemic and explanatory power considerations, as Pöyhönen has recently argued.[295]

294 See e.g. Hurley 2010: 106–107. Note also that this is a view that distributed cognition theorists hold.
295 Pöyhönen 2012.

7 Agency in 4ED

In this part of the thesis I will draw upon the previous chapters and aim to provide an analysis of 4ED in the context of *cognitive agency*.[296] I deem the question of cognitive agency to be very relevant and illuminating both because it highlights the overlap and differences in 4ED, and because the playing field between the orthodox cognitivism and the more unorthodox 4ED seems to be levelled (due to issues to do with the mark of the cognitive) thus opening up the analysis of cognitive agency. So, this chapter continues the analysis of 4ED under the guidance of the first research question: I) What are the theoretical commitments the different accounts of cognition in 4ED hold on to and how do these relate to each other? Yet, more specifically this chapter will be directed by the fourth research question: IV) How does the notion of cognitive agency figure in the 4ED approach(es) to cognition? The driving issue behind this relates to whom, or what kind of elements, are cognitive properties attributed and what/who is the *author* or of these properties and processes?[297] In what follows I will go through each theory's take on agency in the order mirroring the structure of this thesis.

7.1 Embodied-Embedded Agency

According to the embodied-embedded approach the cognitive agent is the biological organism. The human being that may utilise and lean on environmental resources in problem-solving tasks for instance. Now, rest of the exposition of the embodied-embedded approach's take on agency will be phrased in terms of a challenge, put forward by Rupert, to views such as enactivism, extended cognition and distributed cognition that do not conceive agency this *organismically-bound* way.[298]

Rupert holds on to a type of embodied-embedded approach, where the human cognitive system does not extend beyond the organism, even if during cognitive processing humans may exploit and lean on environmental and cultural scaffolding structures.[299] Specifically, Rupert's *hypothesis of embedded cognition* (HEMC) maintains that:

296 I would like to thank Anne Salminen for drawing my attention to important issues concerning the notion of agency in the context of 4ED while we were doing a joint presentation on this issue in 2012 (Calonius & Salminen 2012). Needless to say the remaining mistakes in this chapter are all mine.

297 Note that the focus here is not in phenomenological accounts of agency, in who or what experiences agency or how agency might feel like. See Rowlands 2010: 160–161 for this sort of distinction.

298 This challenge may be found in Rupert (2004: 425–428; 2009a: 44–47; 2009b: 102–105; 2010a: 328–334; 2010b: 346–352).

299 Rupert 2010b: 346

> Cognitive processes depend *very* heavily, in hitherto unexpected ways, on organismically external props and devices and on the structure of the external environment in which cognition takes place.[300]

Rupert's argument is explicitly directed against the *hypothesis of extended cognition* (HEC) where

> human cognitive processing literally extends into the environment surrounding the organism, and human cognitive states literally comprise–as wholes do to their parts–elements in that environment [...][301]

Rupert's argument is quite similar to the motley crew argument and his difference argument (which was presented in section 4.2.1). He argues that there is a scientific cost that ensues from adhering to HEC because it loses touch with the core cognitive agent due to the proliferation of various (often only temporary) cognitive systems. Moreover, if HEMC can account for the way current cognitive science explains cognitive phenomena we should just stick with the more conservative HEMC over HEC.[302] So, Rupert aims to deflate HEC.

Rupert maintains that it is the persisting human organism that houses the persisting and relevant cognitive capacities that explain diverse cognitive behaviour. He continues that hybrid extended cognitive systems vary much in their constitution and are often only temporary, i.e. they do not persist beyond the integrated interaction between the internal and external. Furthermore, Rupert notes how cognitive scientists are especially interested in persisting human organisms rather than the temporary hybrid extended cognitive systems. This is due to the *explananda* in cognitive science being mainly regularities in the organism's behaviour (such as patterns of intelligent behaviour), that are relatively independent of the surrounding environment and the organism's interactions with it. Rupert maintains that abilities that are of interest to cognitive scientists, such as: remembering, reading, perceiving, language use, reasoning, social cognition, are all explained by the persisting abilities, mechanisms, and capacities etc. of persisting systems, not by fleeting temporary coupled systems.[303]

Now, if it is really so that what cognitive science is interested in is explained in terms of

300 Rupert 2004: 393 original emphasis.

301 Rupert 2004: 389. Note that where Rupert presents this argument the main focus and dialogue is with Clark's exposition of extended cognition. Yet, in his (2010b: 343n.1 and 2) he explicitly refers to some of the authors of enactivism, extended cognition and distributed cognition (as well as some that have here been presented as proponents of the embodied-embedded approach) as advocating HEC. Therefore, in this chapter I will keep in line with Rupert's way of conceiving matters and will take enactivism, extended cognition and distributed cognition in general as proponents of HEC.

302 Rupert 2004: 425–428; 2009a: 44–47; 2009b: 102–105; 2010a: 328–334; 2010b: 346–352

303 Rupert 2010b: 346; 2010a: 330–334; 2009b: 103. I will not go through the specific empirical examples Rupert uses to support his argument. If one is interested in these see the references here. Note also that just this sort of view of cognitive behaviour is highly contested by distributed cognition for instance.

persisting organism-bound properties and HEC does not make use of this type of organism-bound cognitive architecture, then according to Rupert, HEC 'loses out on grounds of accuracy and explanatory power.'[304] Presumably this is where the scientific cost that Rupert is getting at ensues: if there is no core cognitive agent to refer to in explanations because the systems are so vastly varied and temporary, the required explanations for (given) cognitive phenomena would also be vastly numerous and diverse. Yet, Rupert continues:

> Once, however, the HEC-theorist's model has been appropriately articulated–so that it includes the integrated internal architecture necessary to explain such organismically local behavior as reading [etc.]–both HEC- and HEMC-theorists will have embraced structurally similar explanations of interactive cognitive processing, that is, of the cases that motivate HEC.[305]

According to Rupert two unappealing consequences for HEC ensue from this. Either 1) HEC loses on grounds of conservatism or 2) it loses on grounds of simplicity. 1) Rupert argues that HEC is an uninteresting position since it merely adds the label "cognitive" to the external resources (in addition to positing an internal (cognitive) system and interactions between the internal and external). This is in contrast to HEMC which recognises the same three structures/phenomena without the need of relabelling the external resources as cognitive. 2) Rupert notes that, on the other hand if HEC's contribution is to posit a single unified system over and above the distinct elements (that all the proponents in the debate recognise and with which HEMC so successfully accounts for the relevant cognitive phenomena) then HEC merely adds unnecessary and gratuitous complexity.[306]

7.2 Enactive Agency

The account of enactivism presented in section 3.2 may be understood essentially as a broad account of cognitive agency.[307] Hence, in what follows I will rehearse the relevant aspects of enactive cognition already presented in 3.2. with respect to agency and Rupert's challenge. Much of the application of enactivism's conception of agency to Rupert's challenge is my own extrapolation since there is yet to be found a systematic account of enactive agency that discusses Rupert's

304 Rupert 2010b: 347. Recall here Pöyhönen's (2012) point about the lack of a systematic account of explanatory power by the opponents and proponents of HEC.
305 Rupert 2010b: 347
306 Rupert 2010b: 347
307 Note how Torrance (2005) has a similar reading of the general projects within enactive cognition: a broader account examining cognition and a more focused one examining the nature of perception.

challenge and criticisms.

Recall that there were five essential notions in enactive cognition: autonomy, sense-making, emergence, embodiment, and experience. A key point of enactive cognition was to get rid of the cognitivists' sandwich model of cognition that divorced cognition, perception and agency. Instead, cognition, perception and agency should be thought of as intertwined. Embodied biological organisms are autonomous through their self-generating and self-maintaining activity. In addition, these organisms are sensorimotorly coupled to the environment, thus allowing them to bring forth (or enact) relevant information for them. Thus, for enactivism cognitive agency is found in these sort of autonomous biological organisms.[308]

So, contrary to Rupert's conception of agency I take it that enactivism conceives cognitive agency as organism-centered, rather than organism-bound.[309] As Thompson describes cognition as 'a knot or tangle of recurrent and re-entrant processes centered on the organism [...]'.[310] Thus if one were to extrapolate a response to Rupert's challenge from all this, one could argue that enactivism does not lose touch with the core cognitive agent and become explanatory simple or gratuitous since agency remains centered on the organism. Moreover, one could argue that Rupert's way of conceiving agency as organism-bound is wrongheaded since it misses the point about the intertwined nature of cognition, agency and perception, and thus mistakenly divorces agency from cognition and perception. It is Rupert's account that misses something explanatorily crucial. Furthermore, enactivism need not make such a strict three-fold division of distinct phenomena involved in cognition as Rupert claims is needed in HEC (i.e. internal cognitive system, external cognitive resources and interaction between the internal and external). Recall from 3.2 that for enactivism cognition is more holistic and emerges in the relational domain between the organism and the environment. This is opposed to the internalist or individualistic picture of cognition Rupert champions, whereby cognition is primarily within the organism (its brain/central nervous system (CNS hereafter)) that may then be influenced by external resources found in the environment.

Therefore, even if the brain/CNS plays an important role in cognition for enactivism, because of the relational conception of cognition it is not simply reduced to the brain/CNS. I take it applies to the question of agency alike. It is through the autonomous sense-making activity of the organism and the integration with the environment through the sensorimotor couplings that agency is *centered* on the organism (or perhaps more aptly in the relational domain of organism-environment for that matter) without it being *bounded* within the organism and its nervous system. Agency is not a static property of closed organism, but results from the enactive *activity* of the

308 Thompson 2005: 417–418

309 Note that this should be distinguished from Clark's hypothesis of organism-centered cognition that will be shortly discussed below.

310 Thompson 2005: 408 referencing Hurley 1998.

organism. Thus, Di Paolo's characterisation of enactive agency as the 'modulation of the coupling between the autonomous entity and its environment'[311] applies well here and captures the essence of what is to be explained in scientifically relevant phenomena.

7.3 Extended Agency

I will now turn my attention to extended cognition's take on agency. I will start with first-wave extended cognition, namely Clark's account of agency, and then continue with second-wave extended cognition. The second-wave account of agency is only a tentative analysis of how it might conceive the matter since there is no systematic account of agency found in the second-wave extended cognition literature.

Clark explicitly responds to Rupert's criticism by arguing for the *hypothesis of organism-centered cognition* (HOC) over Rupert's organism-bound cognition.[312] Clark says the following:

> Human cognitive processing (sometimes) literally extends into the environment surrounding the organism. But the organism (and within the organism, the brain/CNS) remains the core and currently the most active element. Cognition is organism centered even when it is not organism bound.[313]

Clark's aim is to show that extended cognition, when properly understood, does not loose grip of the persisting core of cognitive agency that is present in softly assembled hybrid systems, and that it is in fact the conservative HEMC that threatens to obscure that what is important with regards to explaining cognitive phenomena.[314]

First of all Clark maintains that cognitive agents are not individuated by first finding their cognitive mechanisms, but by identifying 'a reliable, easily identifiable physical nexus of perception and action, apparently driven by a persisting and modestly integrated body of goals and knowledge.'[315] It is only after this identification that the location and type of the relevant underlying mechanisms in/for a given cognitive task are identified. Secondly, Clark makes the “uncontroversial” assumption that the brain is at least currently the essential core element of individual human cognitive activity. Clark then asks whether the brain cares about 'the nature (biological or non-biological) or the location (organism bound or organism external) of the processing and storage

311 Di Paolo 2009: 15 fig.1
312 Note that this is independent of the enactive organism-centred agency.
313 Clark 2008a: 139
314 Clark 2008a: 111, 116
315 Clark 2008a: 118

resources soft-assembled to tackle some cognitive task'.[316] In responding no to this question Clark relies on Gray et al.'s[317] studies and their subsequent time-cost-based model of cognitive task performance, to which was briefly alluded to in section 6.3.1. The key idea in Gray et al.'s model that arises from their experiments is that our brains are indifferent to the information source and what ultimately matters when recruiting internal or external resources when completing cognitive tasks is the cost-effectiveness of the (mix of) elements. Hence, they argue for a "level playing field" with regards to the use of external and/or internal resources. Gray et al. took the time taken to complete a task as the prime cost function, but Clark maintains that Gray et al's conclusions are compatible with other cost functions as well (that may be determined depending on the given context and goals) and hence Gray et al.'s conclusions should be taken more generally.[318] This empirical evidence leads Clark to a *hypothesis of cognitive impartiality*, that resembles his 007-principle and the principle of ecological assembly presented in section 2.2:

> Our problem-solving performances take shape according to some cost function or functions that, in the typical course of events, accord no special status or privilege to specific types of operations (motoric, perceptual, introspective) or modes of encoding (in the head or in the world).[319]

The relevance of the hypothesis of cognitive impartiality for HOC, the role of the brain/CNS and the interplay of the elements within a hybrid system becomes clearer, when Clark makes two qualifications to the hypothesis of cognitive impartiality. Or rather Clark teases out two explanatory targets in play so as to prevent HOC from slipping into Rupert's HEMC. 1) The role of the brain/CNS is to "recruit" the extended cognising; i.e. it combines the various elements into a softly-assembled extended device. 2) Once this sort of device is in place cognitive processing, namely the flow and transformation of information, provide 'the machinery of ongoing thought and reason.'[320] So, Clark's point is that the brain/CNS plays an important role in building up and controlling the soft-assembled system through combining the influences of various different elements; whereas HEC brakes the boundaries of skin and skull and provides an account of the information flow and cognitive processing relevant for a given problem-solving (cognitive) task once the assembly is complete. Clark argues that HEMC threatens to obscure this (scientifically important) distinction between the two explanatory projects by erecting strict boundaries.[321] I will finish Clark's account of agency with the following rather long quote.

316 Clark 2008a: 118
317 Gray & Fu 2004; Gray & Veksler 2005; Gray et al. 2006
318 Clark 2008a: 118–122
319 Clark 2008a: 121
320 Clark 2008a: 122
321 Clark 2008a: 122, 137. Note that Clark (2008a: 123–136) steers away from any possible picture of the brain with an inner homunculus that controls the building of the soft-assembled system.

> [...] in rejecting the vision of human cognitive processing as *organism bound*, we should not feel forced to deny that it is (in most, perhaps all, real-world cases) *organism centered*. It is indeed primarily (though not solely) the biological organism that, courtesy especially of its potent neural apparatus, spins and maintains (or more minimally selects and exploits) the webs of additional structure that then form parts of the machinery that accomplishes its own cognizing. Just as it is the spider body that spins and maintains the web that then (following Dawkins 1982) constitutes part of its own extended phenotype, so it is the biological human organism that spins, selects, or maintains the webs of cognitive scaffolding that participate in the extended machinery of its own thought and reason. Individual cognizing, then, is *organism centered even if it is not organism bound*.[322]

Note the similarity to enactivism's insistence on the organism creating and maintaining its own domain of meaningfulness, but recognise also the slightly distinct take on the notion of organism in use here. Whereas enactivism emphasised the relational nature of the organism and environment without there being strict boundaries marking cognition, Clark seems to be conceiving the organism as some pre-determined entity with boundaries that may be extended.

Second-wave extended cognition have yet to provide a systematically developed, clear and/or unified account of cognitive agency. Hence I will extrapolate three accounts of agency from the second-wave literature. I take it that there are four available options for the second-wave theorist to take. Cognitive agency could be conceived in line with enactivism, Clark's organism-centered agency, and/or as distributed agency. It is also possible for second-wave extended cognition to postulate its own distinct take on agency.

Menary takes the formation of cognitive agency to be a dynamic process that ties intimately with the surrounding environment through coupling rather than starting from a predetermined position. He makes the following claim about cognitive agency:

> We are not just coupling artifacts to pre-existing cognitive agents; the organism becomes a cognitive agent by being coupled to the external environment.[323]

Now, Menary does not develop this claim further in his writings and therefore I take the interpretation of his account of agency to be up for grabs. I take the most viable reading of Menary's claim to be most inline with enactivism's account of agency. On the one hand, even tough Menary's use of the notion of organism could suggest an allusion to Clark's organism-centered view I deem

322 Clark 2008a: 123 original emphasis.
323 Menary 2006a: 342

this not to be a viable option since Menary clearly shies away from the idea of a predetermined (or "pre-existing") organism to which agency is then centered on. Yet it seems that for Menary agency is still somehow centered on the organism. Note for instance how it is only after the coupling that the organism may be said to be an agent; as is clear in the claim: *organism becoming a cognitive agent by being coupled to the external environment*. This passage seems to allude to a similar reading of the intimate relationship between the organism and the environment that enactivism holds. Moreover, just before the quoted passage Menary emphasises that the proper unit of cognitive analysis is the *integration* of internal and external vehicles.[324] I take it that this again echoes enactivism's point about the role of agency in modulating the interactions between the autonomous entity and the environment. On the other hand, Menary's focus on the notion of "organism" in the quoted passage does exclude him from a more distributed or decentralised view of agency, that is championed by Hutchins for instance. Hutchins' view, which will shortly be discussed more thoroughly, maintains that both the assembled cognitive systems and more importantly even the assembly, recruitment and control processes themselves (i.e. both of the explanatory projects that Clark distinguished) should be thought of as distributed. Thus I take it that Menary's account of agency is organism-centered, but not as Clark conceives it but as enactivism may hypothesise.

Yet, this sort of view on agency might not be widely accepted in second-wave extended cognition in general, or by Sutton in particular. As was noted in section 5.3 Sutton alludes to distributed cognition as offering a viable methodological take on the study of cognition that is compatible with extended cognition.[325] So one could extrapolate that Sutton's view on cognitive agency could be more in line with Hutchins' than with Clark's organism-centered view or Menary's enactivist perspective. Moreover, Kirchhoff,[326] inspired by Sutton's gesture at a possible distinct third-wave extended cognition, has recently argued that extended cognition (in both of its waves) relies on a too individualistic notion of agency (e.g. Clark's HOC), and what is missing is a more decentralised account of cognitive agency. In providing this sort of an account Kirchhoff relies heavily on Hutchins' distributed account of agency. Essentially the point that Kirchhoff is making is the same that Hutchins[327] makes in favour of his distributed view of agency to be discussed below.

Rowlands' account could also be a move towards a distributed view of agency. This I take to be manifested in the following quote:

324 Menary 2006a: 342
325 Sutton 2010; Sutton et al. 2010
326 Kirchhoff 2011
327 Hutchins 2011b

> For our purposes, the notion of a subject can, I think, be understood quite broadly. For example, I do not wish to rule out the possibility that the subject in question might be a group rather than an individual.[328]

Yet, Rowlands does develop his own distinct and rather complicated account of agency that is closely tied with his criteria for a mark of the cognitive. He mostly analyses the *ownership* of cognitive states and processes along the lines of *disclosing and revealing activity* of the world to the subject. Without going much deeper into Rowlands' difficult account, I take it that Rowlands also (this is in addition to Hutchins) allows the assembly and recruitment processes themselves and not simply the assembled cognitive system to be extended/distributed.[329]

> [...] the vehicles of cognition are causal disclosers of the world. World disclosure, in general, is entirely neutral over the nature and location of its vehicles. Sometimes they are neural operations, sometimes they are processes taking place in the body, or even processes that extend into the world in the form of manipulation, exploitation, and transformation of environmental structures.[330]

7.4 Distributed Agency

As has been noted in chapter 5 distributed cognition conceives cognition from a systemic point of view as socially, environmentally and culturally distributed. However, with regards to cognitive agency there is disagreement among theorists of distributed cognition. In what follows I will present and examine Ronald Giere's and Hutchins' views of agency. I focus on these two since they explicitly take up the question of agency in the first place and because their analysis and treatment of cognitive agency feeds directly into the current debate of 4ED.

Giere[331] advocates a distributed view of cognition that is much in line with that presented in chapter 5. Yet, it is Giere's view on agency that especially distinguishes him from the other distributed cognition theorists presented in this thesis, Hutchins in particular. Giere alludes to Bruno Latour's notion of hybrid systems; i.e. systems that are composed of combinations of human and non-human parts.[332] Giere conceives distributed cognitive systems essentially as hybrid systems thus accrediting external representations and the artifactual world to be encompassed in distributed cognitive systems. Up to this point Giere is in agreement with Hutchins. Yet, he departs from

328 Rowlands 2010: 135
329 Rowlands 2010: ch. 6–8
330 Rowlands 2010: 218
331 Giere 2004; 2007; 2011.
332 Giere 2007: 319; Latour 1993.

Hutchins' view by maintaining a threefold division of distributed cognitive systems: physical, computational and human. Specifically Giere reserves notions such as mind, intentionality and agency only to the human aspects of the distributed cognitive systems. Hence he has recently aligned himself with Clark's notion of organism-centered but not organism-bound cognition.[333] The main justification for Giere's position, I take it, is similar to Rupert's and Clark's since it is framed in terms of providing a theoretical framework that various scientific enterprises from philosophy, anthropology, history to psychology are able to share. For Giere it seems to be explanatorily superfluous and redundant to extend the notion of agency to parts of distributed cognitive systems other than humans.[334]

Now, Hutchins conceives cognitive agency as essentially distributed. First, recall the way Hutchins argued for cultural practices as key elements in shaping cognition in section 5.2 and how one must be careful when attributing relevant cognitive properties to individuals taking part in cultural practices that might in fact more aptly belong to a distributed system. Recall also that the roots of Hutchins' account of agency lie in his reaction against Clark's organism-centered view. Hutchins argues that Clark's view is too individualistic and centered too heavily on the brain. Hutchins notes how Clark supports the idea that the assembled cognitive systems may be extended (Clark's second explanatory target) but fails to see 'that the assembly process itself is extended and orchestrated by the cultural practices that constitute the cognitive niche.'[335] Hutchins thus takes away the assembly and recruitment processes, that Clark emphasised belonged solely to the brain (Clark's first explanatory target), and distributes them to the surrounding dynamic culture, i.e. cultural practices.

Hutchins argues that Clark is forced to take up his hypothesis of organism-centered cognition due to a misunderstanding of the role of culture in cognition (I take Hutchins' argumentation to apply to Rupert's HEMC and Giere's position alike). First, Clark misses the fact that the cultural world is dynamic and not static like Otto's notebook. It for instance includes dynamic activities of other social people and thus allows the dynamic organising, assembly and recruitment processes to include other possible sources besides the brain and/or the body. Second, culture is falsely reduced simply to internal mental representations and lifeless artefacts. Cultural practices do include internal representations but are not solely identified by them, as was noted in section 5.2.[336] Hutchins maintains that by getting rid of these misunderstandings the organism-centered (or for that matter organism-bound) thinking of cognitive agency could be removed. It is cultural practices that contribute to the organisation and assembly of distributed cognitive agency. Hence, Hutchins proposes his own *hypothesis of enculturated cognition*:

333 Giere 2011: 398
334 Giere 2004; 2007; 2011.
335 Hutchins 2011b: 442
336 Hutchins 2011b: 442–444

> The ecological assemblies of human cognition make pervasive use of cultural products. They are always initially, and often subsequently, assembled on the spot in ongoing cultural practices[337]

Now, Rupert has reacted against this sort of distributed agency since according to him it implies the presence of *group cognition/minds*.[338] Group cognition refers to the idea that a group of people in itself may bear cognitive states over and above its constituent individuals. That is, group cognition supervenes on, is realised by, but is not simply reducible to the cognitive processes, actions and agency of the members of the group. Rupert focuses on group cognitive states as representational states that take the form of expressions in public language. Such as written court verdicts and press releases. These states are supposed to be *cognitive states*, analogous to the thoughts and representations of an individual, even if they appear to be of linguistic form. Rupert focuses on representations because he takes them to be a central feature of cognition, a feature which according to him groups lack. Moreover Rupert focuses group representations as expressions of public language because he takes most of the proponents of group cognition, such as Hutchins and Margaret Gilbert, to conceive them this way.[339] Hence, I will refer to individual cognitive states as *individual representation*s and group cognitive states as *group representations.*

Rupert's argument has two parts: he first offers a general objection to the causal explanatory power of group cognition and then applies these general concerns to particular cases of naturalistic accounts of mental representation. The general objection aims to show that nothing explanatory is gained from positing group cognition over and above what is gained from a more thorough understanding of the individuals involved in the group. Note, here how Giere's argumentation is remarkably similar. Giere is also weary of the notion of group/collective cognition and maintains that what ultimately explains the final cognitive output of the group is the interactions of the individuals of the group acting together.[340] Rupert specifically maintains that it is explanatorily superfluous to equate the expressions of public language with autonomous cognitive states (group representations).[341] Rupert gives two sufficient conditions for this and argues that:

---

337 Hutchins 2011b: 445

338 Rupert 2005. Note that one of the reasons for Giere to restrict cognitive agency to the human part of the distributed system was to avoid the explanatorily redundant postulation of this sort of group cognition, or what he dubbed *collective cognition.* (Giere 2007: 319)

339 Rupert 2005: 177–180. I accept Rupert's way of conceiving group representations only because I want to put forward a counter-argument that meets Rupert at his own theoretical territory. In the end I will point to another dynamic reading of group representations that Rupert neglects. Note also that Rupert takes Gilbert as supporting the ontological claim of group minds, even though this is not necessarily Gilbert's position with regards to her *plural subject* account (Gilbert 2000: 21fn.23, 22).

340 Giere 2007: 319

341 Rupert 2005: 178–180

> After all, every step in the construction of such representations, as well as every step in the causal sequence alleged to involve the effects of those representations, *proceeds either* [(1)] *by brute physical causation* (e.g., photons emitted from the surface of the page stimulate the reader's retinal cells) *or* [(2)] *by causal processes involving the mental states of individuals.*[342]

But an initial worry arises straight out of these conditions. If group cognition is to be dismissed due to lack of explanatory power because they supervene on, or are realised by, individual mental states, individual cognition should also be dismissed as explanatorily superfluous since it supervenes on physical neurological processes. Otherwise Rupert would be simply begging the question against group representations by *a priori* excluding representational systems that for their representations rely on subsystems capable of representation themselves.[343]

Rupert recognises this problem and counters it by maintaining that there is an important difference in understanding "inter-level relations", which allows individual representations to retain their causal-explanatory power. Rupert argues that we do not have any idea how to reduce individual representations to their corresponding neurological realisers, i.e. psychological laws and regularities to physical laws and regularities.[344] That is, to *eliminate* the talk of mental states in favour of talk about brain states. But we do have an idea how to account for group representation, by making reference to the individual representations of the group. For instance, Rupert maintains that a legitimate court verdict is just a 'majoritarian agglomeration of the opinions of the court's members, set down on paper.'[345] So, individual representations are explanatorily useful since they cannot (yet) be explained by reducing them to their realisers, but group representations are explanatorily superfluous because they can be explained by making reference to individual representations of the members of the group. Let's call this further qualification of conditions (1) & (2) with the notion of understanding "inter-level relations", problem (3).[346]

These general considerations lead Rupert to a more particular objection. He views five naturalistic accounts of mental representation[347] and argues that none of them plausibly applies to group representations. Rupert aims to show that group representations as expressions of public language cannot be made naturalistically represent something in the world. I am not going to view this objection in detail since I will object to Rupert's general argument, and by rejecting the general account beg for its clarification before any particular objections may be drawn from it.[348]

Now, Bryce Huebner notes that Rupert's condition (1) applies equally to individual

342 Rupert 2005: 179 emphasis added.
343 Rupert 2005: 179–180; Huebner 2008: 99
344 Cf. Dretske's (1993) emphasis on distinguishing structuring and triggering causes of behaviour.
345 Rupert 2005: 179
346 Rupert 2005: 179–180; Huebner 2008: 99
347 Those of indicator semantics (Dretske), pure-informational semantics (Fodor), teleological semantics (Millikan), causal-historical semantics (Prinz), and teleo-isomorphic semantics (Cummins). For more details on these and references to the original expositions see Rupert (2005: 180–182).
348 Rupert 2005: 180–184

representations. All theories that aim to give a naturalistic account of mental representations (i.e. including just those that Rupert relies on in his detailed objection) necessary require that the production and use of representations proceed by some causal or physical process. But, if group representations are denied on the basis of being explanatorily superfluous because they involve brute physical causation and if individual representations necessarily involve causal or other physical processes, then individual representations should also be denied on the same basis. Rupert cannot accept this conclusion, in so far as he relies on naturalised theories of representation in his other objection. Therefore, (1) fails as a sufficient condition for establishing the explanatory superfluity of group cognition. Hence, there needs to be some other unique problem with group representations.[349]

Problem (3) may be seen as this unique difficulty to group representations. But first, I would like to set aside an initial worry that arises from Rupert's insistence on the proper understanding of "inter-level relations", and then follow up on Huebner's much more charitable reading of (3). Now, if it is so that individual representations are explanatorily useful only because they cannot be eliminated in favour of the talk of their realiser brain states, then this implies that the explanatory value of individual representations is simply up to the current scientific knowledge. That is, if science eventually succeed in the elimination, then individual representations will become as explanatorily superfluous as group representations, maybe it is just a matter of time when this happens. As was noted above Rupert would not accept the (possible) elimination of individual representations.[350]

To resist this eliminativist outcome, (3) may be seen in another way, which directly feeds into condition (2) and Rupert's detailed objection. Huebner emphasises this charitable reading of (3) and notes that the unique difficulty for group cognition may lie in the fact that group representations do not introduce anything new in *kind* as opposed to individual representations. That is, by moving from physical neurological states to representational states individual representations gain *semantic content*. They gain it by having the appropriate causal relations (perceptual processes) between the perceiver and the world so that they may indicate and carry information about the world. Group representations on the other hand lack these appropriate relations. It is specifically because already contentful individual representations realise them, they fail to gain anything new. So, because it is individual representations that explain semantically important states and processes in the world, such as intentional states, they cannot be made explanatorily superfluous by further scientific discovery. Also, since group representations bear only content *derived* from the representational states of the individuals that produce them, group representations may be reduced to the aggregate of its individual representations.[351]

---

349 Huebner 2008: 98–99

350 Huebner 2008: 99. Cf. Churchland's (1981) point about eliminative materialism.

351 Huebner 2008: 100; Rupert 2005: 180–184

Huebner argues that even this charitable reading fails to establish Rupert's general argument because even individual representations derive their content from lover-level states that are already contentful. In order to illuminate this claim Huebner presents a pathological case of Capgras syndrome. Everything in a patient with a Capgras syndrome is functioning correctly with the exception that when she sees someone familiar to her, e.g. her mother, she has the unshakable belief that it is not her mother but an impostor. This phenomenon is thought to arise from a failure to bind visual representations to affective representations of feelings of familiarity. This again suggests that in order to represent someone *as* one's mother: a) the visual system needs to function properly in order to produce the correct representation, b) the affective response to this stimulus must be correct and produce representations of feelings of familiarity, and c) the association of visual and affective representations must follow correct rules. Hence, it seems that individual representations consist of subcomponents that produce representations themselves. That is, individual representations supervene on structures that are already contentful. Therefore, Rupert's argument against group representations is yet again in danger of precluding individual representations. For if it is the case that group representations are explanatorily superfluous because they supervene on already contentful representations, then individual representations are similarly superfluous since e.g. the representation of someone as one's mother *proceeds by causal processes involving semantically contentful subcomponents* of the visual system and the rules of association.[352]

To conclude, recall that Hutchins argued for a dynamic conception of the cultural and social world. I think this idea could be expanded to group representations as well; group representations should be conceived as dynamic representations that arise from and guide further behaviour, not as static public language structures. This aspect of representations is neglected by Rupert; as Huebner notes: 'taking [...] public language structures to exhaust the representational states of collectivities is analogous to taking an individual's utterances to exhaust her mental representations'.[353] Finally consider Hutchins' navigation example presented in chapter 5. In the task of marking the ships location a great deal of coordinated behaviour and problem solving among different individuals and parts of the ship was required. But, even though this task involved numerous representational subsystems (e.g. the particular individuals) no single subsystem was solely responsible for the representation of the ship's location. Rather, this representation came as a whole, and was best explained as being realised by the relevant individuals and their coordinated actions. Hence, one could make the claim that there is group cognition supervening on the crew of the ship and the representation is irreducible to any individual part of the group. Therefore, the representation of the ship's location is best explained as a group representation.[354]

352 Huebner 2008: 101–103
353 Huebner 2008: 105
354 Hutchins 1995: 117–119, 128–131; Tollefsen 2006: 148–149; Huebner 2008: 105–106

7.5 Summary

In this chapter I examined the issue of cognitive agency and applied 4ED's take on it. This further highlighted the overlapping and distinguishing features the different 4ED approaches share with each other. I followed similar structure as was present in chapters 2–5 and started with Rupert's embodied-embedded take on cognitive agency and his proposed challenge to the other 4ED views.

Rupert identified agency as organism-bound and argued from deflationary reasons that no-other take on agency would suffice since they would loose touch with the core cognitive agent. He maintained that the organismically-bound conception of agency accounts for all the relevant cognitive phenomena and nothing (other than confusion and gratuitous complexity) is gained from postulating agency any other way. Enactivism challenged the organism-bound view and maintained that rather than being bound to the organism, agency is merely centred on it. This way one does not loose touch with the core agency. The driving idea behind this was the commitment to the holistic intertwined nature of cognition, perception and agency that enactivism upholds. From an enactivist perspective Rupert's view seems misguided since it mistakenly divorces agency from cognition and perception.

Extended cognition had many takes on agency. Clark responded to Rupert and argued for an organism-centered view by distinguishing two explanatory projects. He interestingly still relied on the brain/CNS having the control and recruitment processes but allowed the information flow processes to be softly-assembled and extended. According to Clark Rupert's view misses this important distinction. Now, second-wave theorists' take on cognitive agency had allusions to both the enactivist organism-centred view (Menary) and Hutchins' distributed conception (Sutton, Kirchhoff and Rowlands).

I finally turned my attention to how distributed cognition conceived agency. I distinguished two takes on the matter. Giere argued along the lines of Clark and justified his organism-centered view similarly to Rupert and Clark by appealing to explanatory reasons. Hutchins reacted against especially Clark's organism-centered view and maintained that it relied on a too individualistic notion of agency and static picture of culture. Hutchins argued that both of the explanatory projects Clark distinguished (the assembled system and importantly the assembly process itself) should be regarded as distributed. He accomplished this through his notion of cultural practice. Finally I viewed how Hutchins' view implied the existence of group cognition and how Rupert failed to dismantle their existence as explanatorily superfluous.

8 Conclusion

The title of this thesis alludes to the location of cognition and paraphrases Clark and Chalmers[355] when suggesting that "It ain't all in the head". But if cognition ain't all in the head, then where is it? The subtitle gestures the possible route to an answer that I have taken in this thesis by "situating cognition to the body and the surrounding world". After briefly describing the orthodox isolationist view of cognition I narrowed down my exploration in to this matter into four views on cognition that do not restrict cognition solely inside the cranium. These were embodied-embedded cognition, enactive cognition, extended cognition, and distributed cognition.

Four research questions have been guiding this thesis: I) What are the theoretical commitments the different accounts of cognition in 4ED hold on to and how do these relate to each other? II) (Why) should distributed cognition be added with the rest of the 4E accounts? III) What critique does 4ED face and how does providing a mark of the cognitive affect both the proponents as well as the opponents of 4ED? IV) How does the notion of cognitive agency figure in the 4ED approach(es) to cognition?

With regards to the first research question it seems that all the views share a similar commitment to challenge the cognitivist sandwich model of cognition. There is also the tendency, with varying degrees, to credit more role to the body (sensorimotor activity) and the environment (including culture, artefacts and instruments) in cognitive processing. The differences in the views relate specifically on how this relationship is conceived and to what degree the body and the environment affect cognition. It is here that lies a very interesting issue. There seems to be an ostensible thematic gradation that, compared to the orthodox cognitivist position, moves from less to more radical with regards to how/where the boundary of cognition may be drawn. This thematic progression also mirrors the structure of the thesis in chapters 2–5. Now, even though the approaches share similar theoretical commitments and rely on many notions and concepts from each other (note how for instance the notions of scaffolding, niches, ecology, and embodiment feature prominently in most, if not all of the views) there seems to be emerging a sort of continuum where a growing influence on matters outside the brain are said to take part in cognition when one moves from embodiment and embeddedness to enaction and extension all the way to the distribution of cognition. To begin with, the body and the environment are said to have facilitating, constraining and coordinating influences on cognitive processing (embodied-embedded cognition). Then cognition is moved to the relational domain between the active organism and the environment (enactive cognition). This is followed by allowing the literal extension of cognition in to environmental resources on functionalist, parity-based and/or integrative complementarity grounds (first- and second-wave cognition). Ending with a change of perspective that does not

355 Clark & Chalmers 1998: 8

extend the cognitive boundaries of organisms but rather *necessitates* the incorporation of culture, history and artifacts etc. into the cognitive ecology of problem solving organisms (distributed cognition). It is important to note that even if this thematic form seems apparent it should not be taken as a given and may very well simplify matters too much as there is a lot of variation within and between the approaches.

This leads to another related and interesting point. It seems that the examination here has highlighted both the distinct character (found in the research interests, background and in the theoretical and conceptual positions of the different 4ED approaches) as well as the shared attitude towards the nature and study of cognition as a non-isolated phenomenon. Hence, this study speaks in favour of not compiling the different approaches into one wholesale position that in doing so looses the theoretical nuances and research tradition of each approach. Note how just these sort of considerations were also prominent in answering the second research question to do with the addition of distributed cognition with the rest of the 4Es in chapter 5. These sort of considerations again underline the question whether 4ED constitutes a framework or a general approach for cognition or merely a group of similarly committed and inspired approaches to cognition. I am not ready to declare that a framework has emerged. Rather, I would like to paraphrase Kiverstein & Clark:[356] even if 4ED does not (yet) form a unified church, the mere questioning of this and the further examination of these sort of approaches to cognitive phenomena will likely open up new vistas of discovery for the study of cognition and its place in the world.

Now, the third research question concerning the mark of the cognitive has its roots in the reactions of the opponents of 4ED type cognition. Chapter 6 ended on a note about how Adams and Aizawa's insistence on a proper *mark of the cognitive* retained its significance in the discussion even if their objections against 4ED that were based on it failed. Firstly the notion of a mark of the cognitive seems to push the orthodox cognitivists to take part in the debate about the proper boundaries of cognition and in doing so raise the viability of 4ED as an available position. Thus the ground between the orthodox and unorthodox positions seems to be levelling. Secondly a host of meta-level questions about the requirement for a mark of the cognitive arise. Does one need a mark of the cognitive in order to draw the boundaries of cognition or even study cognitive phenomena? Could one draw the boundaries on some other grounds (e.g. epistemic or explanatory power considerations)? Or would it be best to just proceed with scientific research and let science discover the boundaries if ever they are to be discovered?

The fourth research question on agency narrows down the examination of 4ED into a more detailed issue. The discussion revolving around the notion of cognitive agency and 4ED's take on it draws out further overlap and differences between the different approaches. For these reasons it is a growingly relevant issue in the 4ED debate. It is also an independently interesting issue since it

356 Kiverstein & Clark 2009: 6

shakes our ordinary understanding of what cognitive agency means. The views on agency varied from organism-bound (cognitivism and embodied-embedded cognition), organism-centered (enactivism, extended and distributed cognition) to distributed agency (extended and distributed cognition). I find it interesting that the notion of agency neither lumps the 4ED together but rather for instance shows some common ground between cognitivism and embodied-embedded cognition. Another surprising issue is to do with Clark's account of agency. Even though he puts forward a view of cognition that radically departs from the orthodox position, his reliance on the role of the brain and the central nervous system in agency puts him again closer to the more conservative positions. Whereas if you look at Hutchins' distributed account, both the notions of cognition and agency boldly depart from the more conservative positions. Distributed agency goes so far as to imply the possible viability of group agency.

I would like to conclude this thesis by raising some open questions and future applications of the 4ED approach. In addition to the notion of cognitive agency one could investigate the normative issues that result from adopting a 4ED position. This is an interesting matter even more so because there is lacking a systematic take on this on the part of the 4ED theorists. What is missing in the discussion of cognitive agency is for instance questions to do with responsibility: who or what elements can be said to carry responsibility of an action credited to a hybrid system, or where lies the responsibility of a group decision? It would be very interesting to see what kind of other issues would arise when the notions of organism-bound, organism-centred and distributed agency would be approached from the perspective of ethics.

One could analyse further the significance of the notion of mark of the cognitive and the possible ways of demarcating the cognitive from the noncognitive. For instance one could bear down on the requirement of such a mark for cognitive practice and the empirical study of cognition. One could also investigate the notion of representation that is used by the various authors here. Should we get rid of representations all together, or only some forms of it? How should representations be understood in the first place? Could they possibly be for instance understood as artefacts[357] and how does this fit with the 4ED view? How does an extended or distributed understanding of representations affect the notion of representation in the sciences?

Still, another area where the 4ED approach could bear fruit is the debate surrounding social cognition, i.e. how we come to understand the thoughts, feelings and intentions of others. For instance Gallagher has applied his embodied-embedded approach to the mindreading/interpersonal understanding debate and re-interpreted mirror neuron findings in favour of his *interaction theory* of social cognition.[358] Also, as was noted in chapter 5.2, it would be highly interesting to bring in the analysis of philosophy of action theorists, such as Gilbert, Bratman,

357 See e.g. Knuuttila (2011).
358 See Gallagher (2001, 2004, 2005, 2007, 2008b, 2009b) and De Jaegher (2009a, b).

Tuomela, Velleman and Tollefsen to issues of cognitive and collective agency. Or the other way around: bring the issues that have risen in the 4ED conceptions to the philosophy of action debate. Lastly, it seems that there is still an open question with regards to whether there emerges a unified framework for the study of cognition from 4ED or whether 4ED remains merely as collection of loosely tied approaches? Suffice it to say that even if 4ED remains (for now) as a collection of distinct approaches there is a kindling of an emerging paradigm. The take-home message I wish to suggest with this thesis is that even if the liberation of cognition from the confines of the head is a complex issue, being open to this kind of possibility will nevertheless bring forth new and interesting ways of understanding cognitive phenomena.

Bibliography:

Adams, Fred & Aizawa, Ken. 2001. The Bounds of Cognition. *Philosophical Psychology.* 14(1): 43–64.

– 2008. *The Bounds Of Cognition.* Malden (MA): Blackwell Publishing.

– 2009. Why the Mind Is Still in the Head. *The Cambridge Handbook of Situated Cognition.* eds. Robbins Philip & Aydede Murat. New York: Cambridge University Press. 78–95.

– 2010a. Defending the Bounds of Cognition. *The Extended Mind.* ed. Menary, Richard. Cambridge (Massachusetts): A Bradford Book. The MIT Press. 67–80.

– 2010b. The value of cognitivism in thinking about extended cognition. *Phenomenology and Cognitive Sciences.* 9: 579–603.

Adams, Fred. 2010. Embodied Cognition. *Phenomenology and Cognitive Sciences.* 9: 619–628.

Adolph, Karen E., Eppler, Marion A. & Gibson, Eleanor J. 1993a. Crawling versus Walking Infants' Perception of Affordances for Locomotion on Slopes. *Child Development.* 64(4): 1158–1174.

– 1993b. Development of Perception of Affordances. *Advances in Infancy Research.* Eds. Rovee-Collier Carolyn & Lipsitt Lewis P. Norwood (New Jersey): ABLEX publishing. 8: 51–98.

Aizawa, Ken. 2007. Understanding the embodiment of perception. *The Journal of Philosophy.* 104(1): 5–25.

– 2010. The coupling-constitution fallacy revisited. *Cognitive Systems Research.* 11: 332–342.

Aizawa, Ken & Adams, Fred. 2005. Defending Non-Derived Content. *Philosophical Psychology.* 18(6): 661–669.

Anderson, Michael. 2003. Embodied Cognition: A field guide. *Artificial Intelligence.* 149: 91–130.

Ballard, Dana. 1991. Animate Vision. *Artificial Intelligence Journal.* 48: 57–86.

Ballard, D., Hayhoe, M. and Pelz, J. B. 1995. Memory representations in natural tasks. *Journal of Cognitive Neuroscience.* 7(1):66–80

Bateson, Gregory. 1972 [1987]. *Steps to an ecology of mind: collected essays in anthropology, psychiatry, evolution, and epistemology.* Northvale (New Jersey): Jason Aronson Inc.

Beer, Randall. 1990. *Intelligence as Adaptive Behaviour.* New York (NY): Academic Press.

Bird, Alexander & Tobin, Emma. 2010. Natural Kinds. *Stanford Encyclopedia of Philosophy (Summer 2010 Edition).* ed. Zalta, Edward N. http://plato.stanford.edu/archives/sum2010/entries/natural-kinds/. Accessed in May 2012.

Block, Ned. 1978. Troubles with Functionalism. *Minnesota Studies in Philosophy of Science.* 9: 261–325.

– 2005. Review of Alva Noë's" Action in Perception". *Journal of Philosophy.* 102(5): 259–272.

Boyd, Richard. 1991. Realism, Anti-foundationalism and the Enthusiasm for Natural Kinds. *Philosophical Studies.* 61: 127–148.

– 1999. Homeostasis, Species, and Higher Taxa. *Species: New Interdisciplinary Essays*. ed. Wilson, Robert. Cambridge (Massachusetts): A Bradford Book. The MIT Press. 141–185.

Bratman, Michael. 1993. Shared Intention. *Ethics.* 104(1): 97–113

– 1999. I Intend That We *J. Faces of Intention: Selected essays on intention and agency.* New York (NY): Cambridge University Press. 142–161

Brooks, Rodney A. 1990. Elephants Don't Play Chess. *Robotics and Autonomous Systems*.6:[3–15], accessed from http://people.csail.mit.edu/brooks/papers/elephants.pdf. 1–12. accessed in August 2011.

– 1991a. Intelligence Without Representation. *Artificial Intelligence*. 47: [139–159]. accessed from http://people.csail.mit.edu/brooks/papers/representation.pdf. 1–12. accessed in August 2011.

– 1991b. Intelligence Without Reason. *MIT: Artificial Intelligence Laboratory Memo*. No. 1293: 1–28.

– 1991c. Integrated Systems Based on Behaviours. *SIGART Bulletin*. (2:4): 46–50.

Brooks, Rodney A., Connell, Jonathan H. and Ning, Peter. 1988. Herbert: A second generation mobile robot. *MIT AI Memo*. No. 1016.

Burge, Tyler. 1979. Individualism and the mental. *Midwest Studies in Philosophy*. 4: 73–122.

Calonius, Lauri & Salminen, Anne. 2012. Mielen "Laajentuminen". Presentation in *Filosofian Akatemian tutkijaseminaari 4.5.2012*. Helsingin Yliopisto

Chemero, Anthony. 1998. A Stroll Through the Worlds of Animats and Humans: Review of *Being There: Putting Brain, Body and World Together Again* by Andy Clark. *Psyche*. 4(14): 1–10. accessed from http://theassc.org/files/assc/2369.pdf. Accessed in September 2011.

Chemero, Anthony & Silberstein, Michael. 2008. Defending Extended Cognition. *Proceedings of the 30th Annual Meeting of the Cognitive Science Society*. Eds. Love, McRae and Sloutsky. 129–134

Churchland, Paul. 1981. Eliminative Materialism and the Propositional Attitude. *Journal of Philosophy.* 78: 67–90

Clark, Andy. 1989. *Microcognition: Philosophy, Cognitive Science, and Parallel Distributed Processing*. Cambridge (Massachusetts): The MIT Press.

– 1997 [1998]. *Being There: Putting Brain, Body, and World Together Again*. Cambridge (Massachusetts): The MIT Press.

– 1998. Author's Response: Review symposium on *Being There*. *Metascience*. 7: 95–103.

– 2005a. Word, Niche, and Super-Niche: How Language Makes Minds Matter More. *Theoria*.54: 255–268.

– 2005b. Intrinsic content, active memory and the extended mind. Analysis. 65(1): 1–11

– 2006. Soft Selves and Ecological Control. To appear in *Distributed Cognition and the Will*. Eds. Spurrett D., Ross D., Kincaid H. and Stephens L. Cambridge (Massachusetts)eThe MIT Press. Accessed from http://hdl.handle.net/1842/1446. 1–37. Accessed in December 2011.

– 2007. Curing Cognitive Hiccups: A Defense Of The Extended Mind. *The Journal of* Philosophy. 54(4): 163–192.

– 2008a. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension.* New York: Oxford University Press.

– 2008b. Pressing the Flesh: A Tension in the Study of the Embodied, Embedded Mind?. *Philosophy and Phenomenological Research.* 76(1): 37–59.

– 2010a. *Memento*'s Revenge: The Extended Mind, Extended. *The Extended Mind.* ed. Menary, Richard. Cambridge (Massachusetts): A Bradford Book. The MIT Press. 43–66.

– 2010b. Coupling, Constitution, and the Cognitive Kind: A Reply to Adams and Aizawa. *The Extended Mind.* ed. Menary, Richard. Cambridge (Massachusetts): A Bradford Book. The MIT Press. 81–99.

Clark, Andy & Chalmers, David J. 1998. The Extended Mind. *Analysis.* 58(1): 7–19

Connell, Jonathan H. 1989. A colony architecture for an artificial creature. *MIT AI Lab Technical Report.* No. 1151.

Cowart, Monica. 2005. Embodied Cognition. *The Internet Encyclopedia of Philosophy.* http://iep.utm.edu/e/embodcog.html. accessed in October of 2011.

Craver, Carl F. 2009. Mechanisms and natural kinds. *Philosophical Psychology.* 22(5): 575–594.

Dawkins, Richard. 1982. *The Extended Phenotype.* Oxford (UK): Oxford University Press.

Dehaene, Stanislav. 2007. A few steps towards a science of mental life. *Mind, Brain and Education.* 1(1): 28–47.

Dehaene S., Spelke E., Pinel P., Stanescu R., & Tsivkin S. 1999. Sources of mathematical thinking: Behavioral and brain imaging evidence. *Science.* 284: 970–974.

De Jaegher, Hanne. 2009a. Social understanding through direct perception? Yes, by interacting. *Consciousness and Cognition.* 18: 535–542.

– 2009b. What made me want the cheese? A reply to Shaun Gallagher and Dan Hutto. *Consciousness and Cognition.* 18. 549–550.

Descartes, René. 1641[2008]: *Meditations on the First Philosophy: With Selections from the Objections and Replies.* Translated with an Introduction and Notes by Michael Moriarty. New York: Oxford University Press.

Dewey, John. 1896. The Reflex Arc Concept in Psychology. *The Psychological Review.* 3(4): 357–370.

– 1916. *Essays in Experimental Logic.* Chicago: University of Chicago Press.

Di Paolo, Ezequiel. 2005. Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences.* 4: 429–452.

– 2009. Extended Life. *Topoi.* 28: 9–21.

Di Paolo Ezequiel, Rohde Marieke, and De Jaegher Hanneke. 2007 [to appear]. Horizons for the Enactive Mind: Values, Social Interaction, and Play. To appear in *Enaction: Towards a New Paradigm for Cognitive Science*. Eds. Stewart J., Gapenne O., and Di Paolo E. Cambridge (Massachusetts): The MIT Press. Accessed from http://www.sussex.ac.uk/Users/ezequiel/DiPaoloetal_csrp587.pdf. 1–56. Accessed in January 2012

Drayson, Zoe. 2010. Extended cognition and the metaphysics of mind. *Cognitive Systems Research*. 11: 367–77.

Dretske, Fred. 1993. Mental Events as Structuring Causes of Behaviour. *Mental Causation.* eds. Heil, John & Mele, Alfred. Oxford (UK): Oxford University Press. 121–136

Ereshefsky, Mark. 2010. Species. *Stanford Encyclopedia of Philosophy (Spring 2010 Edition)*. Ed. Zalta, Edward N. http://plato.edu.archives/spr2010/entries/species/. Accessed in May 2012.

Fodor, Jerry. 1983. *The Modularity of Mind*. Cambridge (Massachusetts): The MIT Press.

– 2000. *The Mind Doesn't Work That Way*. Cambridge (Massachusetts): The MIT Press.

– 2009. Where is my mind?. *London Review of Books*. 31(3): 13–15. Accessed from http://www.lrb.co.uk/v31/n03/jerry-fodor/where-is-my-mind. Accessed in May 2012

Forssberg H. 1985. Ontogeny of human locomotor control I. Infant stepping, supported locomotion and transition to independent locomotion. E*xperimental Brain Research*. 57(3): 480–493.

Gallagher, Shaun. 2001. The Practice of Mind: Theory, Simulation or Primary Interaction?. *Journal of Consciousness Studies*. 8(5–7). 83–108.

– 2004. Understanding Interpersonal Problems in Autism: Interaction Theory as An Alternative to Theory of Mind. *Philosophy, Psychiatry, &* Psychology. 11(3): 199–217

– 2005. *How The Body Shapes The Mind*. Oxford: Clarendon Press.

– 2007. Simulation trouble. *Social Neuroscience*. 2(3–4): 353–365

– 2008a. Intersubjectivity in perception. *Continental Philosophy Review*. 41: 163– 178.

– 2008b. Direct perception in the intersubjective context. *Consciousness and Cognition.* 17: 535–543.

– 2009a. Philosophical Antecedents of Situated Cognition. *The Cambridge Handbook of Situated Cognition*. Eds. Robbins, Philip & Aydede, Murat. New York (NY): Cambridge University Press. 35–53.

– 2009b: Deep and dynamic interaction: Response to Hanne De Jaegher. *Consciousness and Cognition.* 18: 547–548

– 2011. The socially extended mind. *Conference paper at Free University of Berlin.* In 21–22 March 2011. 1–13.

Gallagher, Shaun & Zahavi, Dan. 2008. *The Phenomenological Mind: An Introduction to Philosophy of Mind and Cognitive Science*. London: Routledge

Gangopadhyay Nivedita & Kiverstein, Julian. 2009. Enactivism and the Unity of Perception and Action. *Topoi*. 28: 63–73.

Gangopadhyay Nivedita, Maday Michael & Spicer Finn (eds.). 2010. *Perception, Action, and Consciousness: Sensorimotor Dynamics and Two Visual Systems*. Oxford (UK): Oxford University Press.

Gibson, James J. 1979. *The Ecological Approach To Visual Perception*. Boston: Houghton Mifflin Company.

Giere, Ronald. 2004. The problem of agency in scientific distributed cognitive systems. *Journal of Cognition and Culture*. 4(3–4): 759–774

– 2007. Distributed Cognition without Distributed Knowing. *Social Epistemology*. 21(3): 313–320

– 2011. Distributed Cognition as Human Centered although not Human Bound: Reply to Vaesen. *Social Epistemology: A Journal of Knowledge, Culture and Policy*. 25(4): 393–399.

Gilbert, Margaret. 1990. Walking Together: A Paradigmatic Social Phenomenon. *Midwest Studies in Philosophy*. 15: 1–14

– 2000. What Is It for *Us* to Intend?. *Sociality and Responsibility: New Essays in Plural Subject Theory*. Oxford (UK): Rowman & Littlefield Publishers Inc. 14–36

– 2004. Collective Epistemology. *Episteme*. 1(2): 95–107

– 2009. Shared intention and personal intentions. *Philosophical Studies*. 144(1): 167–187

Glüer, Kathrin & Wikforss, Åsa. 2010. The Normativity of Meaning and Content. *Stanford Encyclopedia of Philosophy (Winter 2010 Edition)*. Ed. Zalta, Edward N. http://plato.stanford.edu/archives/win2010/entries/meaning-normativity/. Accessed in October 2012.

Gray, Wayne & Fu, W.-T. 2004. Soft constraints in interactive behavior: the case of ignoring perfect knowledge in the world for imperfect knowledge in the head. *Cognitive Science*. 28(3): 359–382.

Gray, Wayne, Sims, C. R., Fu, W.-T., and Schoelles, M. J. 2006. The soft constraints hypothesis: A rational analysis approach to resource allocation for interactive behavior. *Psychological Review*. 113(3): 461–482.

Gray, Wayne & Veksler, V. D. 2005. The acquisition and asymmetric transfer of interactive routines. *27Th annual meeting of the Cognitive Science Society*. Eds. B. G. Bara, L. Barsalou, and M. Bucciarelli. Austin (Texas): Cognitive Science Society

Gregory, Richard & Wallace, Jean. 1963. Recovery from early blindness: A case study. *Experimental Psychology Society*. Monograph no.2.

Haugeland, John. 1998. Mind Embodied and Embedded. *Having Thought: Essays in the Metaphysics of Mind*. Cambridge (MA): Harvard University Press. 207–237.

Hayhoe, Mary. 2000. Vision using routines: A functional account of vision. *Visual Cognition*. 7: 43–64

Heidegger, Martin. 1927 [1962]. *Being and Time*. Translated by Macquarrie, John & Robinson Edward. Oxford: Blackwell Publishers.

Hollan, James & Hutchins, Edwin. 2010. Chapter 9: Opportunities and Challenges for Augmented Environments: A distributed Cognition Perspective. *User Friendly Environments: From Meeting Rooms to Digital Collaborative Spaces*. ed. Lahlou Saadi. 237–259.

Hollan James, Hutchins Edwin & Kirsh David. 2000. Distributed Cognition: Toward a New Foundation for Human-Computer Interaction Research. *Transactions of Computer-Human Interaction*. 7(2): 174–196.

Huebner, Bryce. 2008. Do You See What We See? An Investigation of an Argument Against Collective Representation. *Philosophical Psychology*. 21(1): 91–112

Hurley, Susan. 1998. *Consciousness in Action*. Cambridge (Massachusetts): Harvard University Press.

– 2001. Perception and Action: Alternative Views. *Synthese*. 129: 3–40.

– 2010. Varieties of Externalism. *The Extended Mind*. ed. Menary, Richard. Cambridge (Massachusetts): A Bradford Book. The MIT Press. 101–154.

Hurley, Susan & Noë, Alva. 2003. Neural Plasticity and Consciousness. *Biology and Philosophy*. 18: 131–168.

Hutchins, Edwin. 1995. *Cognition in the Wild*. Cambridge (Massachusetts): The MIT Press.

– 2001. Distributed Cognition. *International Encyclopedia of the Social & Behavioral Sciences*. Eds. Smelser Neil & Baltes Paul. Amsterdam: Elsevier. 2068–2072.

– 2008. The role of cultural practice in the emergence of modern human intelligence. *Philosophical Transactions of The Royal Society B*. 363: 2011–2019.

– 2010. Cognitive Ecology. *Topics in Cognitive Science*. 2: 705–715.

– 2011a. Enaction, Imagination, and Insight. *Enaction: Toward a new paradigm for cognitive science*. Eds. Stewart John, Gapenne Oliver, and Di Paolo Ezequiel. Cambridge (Massachusetts): A Bradford Book. The MIT Press. 425–450.

– 2011b. Enculturating the Supersized Mind. *Philosophical Studies*. 152: 437–446.

Järvilehto, Timo. 1998a. The theory of organism–environment system I: Description of the theory. *Integrative Physiological and Behavioural Science*. 33(4): 321–334

– 1998b. The theory of organism–environment system II: Significance of Nervous Activity in the Organism-Environment System. *Integrative Physiological and Behavioural Science*. 33(4): 335–342

– 1999. The theory of organism–environment system III: Role of efferent influences on receptors in the formation of knowledge. *Integrative Physiological and Behavioural Science*. 34(2): 90–100

– 2000. The theory of organism–environment system IV: The problem of mental activity and consciousness. *Integrative Physiological and Behavioural Science*. 35(2): 35–57

– 2009. The Theory of the Organism-Environment System as a Basis of Experimental Work in Psychology. *Ecological Psychology*. 21(2): 112–120

Kant, Immanuel. 1881 [1922]: *Immanuel Kant's Critique of Pure Reason. In Commemoration of the Centenary of its First Publication.* Translated into English by F. Max Mueller (2nd revised ed.). New York: Macmillan

Kim, Jaegwon. 2006. *Philosophy Of Mind*. 2nd edition. Cambridge (Massachusetts): Westview Press.

Kirchhoff, Michael D. 2011. Extended cognition and fixed properties: steps to a third-wave version of extended cognition. *Phenomenology and the Cognitive Sciences*. Published online 18 October. DOI: 10.1007/s11097-011-9237-8. 1–22.

Kirsh, David. 1995. The Intelligent Use of Space. *Artificial Intelligence*. 73: 31–68.

– 2006. Distributed cognition: A methodological note. *Pragmatics & Cognition*. 14(2): 249–262.

Kirsh, David & Maglio, Paul. 1994. On Distinguishing Epistemic from Pragmatic Action. *Cognitive Science*. 18: 513–549.

Kiverstein, Julian. 2010. Sensorimotor knowledge and the contents of experience. *Perception, Action, and Consciousness: Sensorimotor Dynamics and Two Visual Systems*. Eds. Gangopadhyay Nivedita, Maday Michael & Spicer Finn. Oxford (UK): Oxford University Press.

– 2011. A social externalist account of cognitive agency, or why cognition isn't organism centred. Extended cognition workshop. 27.6.–28.6.2011. University of Amsterdam

Kiverstein Julian & Clark, Andy. 2009. Introduction: Mind Embodied, Embedded, Enacted: One Church of Many?. Topoi. 28:1–7.

Kiverstein Julian & Farina, Mirko. 2011. Embraining Culture: Leaky Minds and Spongy Brains. *Teorema*. 30(2): 35–53.

Knuuttila, Tarja. 2011. Modelling and representing: An artefactual approach to model-based representation. *Studies in History and Philosophy of Science*. 42: 262–271

Konner Melvin. 1980. Universals of Behavioral Development in Relation to Brain Myelination. *Brain Maturation and Cognitive Development: Comparative and Cross-Cultural Perspectives*. Eds. Gibson Kathleen R. & Petersen Anne C. Hawthorne (New York): Aldine de Gruyter.

Lakoff, George & Johnson, Mark. 1980. *Metaphors We Live By*. Chicago: University of Chicago Press.

Latour, Bruno. 1993. *We have never been modern*. Cambridge (Massachusetts): Harvard University Press.

Levin, Janet 2010. Functionalism. *Stanford Encyclopedia of Philosophy (Summer 2010 Edition).* ed. Zalta, Edward N. http://plato.stanford.edu/archives/sum2010/entries/functionalism/. Accessed in December 2011.

Maravita, Angelo & Iriki, Atsushi. 2004. Tools for the body (schema). *Trends in Cognitive Sciences*. 8(2): 79–86.

Marr, David. 1982. *Vision*. San Francisco: W.H. Freeman & Co Ltd.

Maturana, Humberto & Varela, Francisco. 1980. *Autopoiesis and cognition: the realization of the living*. Dordrecht: Reidel Publishing.

McBeath, M., Shaffer, D. and Kaiser, M. 1995: How Baseball Outfielders Determine Where to Run to Catch Fly Balls. *Science*. 268: 569–573.

Menary, Richard. 2006a. Attacking the Bounds of Cognition. *Philosophical Psychology*. 19(3): 329–344.

– (ed.). 2006b. *Radical Enactivism: Intentionality, Phenomenology and Narrative: Focus on the philosophy of Daniel D. Hutto*. Amsterdam: John Benjamins Publishing Company.

– 2006c. Introduction: What is radical enactivism?. *Radical Enactivism: Intentionality, Phenomenology and Narrative: Focus on the philosophy of Daniel D. Hutto*. ed. Menary Richard. Amsterdam: John Benjamins Publishing Company: 1–11.

– 2007. *Cognitive Integration: mind and cognition unbounded*. London: Palgrave Macmillan.

– 2009. Intentionality, Cognitive Integration and The Continuity Thesis. *Topoi*. 28: 31–43.

– 2010a. Cognitive Integration and the Extended Mind. *The Extended Mind*. ed. Menary, Richard. Cambridge (Massachusetts): A Bradford Book. The MIT Press. 227–243.

– 2010b. Dimensions of Mind. *Phenomenology and the Cognitive Sciences*. 9(4): 561–578.

– 2010c. The holy grail of cognitivism: a response to Adams and Aizawa. *Phenomenology and the Cognitive Sciences*. 9(4): 605–618.

Merleau-Ponty, Maurice. 1945 [2002]. *Phenomenology of Perception*. Translated by Smith, Colin. London: Routledge Classics.

Nieder A. & Dehaene S. 2009. Representation of number in the brain. *Annual Review of Neuroscience*. 32: 185–208.

Noë, Alva. 2004. *Action in Perception*. Cambridge (Massachusetts): The MIT Press.

– 2005. Against intellectualism. *Analysis*. 65(4): 278-290.

Norman, Donald. 1993. *Things That Make Us Smart: Defending Human Attributes in the Age of the Machine*. Reading (MA): Addison-Wesley Longman Publishing Company.

O'Regan, Kevin & Noë, Alva. 2001a. A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*. 24: 939–1011.

– 2001b. Authors' response: Acting out our sensory experience. *Behavioral and Brain Sciences*. 24: 1011–1031.

Port, Robert & Van Gelder, Tim (eds.). 1995. *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge (Massachusetts): A Bradford Book. The MIT Press

Prinz, Jesse. 2006. Putting the Brakes on Enactive Perception. *Psyche*. 12(1): 1–19.

Putnam, Hilary. 1975. The meaning if 'meaning'. *Language, Mind, and Knowledge*. ed. Gunderson K. Minneapolis: University of Minnesota Press.

Pöyhönen, Samuli. 2012 (draft). Explanatory turn in extended cognition. Paper presented in *Philosophy of science (POS) seminar*. 26.03.2012. University of Helsinki. 1–23.

Robbins, Philip & Aydede, Murat. 2009. A Short Primer on Situated Cognition. *The Cambridge Handbook of Situated Cognition*. Eds. Robbins, Philip & Aydede, Murat. New York (NY): Cambridge University Press. 3–10.

Rowlands, Mark. 1999. *The body in mind: Understanding cognitive processes*. Cambridge (UK): Cambridge University Press.

– 2006. *Body Language: Representation in Action*. Cambridge (Massachusetts): A Bradford Book. The MIT Press.

– 2009a. Extended cognition and the mark of the cognitive. *Philosophical Psychology*. 22(1): 1–19.

– 2009b. Enactivism and the Extended Mind. *Topoi*. 28: 53–62.

– 2009c. The Extended Mind. *Zygon*. 44(3): 628–641.

– 2010. *The New Science of the Mind: From Extended Mind to Embodied Phenomenology*. Cambridge (Massachusetts): A Bradford Book. The MIT Press.

Rumelhart, D. E., Smolensky, P., McClelland, J. L., and Hinton, G. 1986. Schemata and sequential thought processes in PDP models. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition.* Vol. 2: *Psychological and Biological Models*. Eds. McClelland J. L. & Rummelhart D. E. Cambridge (Massachusetts): MIT Press. 7–57.

Rupert, Robert D. 2004. Challenges to the hypothesis of extended cognition. *Journal of Philosophy*. 101(8): 389–428.

– 2005. Minding one's cognitive systems: When does a group of minds constitute a single cognitive unit?. *Episteme*. 1(3): 177–188

– 2009a. *Cognitive Systems and the Extended Mind*. New York (NY): Oxford University Press.

– 2009b. Innateness and the situated mind. *The Cambridge Handbook of Situated Cognition*. Eds. Robbins, Philip & Aydede, Murat. New York (NY): Cambridge University Press. 96–116.

– 2010a: Representation in Extended Cognitive Systems: Does the Scaffolding of Language Extend the Mind?. *The Extended Mind*. ed. Menary, Richard. Cambridge (Massachusetts): A Bradford Book. The MIT Press. 325–353.

– 2010b. Extended cognition and the priority of cognitive systems. *Cognitive Systems Research*. 11: 343–356.

Sacks, Oliver. 1995. *An Anthropologist on Mars: Seven Paradoxical Tales*. New York (NY): Knopf.

Saloman, Gavriel (ed.). 1993. *Distributed Cognitions: Psychological and Educational Considerations. Learning in Doing: Social, Cognitive, and Computational Perspectives*. New York (NY): Cambridge University Press.

Sellars, Wilfrid. 1956 [1963]. Empiricism and the Philosophy of Mind. *Science, Perception and Reality*. London: Routledge & Kegan Paul. 127–196.

Shanahan, Murray. 2009. The Frame Problem. *Stanford Encyclopedia of Philosophy (Winter 2011 Edition).* ed Zalta, Edward N. http://plato.stanford.edu/archives/win2009/entries/frame-problem/. Accessed in December 2011.

Shapiro, Larry. 2007. The Embodied Cognition Research Programme. *Philosophy Compass.* 2(2): 338–346.

– 2010. *Embodied Cognition*. London (UK): Routledge.

Sprevak, Mark. 2009. Functionalism and extended cognition. *Journal of Philosophy*. 106: 503–527.

– 2010. Inference to the hypothesis of extended cognition. *Studies in History and Philosophy of Science*. 41: 353–362.

Sterelny, Kim. 2010. Minds: Extended or Scaffolded?. *Phenomenology and the Cognitive Sciences*. 9(4): 465–481.

Stewart J., Gapenne O., and Di Paolo E (eds.). 2011. *Enaction: Towards a New Paradigm for Cognitive Science*. Cambridge (Massachusetts): The MIT Press.

Sutton, John. 2006. Distributed cognition: Domains and dimensions. *Pragmatics & Cognition*. 14(2): 235–247.

– 2008. Chapter 3: Material Agency, Skills and History: Distributed Cognition and the Archaeology of Memory. *Material Agency: Towards a non-anthropocentric approach*. Eds. Knappett C. & Malafouris L. Accessed from http://www.phil.mq.edu.au/staff/jsutton/SuttonMaterialAgency.pdf. Accessed in February 2012. 37–55.

– 2010. Exograms and Interdisciplinarity: History, the Extended Mind, and the Civilizing Process. *The Extended Mind*. ed. Menary, Richard. Cambridge (Massachusetts): A Bradford Book. The MIT Press. 189–225.

Sutton John, Harris Celia B., Keil Paul G., Barnier Amanda J. 2010. The psychology of memory, extended cognition, and socially distributed remembering. *Phenomenology and the Cognitive Sciences*. 9: 521–560.

Thelen, Esther. 1995. Time-Scale Dynamics and the Development of an Embodied Cognition. *Mind as Motion: Explorations in the Dynamics of Cognition*. eds. Port Robert & Van Gelder Tim. Cambridge (Massachusetts): A Bradford Book. The MIT Press. 69–100.

Thelen Esther, Fisher Donna M., Ridley-Johnson Robyn, Griffin Nora J. 1982. The effects of body build and arousal on newborn infant stepping. *Development Psychobiology*. 15(5): 447–453.

Thelen Esther, Fisher Donna M., Ridley-Johnson Robyn. 1984. The relationship between physical growth and a newborn reflex. *Infant Behavior and Development*. 7(4): 479–493.

Thelen, Esther & Smith, Linda. 1994. *A Dynamic Approach to the Development of Cognition and Action*. Cambridge (Massachusetts): The MIT Press.

Thelen, Esther & Ulrich, Beverly D. 1991. Hidden skills: A dynamic system analysis of treadmill stepping during the first year. *Monographs of the Society for Research in Child Development*. No. 223. 56(1).

Thompson, Evan. 2004. Life and Mind: From autopoiesis to neurophenomenology. A tribute to Francisco Varela. *Phenomenology and the Cognitive Sciences*. 3. 381–398.

– 2005. Sensorimotor subjectivity and the enactive approach to experience. *Phenomenology and the Cognitive Sciences*. 4(4): 407–427.

– 2007. *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge (Massachusetts): The Belknap Press of Harvard University Press.

Thompson, Evan & Stapleton, Mog. 2009. Making Sense of Sense-Making: Reflections on Enactive and Extended Mind Theories. *Topoi*. 28: 23–30.

Thompson, Evan & Varela, Francisco. 2001. Radical embodiment: neural dynamics and consciousness. *Trends in Cognitive Sciences*. 5(10): 418–425.

Tollefsen, Deborah. 2002a. Collective Intentionality and the Social Sciences. *Philosophy of the Social Sciences*. 32(1): 25–50

– 2002b. Organizations as True Believers. *Journal of Social Philosophy*. 33(3): 395–410

– 2006. From extended mind to collective mind. *Cognitive Systems Research*. 7: 140–150

Torrance, Steve. 2005. In search of the enactive: Introduction to special issue on Enactive Experience. *Phenomenology and the Cognitive Sciences*. 4(4): 357–368.

Tuomela, Raimo. 1977. *Human Action and its Explanation*. Dordrecht: D. Reidel

– 2003. "The We-Mode and the I-Mode," in *Socializing Metaphysics: The Nature of Social Reality*. ed. Frederick Schmitt. Lanham: Rowman & Littlefield: 93–127.

Valvo, Alberto. 1971. *Sight restoration after long-term blindness: The problems and behavior patterns of visual rehabilitation*. New York (NY): American Federation for the Blind.

Van Gelder, Tim. 1995. What Might Cognition Be, If not Computation?. *The Journal Of Philosophy*. 92(7): 345–381.

Varela, Francisco. 1979. *Principles of Biological Autonomy*. New York (NY): Elsevier North Holland.

– 1997. Patterns of life: intertwining identity and cognition. *Brain and Cognition*. 34: 72–87.

Varela Francisco J., Thompson Evan, Rosch Eleanor. 1991 [1993]: *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge (Massachusetts): The MIT Press.

Velleman, David. 1997. How To Share An Intention. Philosophy and Phenomenological Research. 57(1): 29–50

von Uexküll, Jakob. 1934. A Stroll Through the Worlds of Animals and Men: A Picture Book of Invisible Worlds. *Instinctive Behavior: The Development of A Modern Concept*. Ed. Schiller, Claire H. New York: International Universities Press. 5–80.

Von Wright, George Henrik. 1963. *Norm and Action*. London: Routledge & Kegan Paul.

Vygotsky, Lev. 1934 [1986]. *Thought and Language*. Translated, revised and edited by Kozulin, Alex. Cambridge (Massachusetts): The MIT Press.

– 1979. *Mind in Society*. Cambridge (Massachusetts): Harvard Press.

Walter, Sven. 2010: Cognitive extension: the parity argument, functionalism, and the mark of the cognitive. *Synthese*. 177: 285–300.

Walter, Sven & Kästner, Lena. 2012. The where and what of cognition: The untenability of cognitive agnosticism and the limits of the Motley Crew Argument. *Cognitive Systems Research*. 13: 12–23.

Ward Dave, Roberts Tom & Clark Andy. 2011. Knowing what we can do: actions, intentions, and the construction of phenomenal experience. *Synthese*. 181: 375–394.

Ward, Dave & Stapleton, Mog. To appear. Es are Good: Cognition as Enacted, Embodied, Embedded, Affective And Extended. *Consciousness in interaction: The role of the natural and social environment in shaping consciousness*. Eds. Paglieri, F & Castelfranchi, C. Part of the John Benjamins series *Advances in Consciousness*Research. Accessed from http://edinburgh.academia.edu/MogStapleton/Papers/657979/Es_are_Good_Cognition_as_Enacted_Embodied_Embedded_Affective_and_Extended. Accessed in January 2012. 1–22.

Webb, Barbara. 1994. Robotic experiments in cricket phototaxis. *From animals to animals 3: Proceedings of the Third International Conference on the Simulation of Adaptive Behaviour*. (Eds.) Cliff D., Husbands P., Meyer J. & Wilson S. Massachusetts: MIT Press. 45–54.

Weber Andreas, & Varela, Francisco. 2002. Life after Kant: natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences*. 1: 97–125.

Wheeler, Michael. 2005: *Reconstructing the Cognitive World: The Next Step*. Cambridge (Massachusetts): A Bradford Book. The MIT Press.

– 2010. In Defense of Extended Functionalism. *The Extended Mind*. ed. Menary, Richard. Cambridge (Massachusetts): A Bradford Book. The MIT Press. 245–270.

– 2011. In search of clarity about parity. *Philosophical Studies*. 152: 417–425.

– to appear. Minds, Things and Materiality. *The Cognitive Life of Things: Recasting the Boundaries of the* Mind. eds. Renfrew C. and Malafouris L. Cambridge (UK): McDonald Institute for Archaeological Research Publications. 1–14.

– manuscript. *Extended X: Recarving the Biological and Cognitive Joints of Nature*. Accessed from http://www.philosophy.stir.ac.uk/staff/m-wheeler/ExtendedX.php. Accessed in January 2012.

Wilson, Robert A. 1994. Wide Computationalism. *Mind*. 103(411): 351–372.

– 2004. *Boundaries of the Mind: The Individual in the Fragile Sciences*. Cambridge (UK): Cambridge University Press.

Wilson, Robert A. & Clark, Andy. 2009. How To Situate Cognition: Letting Nature Take its Course. *The Cambridge Handbook of Situated Cognition*, ed. Aydede M. & Robbins P. New York: Cambridge University Press. 55–77.

Wilson, Robert A. & Foglia, Lucia. 2011. Embodied Cognition. *Stanford Encyclopedia of Philosophy (Fall 2011 Edition)*. ed. Zalta, Edward N. http://plato.stanford.edu/archives/fall2011/entries/embodied-cognition. Accessed in November 2011.

Wittgenstein, Ludwig. 1953. *Philosophical investigations*. Oxford: Blackwell.

Zelazo, Philip. 1984. The Development of Walking: New Findings and Old Assumptions. *Journal of Motor Behavior*. 15: 99–137.