Spring 1-1-2016

Action, Agents, and Extended Cognition

Rebecca Moore Renninger
University of Colorado at Boulder, rebecca.renninger@gmail.com

Action, Agents, and Extended Cognition

by

REBECCA MOORE RENNINGER

B.A., Davidson College, 2008

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Doctor of Philosophy

Department of Philosophy

2016

This thesis entitled:
Action, Agents, and Extended Cognition
written by Rebecca Moore Renninger
has been approved for the Department of Philosophy

_____
Robert D. Rupert

_____
Carol Cleland

Date_____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Renninger, Rebecca Moore (Ph.D., Philosophy)

Action, Agents, and Extended Cognition

Thesis directed by Professor Robert D. Rupert

Some authors in philosophy of mind have sought to defend a theory according to which cognitive processing and/or mental states substantively extend beyond the bounds of an individual's body. The Extended Mind Thesis (EMT) asserts that the vehicles of cognition and mental states exist at least partly outside of the human thinker. This dissertation introduces and defends one new objection to EMT. My focus here will be on what I call the agency problem for the extended mind, which asserts that EMT cannot account for intentional action or human agency. I shall argue that EMT cannot explain genuine intentional actions, the ability to perform which is a central part of our identity as human persons. If we believe EMT, then we give up our last grasp of causal sovereignty when it comes to our actions. I use these arguments to motivate a conservative view of the mind as an embodied, but not substantively extended, entity.

For My Family

Many thanks to Carol Cleland, Graham Oddie, Iskra Fileva, Clayton Lewis, and especially Rob Rupert for all of their help and insight.

Table of Contents

Preface:

The issue of what constitutes the mind has received plenty of philosophical attention historically and in contemporary philosophy of mind. Recently, some philosophers and cognitive scientists have argued that the mind, or the cognitive system that makes it up, extends beyond the boundaries of the body. The mind, these authors argue, is constituted by not only the brain, but also elements of the non-neural body and even the external world. What we once thought were cognitive aids out there in the world (calculators, notebooks, smartphones, etc.) are actually parts of cognitive systems. It is this position that I take as an object of criticism in this dissertation. While this extension of the mind out into the world has a certain motivational pull, this ought to be resisted.

The details of the extended mind thesis (EMT) have been developed somewhat thoroughly over time, and they continue to develop in interesting and enticing ways. And, the position has been criticized in various ways. However, what is lacking in both the development of the position and its criticisms is a full exploration of personal identity and agency. The problem being addressed in this dissertation is not only what constitutes the mind, but also the implications that mental constitution has for agential responsibility and self-constitution. The intersection of these issues has been given relatively little attention, but its importance should not be understated. Issues having to do with agency and personal identity are integral to a full examination of the extended mind theory, and should be considered integral to any theory of mind and cognition. However, I argue here, a fuller examination of these issues leads us to recognize irreparable problems with the extended theory.

The main problem addressed in what follows in this text is what I refer to as the *Agency Problem for the Extended Mind*. The problem, explicated briefly, is that the extended mind theory forces us to give up our causal sovereignty when it comes to our actions, and therefore renders us not agentially responsible for them. The mental states that cause our actions typically contribute to our agential control because these mental states are internal to the agent who acts. However, when these mental states are not internal to the agent but rather extended out into the world, our own agential control is superseded by worldly enterprises. Because our minds are extended beyond the boundaries of our bodies, the control of our actions is causally outsourced to whatever external elements partially make up or constitute our mental states.

Secondarily, this dissertation addresses the issue of embodiment and what theory of embodiment is most satisfactory. If the mind is not extended, what relation does it have to the body and the external world? I put forth a view of embodiment that connects our reasons for action, and therefore control over our actions, to not only the mind but also the body. This is a relatively new approach, which aims to synthesize literature relating to the mind and literature relating to causal theories of action. This part of the dissertation functions to bring together two philosophical topics that are typically isolated – philosophy of mind and action theory. This relatively unchartered philosophical territory provides an interesting secondary unit of inquiry for the dissertation.

Ultimately, I conclude that the extended mind position is incompatible with any robust view of agency and therefore ought to be rejected. This limits the mind to being constituted by body-bound elements. Agency is integral to our constitution as persons, and it is not something that we can reasonably give up in order to preserve an extended view of the mind. Further, extended theories of mind themselves lean heavily on the idea that cognition is an active agential

process. If the extended mind theory is metaphysically incompatible with this very idea, then the theory includes a deep inconsistency.

The dissertation starts by examining the extended mind position followed by what I take to be a standard reductive view of agency. This Davidonian-causal account is compatible with both extended and non-extended theories of mind, and thus presents a neutral starting point for an examination of action throughout the dissertation. I proceed to examine various alternative views of action and explicate why these are either incompatible with an extended view of the mind or why they fail to solve the problem at hand. I argue that no extended view can straightforwardly avoid the agency problem from the start. I then examine possible remedies to the problem, which I reject in turn. The first is the possibility of providing an internal locus of control for action while keeping the mind extended. According to this idea, the agent or self can be bounded, while the mind is extended. This position most straightforwardly leads to the agency problem, and it may seem incompatible with the idea of the extended mind itself, as some extended theorists argue that the self must be extended if the mind is extended. The second is the possibility that the self is extended. If the self is extended out to the boundaries of the mind, then the mental states that cause our actions can be extended but still remain proper parts of the agent, thus preserving agential control and responsibility. I argue against this position at some length. The third possibility for the extended mind theorist is to give up a personal ontology that includes an agential self. I reject this as not only counterintuitive but also as deeply incompatible with the extended mind position and therefore not a viable option to escape the problem at hand.

Later in the dissertation I survey views of embodiment and develop an embodied view of the mind that, while not extended, preserves the idea that the mind and body are intimately connected. In the end, I tie together my initial examination of a reductive causal account of

agency and an embodied view of the mind, arguing that the two dovetail together. Thus, this embodied view of the mind preserves many virtues of embodiment without sacrificing agential responsibility, which is integral to our identity as persons. These sections conclude the dissertation. Thus, my aim is to show not only that the extended mind theory is incompatible with agency, and that it suffers as a result, but also to give a positive alternative view of embodiment and provide some new theory as to how this view of the mind, and philosophy of mind more generally, can fundamentally intersect with issues relating to action theory and agential responsibility.

Chapter 1: Reasons Causalism and The Agency Problem for the Extended Mind

§1.1: Introduction

Traditionally, the boundaries of the mind have been thought to be internal to the human person, in that the mind, however it is constituted, is made up of things internal to a person's body. This conception comes in part from the Cartesian tradition, which includes the idea that the mind is something purely inner. The notion that the brain contains the vehicles of mental states has until somewhat recently been widely accepted. This belief, though, has been challenged by philosophers who theorize that mental states and cognitive processing may loop out into the world or even be largely constituted by objects outside of a person's brain and body. This hypothesis, which I will call the Extended Mind Thesis (EMT), has gained a philosophical following.[1] Many theoretical consequences of EMT, however, remain unexplored. I will argue that one of these consequences is that when the human mind outstrips the body of an agent, she loses her power to execute genuine intentional actions. Once the mind extends out into the world, the production of our actions becomes outsourced, no longer depending on our agential outputs, but instead constitutive elements from the world.

§1.2: Overview and Chapters Outline

In Chapter 1, I proceed by giving an overview of the basic theories of action and the mind that I will be addressing throughout this dissertation. Chapter 1 focuses on the broadly Davidsonian[2] causalist account that is prevalent in the philosophical literature on action, and the

[1] There are many versions of the Extended Mind Thesis, a selection of which I will review in the following chapters.

[2] For a concise explanation of this account see Davidson, Donald (1963). Actions, Reasons, and Causes. *Journal of Philosophy* 60, 23: 685-700.

view of the extended mind originally put forth my Clark and Chalmers (1998) in their paper on the subject. In the end, I introduce and give a brief sketch of the problem that this work seeks to shed light on – the agency problem for the extended mind.

Chapter 2 gives a broader taxonomy of theories of action. This includes both causal and non-causal theories of action. I argue that these views when combined with the extended mind theory either do not avoid the agency problem or are themselves problematic. This may be because the view is inherently problematical and lacking in merit sufficient to make it plausible in its own right or because the view does not respect the naturalistic commitments of EMT. This leaves the extended mind theorist with a problem of how to reconcile an extended view of the mind with an acceptable account of agency.

In Chapters 3 and 4, I address possible solutions to this problem including the possibility of re-bounding an internal locus of control for action while keeping the extrapersonal extension of mental states, and the possibility that the agent, or the self, is extended. In Chapter 3, I review multiple versions of the extended mind thesis that aim to retain an individualized locus of control for action while still extending cognition. I argue that these views do not solve the agency problem at hand. An internalized agent or locus of control for action with externalized mental states may even make the problem more salient. In Chapter 4, I argue against the plausibility of the idea of an extended self at some length. And, I argue that any extended mind theory the natural extension of which is extended agents ought to be rejected. If the thinking agent is not extended, though, and the mind is, she becomes inert and unable to perform genuine actions. So, the extended mind theorist has two choices, both of which fail. Either the agent is extended, and can act, in which case radically implausible consequences follow, or the self is narrow and

internally bounded while the mind extends into the environment, in which case it cannot perform truly intentional actions.

One of the main ideas addressed in these chapters is that some authors defending an extended view of the mind have stated that EMT entails a personal ontology that is also extended. This would be one way in which the extended mind theorist could try to obviate the agency problem. There are, in fact, some reasons to think that if we accept EMT we also ought to accept an extended version of the self. However, the extended self thesis is implausible and, I argue, irreparably problematic. So, not only will this solution to the agency problem fail, this may in itself constitute one additional reason to reject EMT. In attempting to solve the agency problem using an extended theory of the self, the extended mind theorist will in doing so create further problems for her own theory.

Chapter 5 defends the idea that both agency and the self are important features of our personal ontology that cannot be abandoned by the extended mind theorist. This chapter functions in part to motivate the idea that this problem is not one that the extended theorist can get around simply by rejecting the existence of agents out of hand. I argue that agency is an important part of an individual, and it is a part that cannot plausibly be eliminated. I also argue that agents as a whole cannot be eliminated from the metaphysical apparatus necessary for supporting the extended mind theory. The chapter concludes that it would be wrong for the extended mind theorist simply to jettison agency as a feature of human persons.

Chapter 6, finally, uses the above considerations to motivate a bounded view of the mind and the self. The view that I argue for is of a mind that is embodied, as well as embedded in and scaffolded by features of our environment in important ways. It is not, however, substantively extended. The moderate approach to embodiment I give focuses on the neural redeployment of

bodily representations. This view of the mind can account for important benefits of EMT and the overall importance of the body without causing us to lose our grip on the most central agential features of our selves.

In Chapter 7 I continue this examination of embodiment by exploring its implications for actions (embodied actions). We are not only embodied agents according to this view; our actions are owned by us in a deep sense as they are embodied by us as agents. Action production is a form of embodied cognition. In this chapter I thoroughly examine what it means for an action to be agentially controlled and what it means for an agent to be responsible (in the metaphysical sense) for her actions. I argue that the mechanism that produces the action in question must be properly owned by an agent, or part of an agent's constitution. And, the previously examined theory of embodiment gives us just this kind of ownership. Finally, this chapter ties together the Davidsonian account of action explicated earlier (in Chapter 1) and this neural account of embodiment, showing how an agent can embody the reasons that cause her actions.

In conclusion, I offer some prospects for further research in these areas. The final section offers some concluding thoughts meant to tie together these arguments, motivating a rejection of the extending mind thesis, and suggesting future research in this area. The preceding arguments support theories of mind according to which the mind is scaffolded, embedded, and embodied, but not substantively extended out into the world. The world is not poised to participate in the cognitive machinery of the person as the extended mind thesis suggests. We, as persons and organisms, are cognitive agents that act on the world not from apart, but from within.

§1.3: Internal and External Theories of Mind

The standard view of the mind holds that mental states are internal to the brain and/or body of the human organism. According to traditional internalism, mental states, while they may gain their meaning by virtue of external entities, are wholly located inside of the head (or, maybe, bounded by the body). The physical individual is conventionally seen as sacred when it comes to her mental functioning. Mental states, at least to the extent that they are reducible to or constituted by physical states, are not constituted by features outside of us.[3]

Internalism, or individualism about the mind asserts that mental states either supervene on or are realized by the internal, intrinsic, physical properties of the individual who bears them. These mental states are fully constituted by the internal physical states of an individual, ordinarily conceived of as brain states and/or states of the body. What separates individualists and extended theorists about the mind is the question of whether there is a more direct sense in which the nature of the mind is determined by an individual's environment.[4] The individualist may still hold that the environment makes a difference in the content of our mental states in that a difference in the world can make a corresponding difference to what occurs inside of the mind. However, the environment is merely causally determinative in this context. The entirety of what determines the nature of our mental activity in the metaphysical or constitutive sense is what exists inside of the boundaries of the individual.

If mental states are constitutively determined by internal features of the person, and are thus constitutively autonomous from the external world, the realization base of mental properties and processes is wholly internal to the individual. Her intrinsic, physical properties are determinative of her mental states and metaphysically sufficient for any mental properties that

[3] As most extended mind theorists do, I will be assuming a physicalist stance when it comes to the mind.

[4] For more on this division, see Wilson, Robert (2004). *Boundaries of the Mind: The Individual in the Fragile Sciences*. New York: Cambridge University Press.

they realize. External features of the environment may causally affect these realizers, but they are not constitutive of them. So, one's perceptions of the world, say, seeing a cat, may cause one to have certain occurrent mental states. For instance, you might form the belief that there is a cat in the room or the desire to pet the cat nearby. The cat itself is instrumental to your forming this belief, but such instrumentality is exhaustive of its role. These mental states are merely caused by one's environment, through one's perceptual faculties. The physical realizers of our mental states are fully internal to the thinking agent – most likely located in the brain, and possibly extended to a limited degree throughout the rest of the body.

Externalism about meaning is compatible with a bounded view of the mind. The meaning of a term may be determined by factors external to the speaker without beliefs involving that term being constituted by external features of the world. In twin cases, for instance, there are external factors responsible for differences in the meaning of terms used by twins in different circumstances.

> When I believe that water is wet and my twin believes that twin water is wet, the external features responsible for the difference in our beliefs are distal and historical, at the other end of a lengthy causal chain…In these cases, the relevant external features are *passive*. Because of their distal nature, they play no role in driving the cognitive process in the here-and-now.[5]

The contents of our mental states can gain their meanings by virtue of external features of the world, possibly historically and causally. However, the external environment does not play a constitutive role in the cognitive processing that leads to mental states or behavior. According to this type of semantic externalism, the internal structure of our minds does the crucial work of generating beliefs, desires, behavior, etc. The environment does not constitute the vehicles of these mental states and processes. While the meaning associated with the content of our mental

[5] Clark, Andy and Chalmers, David (1998). The Extended Mind. *Analysis* 58: 7-19. 9.

states may have external determinants, the vehicles of that content remain internal to the thinking subject.

Semantic externalism is a thesis about the individuation of our mental states, and as such it is not typically seen as a thesis about the material realizers of their content. As elucidated by Andreas Elpidorou (2012):

> It holds that the contents of some mental states of a subject *S* depend for their individuation on factors external to *S*. Assuming that there are mental states that possess their content essentially, then semantic externalism turns out to be a thesis about the *individuation* of those mental states. And although it can also be understood as a thesis about the *constitution* or *location* of mental states (given certain assumptions about the nature of mental content), typically it is not. As a thesis about mental or cognitive states, semantic externalism does not commit one to an externalist thesis concerning the nature of mental or cognitive *processes*.[6]

The contents of mental states may depend on external factors while the physical states that realize these mental states remain internal to the individual. Cognitive processing may depend upon external items, but these items are not constitutive of its vehicles.

The question, then, is whether or not the vehicles of content, or the material aspects that enable an individual to possess (occurrently or dispositionally) a given mental state are extended beyond the bounds of that individual.[7] Processes and structures that lie outside of the human brain and body can realize these mental states, at least in part. The goal of many extended theories of mind is to show not that the content of our mental states could be context sensitive, but that the vehicles of these mental states could themselves be externalized. The extended mind thesis (EMT) claims that the environment partially constitutes the vehicles of an agent's mental states or cognitive processing. If the mind does not extend, the world may still play a causal role in the production of mental states and cognitive processing, but it does not constitute the material

[6] Elpidorou, Andreas (2012). Where is My Mind? *Avant* 3, 1: 145-160. 145.
[7] Clark, Andy (2008b). *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. New York: Oxford University Press. 76.

realizers of those states or processes. These vehicles are the physical states that realize our mental states, and, according to EMT, these vehicles are spatiotemporally located outside of the head. The distinction between constitution and causation is central to the debate over the extended mind. If the environment plays a merely causal role in the production of mental states and processes, then the material vehicles of cognition remain within the boundaries of the individual. However, if the environment partially constitutes our mental states and processes, then these states extend outwardly beyond the body. While traditional accounts of mind and cognition agree that external causes can influence the content of mental states, EMT argues that the web of environmental relations in which a subject is embedded can function as mental entities themselves.

Something that is part of an extended cognitive system must be more than a cognitive aid; it must be coupled to an individual such that they make up one such cognitive system. Hence, the main burden of proof for the extended mind theorist is to show that some external items are at least partially constitutive of a cognitive system, rather than being mere external causally efficacious or causally relevant instruments used by a cognitive agent. The Parity Thesis, sometimes used to support this, starts with the claim that there are parts of the world that are functionally equivalent to internal parts of the mind. These items are poised to have the same type of interaction with recognized parts of the mind as those parts of the mind have with other (internal) mental entities. Therefore, these parts of the world themselves can plausibly be considered parts of the mind as well. Clark (2010) uses the Parity Thesis to motivate an extended view of the mind:

> 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 accepting as part of the cognitive process, then that part of the world is (for that time) part of the cognitive process.[8]

This principle is meant to pave the way for an extended view of the mind by guarding against internalist prejudices. The worldly process itself need not be like anything actually physically realized the human brain. It only needs to be the case that if (counterfactually) the process were done in the head we would treat it as cognitive.

Clark states that it is his own suspicion that the differences between external processes and inner ones will be no greater than the differences between differing inner processes themselves. This kind of extended systemic coupling occurs when there is an external informational resource that is constant, directly available, and automatically endorsed. All of the components of a coupled system play an active causal role, including a role in the governance of behavior.

Though the distinction is not always made explicit, there are modest differences between extended cognition and the extended mind. Drayson (2010) states that the difference between extended cognition and the extended mind is, at least presumably, derivative on the distinction between cognition and mind itself. Mental states are often thought of as the "ascriptions of 'folk-psychology'."[9] These are the commonsense states that we use to describe, predict, and explain our human behavior. Beliefs and desires are some archetypical examples of these. Mental concepts are also often associated with our concepts of the self, personhood, free will, and responsibility. Cognitive science, on the other hand, explains behavior not by associating it with

[8] Clark, Andy (2010). Memento's Revenge: The Extended Mind, Extended. In R. Menary (ed.), *The Extended Mind* (pp. 43-66). Cambridge, MA: MIT Press. Also see Clark and Chalmers (1998), 8.
[9] Drayson, Zoe (2010). Extended Cognition and the Metaphysics of Mind. *Cognitive Systems Research* 11, 4: 367-377. 374.

mental states but by attempting to understand the underlying mechanisms that lead to it. Cognition can be identified with these mechanisms.

Drayson (2010) states that cognitive science expands upon the traditional notion of the mental by studying not just what it consciously accessible, but also the unconscious processes that lead to these consciously accessible states and to the behavior that we associate with minded creatures. Cognitive science "is not just interested in the consciously accessible states we can attribute to a person, but also the inaccessible states we can only attribute to parts of a person's cognitive system."[10] This may not be entirely on point, as there are times that we attribute unconscious, dispositional, or nonoccurrent mental states to persons, which also serve to explain human behavior. Overall, though, the scope of the cognitive may be broader than what we ordinarily think of as the mental, including in its scope operations that underlie what we ordinarily understand as mental phenomena. The term 'cognition' has come to be used to refer to the mechanisms responsible for mental phenomena. Cognition encompasses traditional mental phenomena as well as the lower-level information processes that account for them.[11]

There are various ways that cognition and the mind might relate. Cognitive processes may be the precursors to genuine mentality. They may be responsible for the phenomena that we refer to as 'mental' phenomena. Perhaps, the mind is a collection of properties that 'emerge' from cognition.[12] When it comes to extended cognition and the extended mind, one might claim to show that cognition is extended "while claiming that the relationship between mind and cognition is one which entails that if cognition is extended, then mind is extended."[13] Or, at the very least, extended cognition is necessary for the extended mind. One way to do this is by

[10] Drayson (2010), 375.
[11] Ibid.
[12] Ibid., 377.
[13] Ibid.

stipulation, claiming that the mind is simply whatever cognitive science determines it to be. However, this would not have any particularly interesting philosophical implications for ideas about, for example, the self or personhood.[14]

The mind is extended just in case mental states are constituted by objects or features of our environment that are external to our bodies. When this happens the world constitutes our mental states, rather than the processing that might give rise to or affect these states. Cognition is extended if and when we utilize parts of the world in order to perform cognitive tasks. When cognition extends, the cognitive agent is coupled with the external world in such a way as to extend the cognitive system in its own right. These systems are often said to be nondecomposable or uncoupleable in that their behavior is not explainable in terms of smaller individualizable parts. Because mental states are at times causally or functionally defined, occurrent mental states often seem to dissolve into cognitive processes themselves. If all the mind *is* is a collection of cognitive processes, then extended cognition is sufficient for the extended mind thesis to be true.

§1.4: Clark and Chalmers: Active Externalism

Clark and Chalmers, in their original paper "The Extended Mind," give an outline of a view that they call *Active Externalism*. According to Active Externalism, the relevant external features in a person's environment are *active*, playing a crucial role in the here-and-now. These features are coupled with a cognitive agent, as they have a direct impact on her cognition and her behavior. And, without them, the person's capacities would be limited. Clark and Chalmers state, "in these cases, the relevant parts of the world are *in the loop*, not dangling at the other end of a

[14] Drayson (2010), 377.

long causal chain."[15] These external features play an ineliminable role in cognitive processing and in the production of our behavior.

The authors give three cases to illustrate the extension of cognitive processing. In the first case, a person sits in front of a computer screen playing a game similar to Tetris in which she must mentally rotate shapes to fit them into the appropriate sockets. In the second case, the person has the option of physically rotating the shapes on the screen using a button. In the third case, taking place some time in the future, she sits in front of a similar computer screen. This time, she has the option of using a neural implant (implanted in her brain) to rotate the shapes, or, like the previous example, she has the option of mentally rotating the images.

Clark and Chalmers ask how much cognition is present in these cases. They state that "Case (3) with the neural implant seems clearly to be on a par with case (1). And case (2) with the rotation button displays the same sort of computational structure as case (3), distributed across agent and computer instead of internalized within the agent."[16] If the third case (in which the neural implant is implemented) is cognitive, then we need a justification for denying that the second is as well. The boundaries of skin and skull are not an appropriate justification for denying this (as the validity of these boundaries is exactly what is at issue for the authors). However, nothing else between the cases seems relevantly different. Thus, cognition must be extended in the second case across the boundary of the agent and be distributed between the agent and the computer. This is one example of an epistemic action, which alters the world so as to aid and augment cognitive processes.

The authors then go on to assert that such actions deserve epistemic credit. They emphasize the previously mentioned Parity Thesis, which tells us that "if, as we confront some

[15] Clark and Chalmers (1998), 9.
[16] Clark and Chalmers (1998), 7.

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*…part of the cognitive process."[17] Case (2) is just such a case. Part of the world in this case is functioning as a process which, were it done in the head as in the other case, we would have no hesitation recognizing as cognitive.

Clark and Chalmers additionally argue that *mental states* – specifically beliefs – can be located outside of the head. As with cognition, the mind itself branches out into the external world. This is what Clark and Chalmers call the extended mind. The authors give the case of Otto in order to demonstrate the way in which the world might constitute some of our mental states.

> Otto suffers from Alzheimer's disease, and like many Alzheimer's patients, he relies on information in the environment to help structure his life. Otto carries a notebook around with him everywhere he goes. When he learns new information, he writes it down. When he needs some old information, he looks it up. For Otto, his notebook plays the role usually played by a biological memory. Today, Otto hears about the exhibition at the Museum of Modern Art, and decides to go see it. He consults the notebook, which says that the museum is on 53rd Street, so he walks to 53rd Street and goes into the museum.[18]

Otto's notebook plays the same role that (ordinarily) someone's nonoccurrent beliefs would play, in that what is written in the notebook can be used in order to inform his behavior. Since these roles are isomorphic, Clark and Chalmers contend that there should be nothing to keep us from giving the information stored in the notebook the label of a belief. Going back to the Parity Thesis, the notebook and Otto form a system that behaves in a way such that it cannot be called substantially different from the system of biological memory ordinarily captured within the bounds of skin and skull. Otto is coupled to his notebook in the same manner in which we are ordinarily coupled to the information stored in biological memory. When we need to remember

[17] Clark and Chalmers (1998), 8.
[18] Ibid., 12-13.

something, we simply consult our memories as Otto consults his notebook. The authors conclude that there is nothing sacred, when it comes to the mind, about the "boundaries of skin and skull."[19] What gives a mental state its status is the role that it plays – not the substance of which it is composed or the location of that substance.

Clark and Chalmers suggest, at the end of the paper, that the self *itself* extends beyond the bounds of the body.

> Most of us already accept that the self outstrips the boundaries of consciousness; my dispositional beliefs, for example, constitute in some deep sense part of who I am. If so, then these boundaries may also fall beyond the skin. The information in Otto's notebook, for example, is a central part of his identity as a cognitive agent. What this comes to is that Otto *himself* is best regarded as an extended system, a coupling of biological organism and external resources. To consistently resist this conclusion, we would have to shrink the self into a mere bundle of occurrent states, severely threatening its deep psychological continuity. Far better to take the broader view, and see agents themselves as spread into the world.[20]

The issue of whether the thinking individual herself ought to feature this form of extension is somewhat contentious. Clark and Chalmers put forth the idea here, which I will address at more length later on, that the individual who has mental states must be wherever these mental states are. This holds especially true when these states are a central part of her cognitive identity, though which states would be identity constituting remains a standing question. Issues of personal identity and the extended self thesis turn out to be a critical part of EMT a broader examination of which I will give later in this dissertation.

Clark argues that mental states can be realized, at least in part, by structures and processes outside of the head. Clark calls his theory, simply, EXTENDED. Clark's theory relies on a distinctive kind of embodiment, in which the body acts as a resource that allows cognition and the mind to extend. The body (including the brain) allows for fluid and efficient interaction

[19] Clark and Chalmers (1998), 7.
[20] Ibid., 18.

with the environment. It permits appropriate cognitive inputs and can itself act "as part of the problem-solving economy", as it allows us to outsource our cognitive and computational routines to external resources, thus extending an integrated cognitive system out into the world.[21] According to Clark, (extended) minds like ours emerge from diverse, dynamic, surprisingly seamless wholes of heterogeneous elements and processes. While overextended compared to the traditional internalistic view of the mind, these elements and processes are extended only relative to the internalistic viewpoint.[22]

Clark emphasizes that the body and world act as 'participant machinery'. That is, the body and the world come together to form part of the machinery by means of which mind and cognition are physically realized.[23] The body functions to co-opt bioexternal resources into a deeply integrated but outwardly extended cognitive system. Mental states are spread out over this integrated body-world participatory machinery without any principled biological boundary. The cognitive system, including different modes of encoding and types of cognitive operation, is impartial to the boundaries of the brain or body of an agent. Instead of acting as a boundary, the body brings the entire system together by functioning as a gateway that allows for the cognitive use of external resources. It does so without limiting the types of things that can be fused into one cognitive economy.

Importantly, these external components are poised to govern our behavior in the same way that our internal mental states are. They play an active causal role in our cognizing. This is the same role, in fact, that our internal cognitive processes play. They are therefore poised to govern our behavior in the same way that our internal states are. The environment is more than

[21] Clark (2008b), 196-197.
[22] Ibid., 219.
[23] Ibid., 207.

an instrumental resource for cognition. It is (in part) *how* cognition happens and *where* it happens. The environment, consequently, plays a significant role in the production and guidance of our actions.

§1.5: The Standard View of Action: Reasons-Causalism

Any explanation of action must properly tie together an agent, her intentions, and her behavior. What I will call the Standard View is a broadly Davidsonian picture of action that has been relatively resilient in accomplishing this. Davidson defends what he calls the 'ancient' and 'commonsense' view according to which actions are rationalized by a kind of causal explanation involving reasons.[24] According to Davidson, it is reasons that rationalize, or explain, our actions. When we act, we are doing something for a reason, and the fact that it is done for a reason is one of the most important features of the act itself. When someone acts for a reason, she must have two features as an agent. First, she must have a pro-attitude toward the act type of which she is partaking. This pro-attitude simply indicates that there is something about the action (or the outcome related to it) that appeals to her. Second, the agent must believe (or know, perceive, remember, etc.) that her action is of that type. Thus, when we give reasons for why someone behaved in a given manner, we often appeal to this pro-attitude, or the belief related to it, or both.

Davidson contends that this belief-desire pair is the primary reason for an action, and that the primary reason for an action is its *cause*. So, if I move my hand thus flipping a light switch in my house, thereby turning on the light in my study, my primary reason for moving my hand in such a way is that this movement turns on the light in my study, and I have a pro-attitude towards turning on the light and a belief that my moving my hand in such a manner will do so. Different

[24] I will assess other options for addressing the constitution and explanation of actions in later chapters.

descriptions may give the primary reason for my action, depending on what my pro-attitudes and beliefs are at the time.

> *R* is a primary reason why an agent performed the action *A* under the description *d* only if *R* consists of a pro attitude of the agent towards actions with a certain property, and a belief of the agent that *A*, under the description *d*, has that property.[25]

Reasons for action must therefore comprise these two distinct types of mental state – a pro-attitude and a belief, both relating to a description of the behavior enacted in the world.

Reasons play a multifaceted role when it comes to action explanation. The pro-attitudes in my reasons for action imply that such actions are intentional. Specifically, according to Davidson, to know the primary reason for which someone acted is to know the intention with which the action was done. Intentions are thus part of primary reasons for actions, though not the other way around. Actions seem coherent in light of their primary reasons, as a reason also explains and thereby justifies an agent's acting in a given manner. An agent performs an action *because* he has a reason, where 'because' is a causal concept. The behavior is not only explained by and justified by its primary reason; it is concretely caused by it. Our intentional movements, whatever description they fall under, are caused by the primary reason for their fulfillment.

Davidson also introduces the important distinction between an action and a mere causal happening. Intuitively, the idea is that some events in my life are *my* doings, they are my actions, while other events in my life are things that happen *to* me. They are mere 'happenings in my history'.[26] Some things I do to the world, and some things the world does to me, or, at least, independently of me. When our reasons cause our primitive bodily movements (or some external change) in the right way, we not only have behavior, but also action. It is easy to make an

[25] Davidson, Donald (2001). *Essays on Actions and Events*. 2nd ed. New York: Oxford University Press. 5.
[26] Ibid., 43.

intuitive case for this distinction, but explaining exactly what constitutes it is less simple. At first glance, it seems as if intentionality is the mark of agency. Actions are those behaviors that are done with an intention, which is at least partly constituted by an agent's reasons for acting. Intentions can be satisfied or fail to be satisfied based on whether the action that the intention is about conforms to what is represented in the content of the intention. Davidson, however, argues that this alone is insufficient to give us a precise enough picture of what action or agency is. Davidson states that someone is the agent of an action when what she does can be described under an aspect that makes it intentional.[27] If there is a description of the action, connected to the belief component of an agent's reasons for acting, such that the agent of it acted intentionally, this will suffice to elucidate his agency at the time. A person does, as an agent, whatever she does intentionally under a given description.

To provide a causal explanation of an event is to give its cause under an appropriate description. An *appropriate* description is one that specifies a causally relevant property for the event in question. To provide a causal explanation of *action* is to give an action's cause under an appropriate description, or one that specifies a causally relevant property for the action.

In doing an action intentionally, an agent acts "under a description". The descriptions under which an action is intentional "are determined by its primary reason (or primary reasons if there are more than one)." They are determined by the features under which the agent subsumes the action in the belief component of her primary reason for action.[28] If an agent flips a switch in order to illuminate a room in her house, and she believes that that action *as a particular flipping of the switch* is expected to bring about the illumination of the room, then the action is intentional

[27] Davidson (2001), 46.
[28] Ferrero, Luca (2013). Intention. In E. Lepore and K. Ludwig (eds.), *A Companion to Donald Davidson* (pp. 75-90). Malden, MA: Wiley. 76.

under the description "flipping the switch". She flips the switch in light of her belief that such flipping of the switch is going to satisfy her pro-attitude toward illuminating the room.

Let's say, however, that a prowler becomes alert to the agent's presence as an effect of her illuminating the room by flipping the switch. "By contrast, the prowler becoming alert to the agent's presence in the room is a causal effect of the flipping no conception of which plays a role in the primary reason and which is unknown to the agent."[29] As such, alerting the prowler is not an aspect under which her action is intentional. It is something the agent does, but not something she does intentionally.

The concept of intentionality in action is often difficult to reduce to more primitive features of behavior.[30] According to Davidson, intentional actions are behaviors that are done for a reason. On the causal account, reasons cause actions, while on some non-causal accounts reasons explain actions without causing them. So, if we take the Standard View, intentional actions just are those actions that are caused by a primary belief-desire (or pro-attitude) pair. Other times, the term 'intentional' is used to denote an action that is done with intention, or whose results are intended. If the term is used in the first way, we must then make sense of what reasons for action are. If used in the second way, we need to determine what counts as an intended action and an intended result, and this may or may not be traceable back to one's reasons for acting, their desires, or their motivational states. Thus, even if we take a causal view, there may be more work needed in order to determine what constitutes our motivating reasons and what separates intentional actions from unintentional behaviors.[31]

---

[29] Ferrero (2013), 76.

[30] For more on the relationship between the broader notion of intentionality and intention in action, see Searle, John R. (1979). The Intentionality of Intention and Action. *Inquirey* 22: 253-280.

[31] I will say more on agential control, responsibility, and intentionality in Chapter 7 of this

On the Standard View, causality can explain intention when it comes to primitive actions (often simply bodily movements) as well as non-primitive ones. The agent causes her body to move, and in doing so intentionally causes something under a given description, whether this is the bodily movement itself or what that movement causes. When I move my hand in order to turn on a light it is the movement that I am causing, and the turning on of the light that I intend, and which is caused by this movement. Simply, an agent causes what her (primitive) actions cause. It is this relation of causality that holds between our primitive bodily movements and what these movements do in the world. The reasons involved in this causal chain ensure that the agent is aiming for some state of the world and therefore acting intentionally.

If all actions are intentional under some description, and acting with intention is defined as performing an action that is caused in the right way by pro-attitudes and beliefs, then all actions will be causally defined. While intending is not itself a belief *that* one will act, it does seem to involve wanting to act in a given way. Intention involves the desirability of the intended action on some level. In a broad way, we intend to do what we want to do. Actions involve a kind of judgment about the way the world is and the way we can best navigate it.

Mental states and events constitute most (if not all) of the agent's role in the event causal process of an action. The idea of control is often used to elucidate the difference between intentional action and other events. According to the Standard View, agent-involving causal processes constitute the exercise of control such that the action produced is caused by some reason-states of an agent, which rationalize its performance.[32] There is an often-raised problem for the Standard View that a behavior might be caused by reasons and produce intended results

---

dissertation.

[32] Schlosser, Markus (2008). Agent-Causation and Agential Control. *Philosophical Explorations* 11, 1: 3-21. 10.

though still not count as an action. The basic problem of causal deviance, as it applies to basic actions, has this structure:

> There is an agent's mental state or event, R, that rationalizes the performance of an action of type A. R causes an intermediary state, N, which causes a movement of type M. This movement would constitute or realize an instance of A-ing, were it not caused by N. Typically, N is a state that appears to undermine the agent's control over the movement, such as a state of severe nervousness, and so it seems clear that the agent does not perform an action at all. But according to a causal theory, the agent performs an A-ing, as M is both caused and rationalized by R.[33]

The basic idea is that there may be extant behaviors that are both caused by and rationalized by a primary reason, but during whose occurrence there is some intervening or causally intermediary state that undermines the agent's control or reasons-responsiveness. Davidson uses the following example:

> A climber might want to rid himself of the weight and danger of holding another man on a rope, and he might know that by loosening his hold on the rope he could rid himself of the weight and danger. This belief and want might so unnerve him as to cause him to loosen his hold, and yet it might be the case that he never chose to loosen his hold, nor did he do it intentionally.[34]

I believe that there are some adequate responses to the problem of causal deviance,[35] which I do not have room to review here. For present purposes, I will stipulate that reason-states must cause an action in a non-deviant manner. The causal chain running from these reasons to the action must not run through a separate event that undermines the agent's control. Thus, the behavior must be guided by and responsive to these reason-states.

Because an action is caused by reason-states, and because these reason-states are the agent's *own* reasons, the behavior of an agent is self-determined, or owned by an agent, even when considering the external influences and causes that affect the agent. Agents whose reason-

[33] Schlosser, Markus (2010). Bending it Like Beckham: Movement, Control, and Deviant Causal Chains. *Analysis* 70, 2: 299-303.
[34] Davidson (2001), 79.
[35] See Stout (2010), Schlosser (2007), and Enc (2004).

states cause their actions in a non-deviant manner, whether or not there are external influences on this causal chain, are in charge of what they are doing. This is an account of control in terms of non-deviant causation by agent-involving reason-states. So, an agent is in control of her behavior just in case her behavior is caused by, guided by, and responsive to mental states and events that rationalize its performance.[36]

Guidance by reason-states is necessary so that actions can be brought about in accordance with the action plan of the agent who performs them.[37] This is a plan that specifies how an outcome is to be brought about, through the agent's body and the circumstances of the world. It involves the same reason-states that cause the action itself. The causal relevance of the content of these reason-states ensures that the agent is reason-responsive. The action is not merely a response to a cause. It is a response to the content of the reason-states that cause it. This accounts for the way in which mental states and internal dispositions not only generate but also guide behavior.

Davidson describes these reasons as being explanatory, but they are also often additionally (or alternatively) seen as *motivating* reasons. When I say that an agent acted for a reason, I might be attempting to explain his action by giving its source (why the action itself came about causally), or I might be explaining why she did what she did by appealing to what motivated her to act. Nevertheless, motive can be reduced, in some sense, to the causal character of reasons. Even if we appeal to reasons in order to explain the motivation or teleological nature of the action, the idea of a reason being motivating for an agent, or having 'purchase' on an

[36] Schlosser (2008).
[37] Ibid., 10-11.

agent, is a causal idea, making any sort of explanation of action by citing the agent's reason a causal explanation.[38]

Many proponents of the Standard View construe desires as directed towards the agent's ends, and beliefs as directed towards means to those ends. Agential behavior typically warrants a description in terms of the agent's knowledge of external circumstances and of his/her wanting to attain, or aiming at, a certain goal. This conceptual link between agency and goals involves a cognitive link between the perceptual and volitional powers of the mind. Our actions are responses to our perceived circumstances given our desires, wantings, and other pro-attitudes. When we perceive an opportunity to achieve a desired goal, we act voluntarily in response to this opportunity, given our predisposition to act in pursuit of it.

These Davidsonian ideas come in part from a general Humean view of causation and action. On the Humean view, some desire is always part of the reason for which one performs an action. This view maintains that desires are an essential part of the reasons for which we act and an essential part of action explanation. Without desires, Humeans argue, there would be no motivation to act and therefore no intentional actions.[39] Without some such desired goal state, there seems to be no reason why we should go through the trouble of acting at all.

The Standard View is just one of many causal views about action. The causalist is simply someone who claims that any explanation of the relation between an action done for a reason and the reason for which it is done must ineliminably invoke the notion of causation. Causalism defines action by reference to its causal history. In other words, an action is such only if and only if it is caused and causally explained by some mental items, such as beliefs and desires, that

[38] Ruben, David-Hillel (2009). Con-Reasons as Causes. In C. Sandis (ed.), *New Essays on the Explanation of Action* (pp. 62-74). New York: Pelgrave Macmillan. 62.
[39] Alverez, Maria (2009). Reasons, Desires, and Intentional Action. In C. Sandis (ed.), *New Essays on the Explanation of Action* (pp. 203-219). New York: Pelgrave Macmillan. 203.

either constitute reasons for action or act as intermediaries between the agent's greater mental and cognitive capacities and her actions.

This is consistent with two claims. The first, and the stronger of the two, is that the reason for which someone performs an action is the cause of that action. This is the Davidsonian view. It typically involves the claim that reasons are mental things (either states or events, or a combination of the two), and that these mental things are the cause of our actions and constituents of the reasons that explain them. Strong causalism claims that a behavioral event A of an agent S is an action "if and only if S's A-ing is caused in the right way *and* causally explained by some appropriate nonactional mental item(s) that mediate or constitute S's reasons for A-ing."[40] The causal antecedents of our actions constitute reasons for action.

The second, weaker, claim is that explanations for actions are causal explanations, though reasons may or may not be part of the action-causing chain.[41] Hornsby, for instance, holds that actions are not wholly accessible from the impersonal point of view and therefore cannot be seen simply as the workings of a causal chain of events in the world. Reasons explanation can be invoked when it comes to action, according to Hornsby, but the primary reason for action is not its cause.

> If we are to make sense of it as 'the cause', then the primary reason must be an item that there is if and only if the relevant agent believes some particular thing and desires some other (related) thing. But why should we think that there is any such item as this?...And if such a fusion (or pair, or whatever) were indeed the cause of an agent's doing what she did, how would it relate to all the other causal truths about the situation?[42]

---

[40] Aguilar, Jesus and Buckareff, Andrei (2010). The Causal Theory of Action: Origins and Issues. In J. Aguilar and A. Buckareff (eds.), *Causing Human Actions: New Perspectives on the Causal Theory of Action* (pp. 1-26). Cambridge, MA: MIT Press. 1.

[41] For more on the causalist viewpoint, see Alverez, Maria (2007). The Causalist/Anti-Causalist Debate in the Theory of Action: What It Is and Why It Matters. In A. Leist (ed.), *Action in Context* (pp. 103-123). New York: de Gruyter.

[42] Hornsby, Jennifer (1997). *Simple Mindedness: In Defense of Naïve Naturalism in the Philosophy of Mind*. Cambridge, MA: Harvard University Press. 134.

The point Hornsby is making is that we do not have to answer these questions if we deny that the causal explanation view relies on the idea of discrete things combining or interacting in the production of action.[43]

The type of strong causalism that Davidson describes attempts to bring us from a personal, to a causal, to an impersonal point of view. This is an attempt that Hornsby describes as fruitless. We ought not to subsume actions completely in the impersonal world of causes.

> When we see an action as a person's initiating a series of events, we recognize a type of event whose causal ancestry is understood from a personal, rational, point of view, and whose causal successors come to be understood from an impersonal, perhaps scientific one. And we appreciate that causality is a concept that we may operate with from both points of view: people make a difference, and do so because their actions are events which make a difference.[44]

Actions are not to be "swallowed up" by causal explanations, and action explanation is not to be understood from this *purely* impersonal, causal, and scientific point of view. From the personal point of view, and action is a person's doing something for a reason, while from an impersonal point of view and action is a link in a causal chain.[45] Reasons become relevant when we view actions from the personal point of view and attempt to explain why a person did what they did, not when we view actions scientifically, as events within a causal chain. Though actions can be explained causally, this is not where reasons come into play.

According to the Standard View, which supports the stronger claim, mental states are constitutive of the reasons states that cause our actions. The Standard View, along with other causalist views, assumes that the concept of agency is a causal concept. An agent causally generates and guides her behavior. Agents may also be seen to cause the results of their actions,

---

[43] Hornsby (1997), 135.
[44] Ibid., 151.
[45] Ibid., 129.

which are states of affairs rather than events, but at the very least they cause the sequence of events leading to these results. According to the Standard View, these agential behaviors are actions just in case they are caused by the right agent-involving mental states and events in the right way. Such a view reduces the entirety agent's role in the performance of actions to a causal role. Causation by an agent is causation by agent-involving states and events – usually construed as mental states and events. There is nothing more that a person has to do to be an agent than to produce the agent-involving beliefs, desires, judgments, etc. that cause actions.

The Standard View is thus a reductive view of action.[46] According to such reductive theories, an agent's ability to act in light of reasons can be both reduced to and explained by reference to agent-involving states and events.[47] These states and events, namely pro-attitudes, beliefs, and the relations between them constitute part of what an action is. Other forms of behavior, those that cannot be explained causally in light of reasons, simply don't count. According to non-reductive theories – oftentimes agent-causalist views – the power to act is irreducible and primitive. It cannot merely be reduced to these agential reasons states. As I will examine in the next chapter, however, these non-reductive theories seem incompatible with the spirit of the extended mind theory.

This causal picture often seems to presuppose mental state internalism, or that mental states (including reasons states) are internal to an agent. However complicated the causal chain associated with our psychological states, the mental states and cognitive processing in the middle of the chain partly constitute the mind, and causally mediate between non-mental inputs (our

[46] This is, however, not a reductive view of the mental states to which action can be (partly) reduced.
[47] Schlosser (2008).

perceptual capacities) and non-mental outputs (our actions).[48] Mental states causally intercede between our perceptions of the world and our active affects in the world. The Standard View tends to assume that beliefs, desires, etc. are physical events in the brain of the actor. This, however, though frequently presupposed, is not necessary for the Standard View to be successful. The view is compatible with extended theories of mind, at least insofar as primary reasons can be external to an agent.

The Standard View is compatible with a physicalist reduction of mental states, it is agreeable to most standard theories of causation, and it is compatible with differing theories of freedom and determinism. Causal views are also widely considered to be intuitive in that reasons for action are the cause of our active behavior, and our behavior is motivated and explained by our own reasons, constituted by our beliefs and pro-attitudes. I will evaluate competing views of action in the next chapter, evaluating whether they are compatible with the extended mind theory and whether they can help to ameliorate the agency problem.

§1.6: The Agency Problem for the Extended Mind

It is initially unclear how intentional actions occur if our minds are extended. Though authors often seem to presuppose the Standard View, many versions of EMT do not specify how we ought to explain human agency or how actions actually proceed. Further, a greater problem arises, for it appears that when we start to explain agency in the extended context, agential control ultimately disappears. The more our minds substantively extend beyond our bodies, the more genuine agency becomes metaphysically unfeasible and conceptually unrecognizable. If we accept EMT, actions, as we generally think of them, do not emanate from agents in any

[48] Stout, Rowland (2007). Two Ways to Understand Causality in Agency. In A. Leist (ed.), *Action in Context* (pp. 137-153). New York: de Gruyter. 137.

discernable fashion. More importantly, genuine agency, in the metaphysical sense, is non-existent in the realm of extended cognition and extended minds. Not only is our standard concept of action eliminated by EMT, but so too is the agential subject herself.

EMT clearly denies the conventional idea that intentional action is the output of neural processing in the brain. Think first of Otto and his notebook from Clark and Chalmers' original paper "The Extended Mind."[49] Otto desires to go to the Museum of Modern Art on 53rd Street in New York. Otto also has a (extended) belief, written in his ever-present notebook, which designates the museum's address. Accordingly, Otto goes to the museum, after consulting his notebook, by going to the address listed. It seems, however, that *Otto* is not really genuinely and intentionally performing this action. His desire is internal to his brain, but his belief that allows him to fulfill it is externally located. If the notebook really does contain Otto's beliefs, then *it* is guiding his actions in a significant way, completely apart from what is internal to Otto himself. Therefore, it seems that it is not Otto himself who is performing an action, but the non-person fusion consisting of Otto-and-the-notebook that in fact performs it.

This is just one example of how EMT denies our traditional understanding of actions. Once the boundaries of skin and skull are no longer privileged with respect to mental states, individual agents seem to lose their power to act. According to extended theorists, the causal relations between the body, including the brain and nervous system, and environment are intricate and continuous. There is nothing special about the causal relations inside the skin, or inside the head, and nothing internal that is especially capable of constituting the causal antecedents to actions.[50] Extendedness about mental states requires that there is nothing special

[49] Clark and Chalmers (1998).
[50] Hurley, Susan (1998). *Consciousness in Action*. Cambridge, MA: Harvard University Press. 336.

about internal states when it comes to action.[51] If EMT is correct, then the causal source of a person's behavior is external to that person, whose sourcehood is fundamentally undermined. And, because the person is no longer the unique source of the reasons states leading to her behavior, her agency is undermined as well. The causal antecedents to our actions are a function of the entire dynamic system which involves a subject, but which also extends out into the world. Actions are not the effects of *internal* reason-states that an agent has on the world; they are also constitutively associated with the external world through the extension of these states.

If this is true, then the autonomous will and causal sourcehood of an agent is significantly impaired. The extended mind theory entails that intentions and behavior are neither autonomously generated nor uniquely sourced in relation to the world. On the contrary, the production of intentions and behaviors is outsourced to the external world. External vehicles determine the causal profiles of our beliefs, pro-attitudes, and bodily movements; these do not merely flow from mental and physical states located in a rational human subject. The subject's external environment is constitutively determinative of her intentions, her reasons states, and the production of her behaviors.[52]

Our intuitive sense of agency involves the ability on the part of ourselves, as persons, to intervene or make a difference in the natural world. Already, the scientific and physicalist pictures of the natural world themselves appear to mildly threaten this conviction. Our actions seem to be mere events that are either causally determined by other events or are chance occurrences that cannot be considered intentional actions. There is less room for the notion that

[51] For more on this, see Hurley (1998), Chapter 3.
[52] Hurley (1998), 272.

we, as agents, are the authors of our actions in a non-reductive sense. Our actions become instead part of the natural order, which we originally conceived of ourselves as effecting from apart.[53]

Even according to the physicalist picture of the world, though, human agents are still the bearers of certain properties that cause events. Since persons possess distinctive causal powers, we can still make space for human agency. We remain persisting wholes with unique properties that cause our behaviors. Agents can still act or abstain as they perceive available opportunities. We act in pursuit of the goals that we construct, cognitively, for ourselves, displaying a relatively wide range of voluntary behavior in response to the circumstances of the environment around us.[54] Even with a reductive picture of agency, agents remain the bearers of the reason-states that cause their actions.

Fundamentally, we view our environment as constituted by a system of objects that exists independently of us. They act upon each other in accordance with causal laws and natural pulls and pushes. These are objects that are independent of our will by being independent of our reason-states. When we act, we act upon the objects in our immediate environment by bringing about changes in this external, independent world. Extended theories of mind, however, take this remaining sovereignty away from the agent. When our reasons states are extended out into the world, our actions become caused, guided, and determined by a set of objects in the world, rather than by our own internal mechanisms.

The problem for EMT is that the theory seems to create unique problems when combined with the Standard View of action. Reasons states, according to EMT, are external to the body of

---

[53] Lowe, E.J. (2009). Free Agency, Causation and Action Explanation. In C. Sandis (ed.), *New Essays on the Explanation of Action* (pp. 338-355). New York: Pelgrave Macmillan. 338-339.
[54] See Hacker, P.M.S. (2007). Thought and Action. In A. Leist, ed. *Action in Context* (pp. 53-74). New York: de Gruyter.

the agent who is supposed to be responsible, in the causal sense, for action performance. EMT and the Standard view are not incompatible at first blush. EMT still allows that there *are* reasons states constituted by beliefs and desires, even if these beliefs and desires are extended. It is part of the extended mind theory that these states are poised to govern behavior. However, once we recognize that agency requires authorship by an agent, as opposed to by the environment, the problem arises that when the central elements of the Standard View are combined with EMT, this essential element is lacking. Once we add the requirement of agential control to the Standard View of action in order to define what it means to be an intentional agent, EMT becomes incompatible with that definition. Because reasons states are external to the agent, according to EMT, and because reasons states cause our actions, our actions will be caused by things that are locationally external to us. While our reasons states are extended, they are still able to govern our behavior in the same way that internal mental states do. It is part and parcel of the theory that these states can and do govern our behavior (functionally) in the same way that inner states do.

§1.7 Concluding Remarks

In what follows, I will evaluate which theories of mind are the most subject to this worry, and the extent to which it is problematic for differing extended views. Interestingly, many extended views seem to favor the Standard View. Clark (2010), for instance, presupposes such a view when describing actions that originate from extended beliefs. There are a few options for the extended mind theorist in terms of accommodating this problem or adopting an alternative view of action that may avoid it. I find these options to be insufficient solutions to the problem. Revised views of extendedness, the self, and action do not save EMT. In the end, the only hope

philosophers of mind have for accounting for agency and control is to reject extended theories of mind.

Chapter 2: Views of Action and the Agency Problem

§2.1 Beyond the Standard View

The Standard View creates problems for EMT in part because it explains our actions with respect the reasons states that cause them. When our reasons states are extended out into the world, or when things in the world constitute them, our actions become caused, guided, and determined by external things, as opposed to agential things. The problem for EMT is that the theory seems to lead to unique problems due to the externalization of our reasons states. Because reasons states are external to the agent, according to EMT, and because reasons states cause our actions, our actions are caused by things that are locationally external to us.

Nevertheless, it might seem possible to avoid these worries if the extended mind theorist rejects the Standard View in favor of some other view of action. The Standard View, while popular, has received certain criticisms from philosophers of action. Causal deviance is one problem for the Standard View that has received the most attention, with some theorists believing the problem to be essentially solved (or, simply, not as problematic as it seems), and others believing it to be the beginning of the end for any Davidsonian-Causalist account of action. Others believe that there are additional problems (apart from causal deviance) for the Standard View, including a lack of adequate acknowledgement of the agent herself. I believe that the central tenets of the standard causal view of action are defensible, but I will not focus on their defensibility here. Instead, I will give a brief overview of differing accounts of action – both reductive and non-reductive in nature – and examine their suitability for combination with extended mind theories. This chapter will examine how other less classical views of action might fare in combination with the extended mind in resolving the agency problem.

Many extended mind theorists often seem to presuppose something very much like the standard view when explicating their theories. Should extended mind theorists seek a revisionary approach, however, it must be the case that there is some other view of action, compatible with extended mental states, that can take the place of the standard causal account. Though there are some contenders for such a view, I see none as being successful at solving the agency problem.

§2.2 The Causal Process Model

Rowland Stout argues that agency is best understood in terms of "making things happen,"[55] in a procedural sense. Agency, according to Stout, ought to be understood in terms of intentionally doing things, and intentionally doing things ought to be understood in terms of making things happen. The difference between this model and the Standard View lies mainly in how 'making things happen' is understood. Stout advocates the following conception of agency:

> An agent intentionally causes things to be G if and only if things being G belongs to a process that is the working mechanism embodied by the agent which results in what should happen according to some version of practical rationality for things be to G.[56]

When an agent intentionally makes something to happen, some process embodied by him causes things in the world to be characteristic of this activity or action. The agent embodies the mechanism that brings about a change in the world, making the world the way in which she intends it to be. This agential process causes the world to conform to a description typical of the chosen action. The agent must embody the mechanism that causes the world to be a certain way, and the way that the world ends up must conform to the way in which it ought to be (normatively) in order for such a goal to be met. An agent intentionally causes things to be a certain way (or, in short, she intentionally causes something) if and only if things being that way

[55] Stout (2007), 142.
[56] Ibid., 148.

(the end result of this causation) belongs to a process that is the working of an intentionally-causing-things-to-be-this-way mechanism embodied by the agent.[57] This belonging represents a responsibility on the part of the agent for making the world conform to her goals. The agent *makes this happen* by embodying the process that causes it to happen.

Stout argues for this approach to agency "based on the Aristotelian model of causal process rather than on the more familiar model of a network of causal links."[58] The Standard View of action is causalist, but it focuses on these causal links. On the traditional picture, however complicated the causal chain that leads to our actions is, "the structure of states and events in the middle of the chain constitute the mind, and causally mediate between non-mental inputs and non-mental outputs."[59] On one prevalent conception of causation the causal world can be described by identifying basic causal processes constituted from two events or complexes of events and states (one the cause, and one the effect) and a generic relation of causation linking them.[60] According to this picture, the causal world looks like a huge chain or network of events and states all linked together. This is what Stout calls the causal chain model of causation.

This, however, is not the only model of causation. Stout points as well to an Aristotelian model that allows individual causal processes in addition to mere causes and effects that generically relate to one another.

> In this model we can ask what the cause is and what the effect is, but in addition we can ask how the effect results from the cause – i.e. what the mechanism is. And this how-question is not to be answered just be introducing more links in a causal chain. It is answered by specifying the potentiality whose realisation is the process that the effect belongs to.[61]

---

[57] Stout (2001), 148.
[58] Ibid., 137.
[59] Ibid.
[60] Ibid., 139.
[61] Ibid.

Our mental states, on this model, are taken to be states of causal mechanisms transforming inputs (from perception, etc.) into outputs (i.e. behavior). States and events in the world are inputs into these causal mechanisms. In turn, these mechanisms constitute our nature as agents and subjects. And, "essential to these mechanisms are the potentialities whose realisations constitute the workings of these mechanisms."[62] These mechanisms transform perceived inputs into produced outputs. However, a change in one's mental states is not a fundamental part of a causal chain that runs between the environment and one's behavior.

Where F stands for a causal notion, the causal process model analyzes causal notions in this way: "*S F*-s *O* if and only if *O* having the *F*-quality belongs to a process that is the working of an *F*-mechanism (the realisation of its *F*-potentiality) embodied by *S*."[63]

Mechanisms, on this view, are identified by their potentiality. Potentiality, here, is an Aristotelian notion.

> A process is the realisation of a capacity or disposition for certain results in certain circumstances. You have to characterize a structure of stages to specify the potentiality. But what is required for the process to be happening is not just that that structure of stages is in train, but that there is a potentiality for such a structure and that this potentiality is being realized.[64]

This model is intended to be opposed to the causal chain model of agency, which depicts an agent's intentional action as being reducible to, or constituted by, causal relations linking the agent's mental states and events to her behavior, or the kind of intentional action performed.[65] The Causal Process Model (CPM) views agency as a matter of process or mechanism instead of the mere causal connection between mental states or events and bodily movements. Rather than

---

[62] Stout (2007), 140.
[63] Ibid., 145.
[64] Ibid.
[65] Bishop, John (2007). On the Importance of Reconciling Two Models of Agency. In A. Leist (ed.), *Action in Context* (pp. 154-162). New York: de Gruyter. 156.

understanding actions merely through the links in the causal chain that produces them, we ought to focus on the processes that constitute our agentive mechanisms.

According to the Causal Process Model, beliefs and desires generate judgments, which produce intentions to perform an action. These intentions then cause the agent to behave in the required way in order to make the world fit the description of such an action's fulfillment. However, according to the CPM, agency cannot be reduced to a causal chain that merely involves relevant beliefs and desires. Rather, agency requires a certain causal mechanism, or causal process, that appropriately functions to link our judgments, intentions, and actions. This causal process cannot be reduced to a series of events or states and their causal connections. Processes can be thought of as the workings of mechanisms, and process causation is different from event causation. However, Stout states, "to explain what it is for one event to result from another is at least as difficult as explaining what it is for an event to belong to a process."[66] One is no more mysterious than the other.

Going back to Stout's definition of agency:

> An agent intentionally causes things to be *G* if and only if things being *G* belongs to a process that is the working of an intentionally-causing-things-to-be-*G* mechanism that is embodied by the agent.[67]

The idea of "an intentionally-causing-things-to-be-*G* mechanism" needs to be spelled out. Stout states that the laws that describe the potentiality of this mechanism "must be characteristic of intentionally causing things to be *G*."[68] The question of what such laws are, and how we should specify the potentiality that is characteristic of intentionally causing things to be a certain way still stands. Stout merely says that this involves sensitivity to reason, in particular teleological

[66] Stout (2007), 147.
[67] Ibid., 148.
[68] Ibid.

reasons. An intentionally-causing-things-to-be-*G* mechanism is "simply a mechanism that results in what should happen for things to be *G*."[69]

The CPM is in part meant to re-introduce the *agent* into the production of actions. It is the agent who must embody the mechanism that causes things in the world to materialize and to fit a given actional goal state. Beliefs and desires, or representations and motivational states, are not enough to explain agency. These may still be causally involved in the production of our actions. However, the processes that involve these need to be embodied by the agent, which gives her agentive responsibility in a more concrete sense.

The Causal Process Model retains several important aspects of the Standard View, including a causal definition of intentional action. Agency, according to the CPM, is still reducible to agent involving states, events, and (most importantly) mechanisms and processes. And, intentional action is still understood as a causal concept. The main difference between the Causal Process Model and the Standard View is the way in which agency is defined – as an embodied mechanism with a potentiality rather than as a series or chain of discreet event-causal states.

According to the Standard View, the mind mediates between the world we perceive and the world we act on. States and events in the world cause mental states and events, which then go on to produce new states and events in the world through producing our actions. We can add feedback loops to this chain, but no matter how complicated we make it the states and events in the middle constitute the mind and causally mediate between non-mental inputs (perception) and non-mental outputs (action).[70] Some call this picture Cartesian, though it isn't committed to any sort of immaterialism. Many philosophers have rejected such a picture because it makes it

[69] Stout (2007), 148.
[70] Ibid., 137.

difficult to give teleological reasons for action a role in agency. Any reasons that cause your action on the Standard View may at best be reasons why you act, but they do not justify your action. Stout states that "the apparent inescapability of construing our causal role as the role of links in a causal chain derives from a prevalent but quite avoidable conception of causation itself."[71] By accepting the above view, which invokes the notion of process causation, we can reject the Cartesian picture without rejecting causalism outright. The process model of causation cannot be pictured as a network of generic arrows, as the causal chain model of causation can. Each arrow would have to represent a different kind of causal process. This is a model in which mechanisms have potentialities that are realized in the workings of these mechanisms.

> And applying this model to our minds' place in the world we can still take minds to stand between the perceived world and the affected world…But our mental states are not taken to be causal intermediaries in the sense of being nodes in the middle of the chain of causal links…Instead they are taken to be the states of causal mechanisms transforming the inputs into outputs.[72]

The way our minds are is the way that these mechanisms transform perceived inputs into produced outputs. Considering how a behavior is caused rather than what it is caused by may appease some anti-causalists without fully giving up a causal model of action.

However, the two views are in some ways very similar. The Causal Process Model also suffers from a similar problem as the standard causal view when it comes to extended cognition and the extended mind. Mental states remain the causes of intentional actions insofar as they play a role in the mechanisms that drive agential behavior. These behaviors are responsive to environmental features that are coupled with the agent. Therefore, extended mental states that are the causes of our behaviors render those behaviors devoid of intentionality. Consequently, a similar kind of causal outsourcing occurs as does with respect to the Standard View. In addition,

[71] Stout (2007), 139.
[72] Ibid., 140.

it is the agent who must embody the mechanism that causes things in the world to fit a given goal state, making the agent extended out into the world. Extended agency poses problems of its own, which I will address in the next chapter. If the agent is not extended, then she cannot embody the mechanism involving extended mental states, and her behaviors will likewise not be genuine actions.

Thus, the Causal Process Model of action seems prone to the same worry as applies to the Standard View when combined with EMT. Simply, causal outsourcing, whether it is interpreted with respect to a simple causal chain or an agential mechanism, undermines agency at its most fundamental level. This is true whether we define actions as caused by reasons-states or by agential mechanisms. The agency problem may be even more salient when considering the Causal Process Model of action due to the fact that causal processes require embodiment by an agent. This leads to a view of the agent that is extended, according to EMT. The extended self thesis turns out to be inherently problematic. According to the CPM, it is the process that matters, rather than causation by reasons states. If this mechanism is both extended and embodied, then the agent must be extended as well in order to embody it. If not, this mechanism will not be agential and will not generate genuine actions. I will examine this view of the extended agent at length later on, at present merely saying that I come to the conclusion that it is an unsatisfying view of agency and the self. However, regardless of whether we look to causal processes or causal chains to explain agency, causal outsourcing still occurs.

§2.3 Control Theories of Action

One important feature of actions is that they are not merely caused, but also voluntary and intentionally controlled. Control Theories of Action focus on this feature of human behavior.

I not only perceive and feel the motions of my body, but I also control them, making my basic actions self-determined in the important sense that they are guided by me. Self-determination requires this kind of control over our actions, such that they are counterfactually guided by our intentions. Barring external constraints, if I had intended differently, I would have acted differently. There are, of course, certain kinds of human behavior that fall short of, or do not qualify as, action. Actions, according to Control Theories, are essentially voluntary behaviors. And, in order to be voluntary, they must be under our control. Human actions, according to such views, are things that we can do or refrain from doing *at will*. There is something about us that determines our actions and puts them solidly under our agential purview such that they would not have happened had we willed differently.

Adams and Mele (1989) model their account of intentional action on control theories of goal-directed behavior.[73] On their view, a piece of behavior is goal-directed only if it is behavior of a goal-directed system – a system which employs measures that control its behavior. A piece of behavior *(B)* is directed toward a specific goal (*G)* due to an information-processing network of control within the system (*S)*. *S* must be causing *B* and comparing information about whether *B* obtains (or is about to obtain) with a representation of goal state *G*, and the system must be prepared to monitor progress or make error-corrections if it detects that *B* is not going to match *G*. The goal-directed behavior of a human system is the bedrock of action. A system, *S*, is a goal-directed system in its behavior *B* toward goal-state *G* if and only if:

1) *S* has an internal representational state *R* capable of fixing *G* and *S*'s goal-state and *S* is capable of detecting *G*'s presence or absence;
2) Information about *S*'s ongoing behavior *B* is fed back into the system as input and is compared with *R*;

[73] Adams, Frederick and Mele, Alfred (1989). The Role of Intention in Intentional Action. *Canadian Journal of Philosophy* 19, 4: 511-532.

3) *S*'s modification of output behavior *B* in response to comparison between *S*'s present state and *S*'s goal-state causally depend on the correction processes of (1) and (2).[74]

According to Adams, all goal-directed behavior fits this model, and intentional action (as a subset of this behavior) fits a very similar one. Intentions are mental states, which set an action plan, causally initiate an action, informationally update progress towards the end of that plan, provide a standard for determining behavioral error, correct errors (or, provide 'damage control'), provide criteria for the success of the behavior, and play a crucial role in the counterfactual dependency of the behavior upon the intention.[75]

Intentions are therefore the causal antecedents for actions. They determine the goals for our behavior, which is then controlled by the system that possesses them. These intentions set a standard for the behavior that the system then continually monitors in order to causally effectuate progress towards a goal state that it represents. The counterfactual dependency at issue is the dependency of actions on the intentions, which specify the goals for agentive behavior. Intentional mental states still do the work of causing actions, though in a different manner than described in the Standard View. Intentions in this sense are causally relevant throughout the action sequence as they set an action plan that is continually checked against for updating behavioral progress towards a goal. They not only initiate, but also guide our behaviors through updating and comparing sensory information against a representational goal.

Control theories of action do not privilege any kind of agential control mechanism. If behavior is controlled by extended states the same kind of agency problem arises. The counterfactual dependency of an action on extended intentional states poses the same problems

[74] Adams, Frederick (2010a). Action Theory Meets Embodied Cognition. In J. Aguilar and A. Buckareff (eds.), *Causing Human Actions: New Perspectives on the Causal Theory of Action* (pp. 229-252). Cambridge, MA: MIT Press. 230.
[75] Ibid., 231.

as are posed with respect to the Standard View. According to EMT, the intentional states that inform our action plans are extended beyond the boundaries of the person. When extended, intentions play an even bigger causal role on this theory through updating than on the Standard View.

§2.4 The Teleological Model of Action

Teleological theories of action start with the certainty that we are actors, apart from any certainty that we have beliefs and desires that cause our actions. The teleological model regards actions as a special kind of activity, but one that is not distinguished by the intentionality of the beliefs and desires that may or may not precede these acts. The teleological model involves more than the claim that actions have aims (or are supposed to fulfill our aims as agents) because they follow from intentions or reasons for acting. The teleological model – as an alternative to the belief-desire model – is made up of various elements which all result from the idea that action is the successful activity of an actor. According to this idea it is the self-perception of the actor that allows her to bring about results intentionally and to achieve these results *successfully*.[76] On this model, the causal relation between reasons and actions is not the basis of action explanation.

Leist (2007) lists three central elements of the teleological model. These are as follows: (1) teleological reasons for action, (2) agent causation, and (3) the priority of understanding over explanation. The teleological concept of practical reasons is that of a conscious aim for or towards something. It is not just a need but a need *for*, not just a motive, but a motive *for*, and not just a desire, but a desire *for*. Reasons are only said to be such against a teleological backdrop. Further, aim-based reasons go beyond the traditional role that desires play in action

[76] Leist, Anton (2007). Cognition and Action. In A. Leist, ed. *Action in Context* (pp. 315-343). New York: de Gruyter. 329.

theory to include the actual intended result of an action.[77] Agent causation is the idea that the actor *herself* causally brings forth her action. While this is done in connection with her reasons for acting, it is not merely the primary reason for her action that causes it. Rather, it is the actor herself. Finally, the priority of understanding is the priority of the actor understanding her own reasons for action and which of her reasons lead her to aim effectively at a teleological goal state. Understanding your own nature as an actor poised to achieve your goals, on the first-personal level, is an inherent component of being an actor. Further, there is no mental or submental event within an actor that could alternatively explain her actions.[78]

Causality cannot be enlisted in order to understand why someone acts. Rather, the self-perception of the actor, how she perceives herself as acting successfully, must explain why an agent behaves in the way that she does. No causal law can explain an action.

> A completely plausible understanding of action can surely only result by embedding the empirical-cognitive elements of acting within the completion of the action – and thus within the view of the actor and the aims behind the action striven for by the actor – and not *vice versa*, by embedding actions within an action-independent causal-event-construction of the world.[79]

It is the goal or purpose of an actor that explains an action, not an antecedent state that caused the behavior.

Insofar as teleological theories of action include agent causation, they are bound to suffer the same problems that agent-causal theories impose (see next section). There are, however, control models of action that may offer satisfactory accounts of goal-directed behavior in the absence of agent causation (see Adams and Mele's account above). These suffer problems of their own when combined with EMT. As I see it, however, the main problem with the

[77] Leist (2007), 330-331.
[78] Ibid., 332.
[79] Ibid.

teleological model is not only its emphasis on irreducible (and seemingly non-naturalistic) aspects of agency, but also its lack of plausible combination with extended mind theories. It may seem as if the teleological picture is consistent with the extended view of the mind, as both contain an appeal (implicit or explicit) to external states being relevant to behavior. But the way in which the teleological picture of action appeals to external contents as part of reasons states and desires (a desire being a desire *for* something, e.g.) is different from the way that the extended mind theory appeals to these contents as being partially constitutive of mental states whose vehicles are external to the agent. States of the world, in the teleological sense, might be constitutive of the content of our intentions or the meaning of them, but this view makes no provision for their constituting the vehicles of our thoughts. Generally speaking, extended mind theorists remain committed to a naturalistic picture of cognitive agents. Teleological models of action, insofar as they invoke the notion of agent-causation and the priority of understanding over event-causal explanation go against this commitment.

§2.5 Non-Causal Theories of Action

The Davidsonian-Causalist approach to action explanation has some features – reductionist, causal features – that have taken hold in the literature on the philosophy of action. There are a limited number of theories that reject the assumptions that actions ought to be explained causally, using a reductionistic approach. Many of the theories that reject these assumptions reject reductionism. Teleological theories of action (above) and agent-causalist theories (see below) are some of these. There are a few theories, however, that focus on rejecting causalism. These are non-causal accounts of action.

Some teleological theories may fall into this category. Valaris (2014) argues that actions are simply teleological processes, or events that are guided towards a goal. Actions are also holistic. They are not mere aggregates of events, but rather structured sequences of components, each of which are performed for the sake of the whole. Actions can be understood teleologically and instrumentally, where each part of an action contributes to a larger whole. Actions cannot be understood by reference to reasons states. Instead, they can be understood by their contribution to this whole, or this overall aim of the agent. This is a non-causal account of agency, as it does not refer to antecedent causes to explain why an action took place, but rather the overall forward-looking aim of the agent who acts.[80] The author's positive account is distinctly non-causal and non-reductive. Our actions are not the causal outcomes of distinct mental states that represent them. Rather, actions *are* the agent's grasping of a way to attain a goal state. Actions are, in fact, cognitive events in their own right. Movements themselves are cognitive in that they are a way in which an agent grasps a way to attain a certain goal state.[81] There is no hope, according to Valaris, of reducing action to any kind of purely causal chain where mental states stand as the causal antecedents to action.

This account focuses on the unity of actions. Instead of explaining bodily movements by reference to their antecedent causes, we can explain them by reference to the greater action that they lead to. So, we can explain the movement of a person's legs not by reference to reasons that cause this movement, but rather by reference to the action of climbing the stairs that these movements allow the agent to accomplish. Because they focus on forward-looking aims as opposed to backward-looking causal chains, teleological theories like this lend themselves to the

[80] Valaris, Markos (2015). The Instrumental Structure of Actions. *The Philosophical Quarterly* 65, 258: 64-83.
[81] Ibid., 78.

rejection of the causalist thesis. This account, though, will suffer from the same problems as a typical teleological theory of action, especially problems arising from teleological reasons for action. It also does not give us an answer to the question of *why* an action occurred when it occurred, i.e. why it came about at all. In order to answer this question we would need to reference what initiated the action in the first place.

Bede Rundle (1997) gives one non-causal theory of action in which he identifies action as purposive behavior, but states that action ought to be identified without reference to knowledge, belief, intention, or desire. Purposive behavior is goal directed, so that an action is the successful accomplishment of something. This way of elucidating action makes no reference to thought at all. Reasons, on this view, accompany actions rather than causing them. There is a reason *with* which one acts but not *because* of which one acts where because is a causal concept. And, while desires and beliefs are manifested or expressed in action, they do not cause actions. Purposive actions are voluntary and controlled, as well as goal-directed. Actions are voluntary in that we can choose to do them or refrain from doing them at will. Though this is not strictly speaking an agent-causalist account of action, it is similar in that it does not look to beliefs and desires to get actions under way. Actions are simply things done for the sake of an end as an expression of a person's will.[82]

This account of action, while it would avoid the problems for EMT that a causalist viewpoint entails, appeals to things that are metaphysically mysterious. While it does not straightforwardly appeal to agent causation, the notion of the will is not fully explicated in the theory. And, there seems to be a tension between actions being done and therefore being *doings* and actions not being caused. It is unclear how actions come about if they are uncaused, or what

[82] Rundle, Bede (1997). *Mind in Action*. New York: Oxford University Press. See chapter 6.

role mental states play if they do not play a causal role. When we ask why someone acted we are not only looking for a reason with which one acted, but also what caused them to act in that manner rather than another one. Further, mental states, according to EMT, play a central causal role in the production of our behavior, which is clearly at odds with this theory of action. Rundle also denies a full-fledged physicalist theory, which seems to go against the central commitments of EMT.

There are a very limited number of non-causal theories of action that are not teleological. The accounts given above in this section may not strictly fall into the teleological camp because they do not involve the notion of agent causation. However, they share many elements with teleological theories. These accounts of action adhere to the claim that reasons for actions are not event-causes of the actions that they rationalize. Many non-causal accounts of action, however, focus mainly on the broader social and moral context of action, and not action explanation at its most fundamental level. "Under this conception, the philosophy of action was called upon to give a systematic and general account of how ordinary agents 'make sense' of actions and those among them that are intentional, culpable, laudable, reasonable or not, and so on."[83] This focus on explaining actions in the context of social life, moral practices, and historical context is a radically different perspective on action explanation. These sorts of explanations do not seek to find an account of action that works with the natural sciences and the sciences of the mind. Rather, their authors seek to give an explanation of action that is completely apart from these naturalistic considerations.

Non-causal accounts of action in this context are not free from the problem of metaphysical dualism that plagues teleological and agent-causalist accounts of action. These

[83] Gustafson, Don (2007). Neurosciences of Action and Noncausal Theories. *Philosophical Psychology* 20, 3: 367-374. 372.

accounts do not take the viewpoint of our selves, as agents, as essentially physical beings. And, explaining actions in virtue of their social and moral context seems to bring about a different form of the agency problem, where agents are thoroughly embedded and are distinctly lacking in independence from their environments.

§2.6 Agent-Causal Theories of Action

Agent-causal theories of action are non-reductive theories that take agents to be ontologically primitive beings. These theories put forth a view of action as being irreducible to a causal process involving mental states and events, while reductionistic causal theories tell us that an event is an action performed by that agent in virtue of being caused by the agent herself. The exercise of the agent's power to act is not reducible to causation by agent-involving states and events, which an agent may also be essentially reducible to. Rather, causation by an agent is causation by the *substance* that is the agent. Thus, such views presuppose substance causation rather than event causation. Agent-causalists insist that genuine actions must involve an ontologically irreducible kind of causal relation where the *causer* is a substantial agent.

Agent-causalist views include the idea that, ontologically, there exist both substances and properties, the former of which constitutes an agent. In short, agents are objects, objects are substances, and substances are the basic entities that are the bearers of properties. Agents (*qua* objects) are not bundles of properties. They *possess* properties, making them the bearers of these properties.[84] Agents are therefore substantial things that possess properties, but they are not reducible to these properties.

[84] Buckareff, Andrei (2011). How Does Agent-Causal Power Work? *The Modern Schoolman: Special Issue on Free Will and Moral Responsibility* 88: 105-121.108.

When an actor is seen as the source of her own activity, as an unmoved mover, so to speak, her actions no longer fit into the natural order of the world. She is the source of her activity, not in the sense of being a link in an unbroken chain of causal events leading up to her action, but in a way that is not determined by the causal events that come before the act in question. An agent's power to act is irreducible and primitive. The agent or person is no less primitive than a physical body in the world. Agent-causal views argue that the agent herself (understood fundamentally as a substance) performs an action, which is caused by this substance. Since a substance isn't itself the sort of thing that can be a causal effect in some causal chain in the world, the agent is the fundamental originator of her actions. There is no event that causes our actions, because it is the substantial agent that causes them. However, this substance does not fit neatly into the naturalistic causal order of the world or the physicalist picture of it.

Agent-causal power is inherently goal-directed in that agents have the power to cause an intention in order to satisfy some aim or goal. This, however, is the causation of an object *qua* object. The existence of an intention in this sense cannot be reduced to agent-involving states and events as it could on a causal theory of action. It is the substantial agent herself that causes her intention to act.[85] This is a deeper sense of agential responsibility in that the agent is (causally) responsible for her action all the way down. She is the end of the line when it comes to explaining her actions. It is not reasons or mental states that causally produce an action, but rather the agent herself (considered as a substance). She not only causes her actions, she also causes her intentions to act in a given manner, making her the ultimate source of all things action-oriented.

> Agent-causal power is a *sui generis* type of causal propensity internal to an agent that is structured by reasons for action and other internal and external influences aimed at

[85] Buckareff (2011), 112.

> *enabling* the agent *qua* substance to causally produce a class of effects of a specified type (coming to have an intention) internal to the agent.[86]

Thus, reasons may structure the causal propensities of an agent, but they do not determine these. They *enable* an agent to act in a certain manner, in that they structure agent-causal powers in a way that makes the action more probable. Reason-states therefore causally contribute to the event that is or constitutes an action, but it is the agent that has the ultimate power to direct her activity.

Only some agent-causal theories acknowledge that reasons states (beliefs, desires, etc.) play a role in the causation of action (a structuring or enabling role, perhaps). But, even on these views, reasons are not the only influences in the causal process that leads to action production. Rather, the agent herself plays an important and decisive role. Other theories deny that such states play any role at all in action-production. Instead, the causal source of an action is entirely the agent who performs the act. On the traditional agent-causal view, "agents are the sole causes of actions, and agent-involving mental states and events do not play any causal role in the performance of action."[87] Whenever we act, we freely choose what to do on the basis of reasons, but these reason-states are not the causes of our action. The agent, while she may act in light of some reason-states, has the ability to exercise power that is distinct from the influence of these reason-states. There is more to the production of action than the simple influence that reason-states have. Agents have the power to directly bring about choices and actions without being held hostage by these states.

Some think that such views of action are preferable because they capture the phenomenon of agency in a way that event-causal views do not, and perhaps cannot. Agent-

[86] Buckareff (2011), 114.
[87] Schlosser (2008), 5.

causalists deny the natural conception of agency that causal theories of action put forth. They deny the idea that agency can be properly understood as constituted wholly by naturalistic events and states of affairs, as the prevailing natural scientific metaphysics interprets these.[88] While this may seem a noble interpretation of agency, as it gives the agent herself a privileged position with respect to her actions, this privilege comes at a significant cost – the cost of a naturalistic ontology itself. When the agent is understood as a unique substance with causal powers that are irreducible to the causal powers of things in the natural causal order of the world, we seem to sacrifice a naturalistic ontology. When agent-causation is understood as something *more* than event-causation (and thus irreducible to it), agency becomes naturalistically mysterious.

One other, related, problem with agent-causal theories of action is that reference to an agent cannot explain why certain types of effect occurred when they occurred. On the agent-causal view, the only thing we can appeal to in order to answer this question is the (metaphysically mysterious) agent herself. An appeal to event-causation can, however, explains this by referring to causation by reasons-states. When we ask, 'Why did xyz happen?' an event-causal explanation in the form of the Standard View can give us an answer to the question. All an agent-causal view can tell us is that the agent caused it to happen. Any further appeal to reason-states will be futile, as these are merely structuring features and are non-determinative of action performance.[89]

A similar problem occurs when we attempt to explain how an agent exercises control over her actions. The exercise of agent-causal power doesn't seem to be guided by anything sufficient to entail that the agent has control over her behavior. The intentional contents of the

[88] Bishop, John. (2010). Skepticism About Natural Agency and the Causal Theory of Action. In J. Aguilar and A. Buckareff (eds.), *Causing Human Actions: New Perspectives on the Causal Theory of Action* (pp. 69-83). Cambridge, MA: MIT Press. 71.
[89] Schlosser (2008), 10.

agent's reason-states are not determinative of her behavior, and her behavior is not a response to them (at least not entirely). Not all causal processes are under control, and the agent-causal theory does little by way of specifying what is special about agent-based behavior that makes it controlled.[90]

Many consider agent-causal models of action untenable for these reasons. The agent-causal model seems to fail with reference to a naturalistic ontology, its lack of a satisfactory explanation of why actions take place, and in its lack of explanation of agential control. Agent-causal theories would, however obviate the agency problem by introducing an irreducible agential component into the action sequence. If actions just are those things that are caused by agents, and agents are irreducible things with irreducible causal powers, then the problem of how we act when our mental states are extended is solved. The briefest answer to the question of how we act is that *we just do*. Problematically, though, agent-causalist accounts of action seem to run afoul of the basic precepts of EMT, as agent-causalism is at odds with the ontological commitments of the theory. Extended theorists take themselves to be espousing a naturalistic account of the mind and of agency. A non-reductive appeal to agency at the ontological level of a substance itself seems incompatible with this naturalistic approach. There is simply no place for an ontologically distinct, substantial agent in the extended mind theory.

§2.7 Dynamical Systems Theories and Agency

Many views of agency (minus, perhaps, agent-causalist views) retain some central aspects of the Standard View. This includes the standard picture of perception and action as inward coming and outward going, respectively. When the agent is seen as a dynamic system,

[90] Schlosser (2008), 12.

however, this inward and outward flow becomes entangled. If the Standard View is taken in conjunction with a traditional internalistic view of the mind, the mind passively receives sensory input from its environment, cognizes using internal vehicles in order to structure that input, and then uses the products of cognition to generate actions that are suitably structured by these inputs and an agent's cognitive goals. Action is a product of internal mental activity.[91]

Hurley (2001b) calls this traditional view of perception, cognition, and action the 'Input-Output Picture'. Perceptions constitute the inputs from the world into our minds, which lead to isolated cognitive processes that constitute internal mental activity. This mental activity is causally efficacious in producing our actions (both mental actions and overt actions), and in doing so constitutes the reason-states for which we act.

> This picture projects the personal level perception/action distinction and the subpersonal level input/output distinction onto one another. This isomorphic projection seems inevitable, given the two presuppositions of a linear, one-way causal structure and of merely instrumental relations between perception and action.[92]

This event-causal bifurcation straightforwardly accommodates the idea that the mental is individualized, or independent of the external world. If perceptions and intentions, beliefs and desires, are independent, then perceptions or beliefs do not lead to or explain any particular behavioral effects on their own. They only do so in conjunction with certain desires or intentions that rationalize them.[93]

Many authors who seek to challenge this picture end up with theories of action that focus on agential systems, as opposed to mental states. Control systems theories of action, which characterize action as the control of inputs into a dynamical system, are one species of these; dynamical theories of action are another. These, because some extended mind theorists have

---

[91] Hurley, Susan (2001b). Perception and Action: Alternate Views. *Synthese* 129: 3-40. 9.
[92] Ibid., 13.
[93] Ibid., 14.

directly put them forth or endorsed them, may cohere well with EMT. However, they may also be highly prone to the agency problem.

§2.7.1: Control Systems Theories of Action

Susan Hurley (2001b) addresses the benefits of control systems theories of action for the extended mind theorist. While Adams and Mele's Control Theory of Action does not presuppose extended mental states, Hurley's view is directly related to the extended mind. So, unlike control theories of action, control systems theories are built around the idea that the agent is an extended cognitive system. Such theories differ from control theories of action in that they invoke the notion of a sub-personal dynamic system that loops out into the world, the behavior of which is controlled within certain parameters. This is different from the output of an agent being controlled in various ways by her intentional mental states. There is a linear progression towards a represented goal state in the case of control theories of action, though this progression is partly a function of feedback, but not in the case of control systems theories. Hurley herself rejects the traditional models of mind and action, especially standard internalist and individualist views of the mind. The idea behind control systems theories of action is that action is control over input into a dynamical, but not internal, sub-personal system.

Some extended theories of mind argue that the vehicles of mental content are dynamic and spatially extended beyond the boundaries of the organism.

> If physical structures serve as vehicles because of their dynamics, then it is reasonable to consider the causal structure that enables the dynamics. If we do so, we find multiple feedback loops, loops that occur at different time scales and involve physical structures both internal and external to the organism.[94]

[94] Madary, Michael (2012). Showtime at the Cartesian Theater? Vehicle Externalism and Dynamical Systems. In F. Paglieri, ed. *Consciousness in Interaction: The Role of the Natural and Social Context in Shaping Consciousness* (pp. 59-72). Philadelphia, PA: John Benjamins. 61.

The brain continuously receives feedback from extended causal loops that involve both the body and the environment. This feedback is part of what constitutes our intentions, which lead to intentional behavior expressed by the system.

Hurley states that action is neither purely stimulus driven nor purely autonomous behavior; it is something in between automaticity and autonomy. Human behavior is the behavior of a system that produces effects in an environment that itself partially contributes to the production of these effects. The environment creates various disturbances after a system generates output. These disturbances can influence the effects that the system has. So the same output from the system can have different effects, reflecting its interaction with different environmental disturbances.[95] Systems that act can produce consistent results in the face of environmental disturbances. Intentional action implies precisely this kind of control, or consistent results in the face of various disturbances. A result is under control just insofar as it is protected from these disruptions, where this protection itself constitutes part of what action is. Intended results are controlled results, thus making intentional action the control of these outcomes.

According to Hurley, there are several necessary parts of a control system: a sensor, a reference signal, a comparator, and an effector. The system uses the sensory signal, compares it to the reference signal, and generates an error signal if the two fail to match. The effector then uses the error signal to produce output that counters the disturbance causing the potential mismatch, so as to keep the sensory and reference signals matching as closely as possible.[96]

> Feedback control is a cyclical and dynamic process, with no nonarbitrary start, finish, or discrete steps; input is as much an effect as a cause of output…Control depends on

[95] Hurley (2001b), 23.
[96] Ibid., 24.

> dynamic relations among inputs and outputs. Information about inputs is not segregated from information about outputs; this blending of information is preserved and extended in the informational dynamics of further layers. Perception and action arise from and share this fundamental informational dynamics.[97]

This allows for results that are under the control of the system, or intentional action.

Hurley also uses dynamic relationships to try to differentiate agential behavior from behavior of the environment. Action, she states, is the control of systemic activity. It is neither purely stimulus driven nor purely autonomous. Intentional action, on control systems theories, implies consistent results in the face of disturbance, or control. A result is under control if it is protected against disturbance *by action*. Intentional action simply is the control exercised when we produce consistent behavior. The agent in this case is the protector and controller. Agency includes the ability to produce controlled behavior, and to produce this behavior in the face of disturbance.

Hurley rejects the 'Input-Output Picture' that divides perception and action into two separate categories. It is a misconception, Hurley states, to view the behavior of the system purely as output, directed outward from an insulated system to the outside world. What a control system controls is input rather than output; output is just an arbitrarily identified component in a causal loop, not the last step in a causal chain. So, when we act we compare the input signals to output signals in order to produce behavior that meets the demands of our goals as cognitive agents. Our intentions are then dependent not only on primitive perceptual input, but also on the greater web of relations, signals, and causal loops that compose this (unbounded) system.

The content of our intentions may depend noninstrumentally on sensory feedback by depending noninstrumentally on a dynamic system of relations between input and output. This

[97] Hurley, Susan (2008). The Shared Circuits Model (SCM): How Control, Mirroring, and Simulation Can Enable Imitation, Deliberation, and Mindreading. *Behavioral and Brain Sciences* 31: 1-58. 12.

dependence is distributed across the entire extended system, not merely internal states of an agent. The content of our intentions is a function of relations within a complex dynamic system. These contents are not determined by internal vehicles of thought, but rather the wider web of cognitive functioning.

Hurley states that this kind of view is an expression of three general ideas. The first is the idea of context-dependence. Content can constitutively depend on an entire network of relations, rather than of the intrinsic properties of discrete, local vehicles. Content can both by carried relationally and determined relationally. This idea is in part a basic expression of vehicle externalism.

> Indeed, the point depends on recognizing that there is no magical causal boundary around the brain that in principle prevents the vehicles of a unified consciousness from spreading beyond it. This is the insight of *vehicle externalism*. Instead, we can think of the subpersonal basis of a unified consciousness as a kind of *dynamic singularity* in the field of causal flows: a tangle of multiple feedback loops of varying orbits, some internal to the brain or body, others involving partly external motor feedback. Such a dynamic singularity is centered on the organism and moves around with it, but it does not have sharp boundaries.[98]

The second idea is that there is no neural boundary that might prevent the first point from applying to relations between input and output. The inputs and outputs of the cognitive system are not distinctly separated, with some signals coming in and other going out. Rather, these signals have a complex interplay and superposition. The third idea is that superposition or co-dependence of contents can make for interdependence.[99] The content of our intentions is interdependent on context, as well as causal loops that leak out into the world.

---

[98] Hurley, Susan (2003). Action, the Unity of Consciousness, and Vehicle Externalism. In A. Cleeremans (ed.), *The Unity of Consciousness* (pp. 72-91). New York: Oxford University Press. 81.
[99] Ibid.

Hurley thinks that this view will actually help to solve the agency problem, which at times she seems to acknowledge. One of Hurley's ideas is that extended mental states might satisfy the criterion of having some privileged inner control mechanism. If this is the case, then the agency problem stands a chance of being resolved. The idea of a dynamic singularity is used in this manner in order to salvage agential control. Hurley (1998) states the following:

> The notion of a dynamic singularity may be the subpersonal functional idea of unity needed to complement a normative approach to the unity of consciousness and to explain the perspectival interdependence of perception and action. Nevertheless, at the personal level, what it is to be a subject and agent is open to view, just where it seems to be, in the normatively constrained lives of creatures whose perception/action system embeds them into and deals directly with the world.[100]

This passage seems to entail a view according to which at the personal level agents are embedded in a system that involves the world, while at the sub-personal level there exists some kind of dynamic singularity that binds and unifies an agent in the way that internalism allows for.

At the same time, however, Hurley defends the theory according to which our basic intentions depend noninstrumentally on external features of our environment.[101] The emergence of agency is therefore contextual in that it depends noninstrumentally on extended environmental context. This contextualization of agency extends control of our actions and outsources that control to the context in which we are embedded as agents. Control systems or subpersonal singularities cannot save agency if intentions remain extended beyond the boundaries of an agent.

There are, moreover, several problems with the view itself. The first is that there are many forms of behavior that such a theory would deem to be intentional that are intuitively non-intentionally performed (digestion, the beating of one's heart, etc.). And, while intended results

[100] Hurley (1998), 336.
[101] See Hurley (1998), chapter 9.

may be controlled results, not all intentional action is controlled (at least not in the sense being used by Hurley). Action, in this sense, *is* a protection from disturbance, or an exercise of control. Mental actions do not seem to adequately fit into this picture, as they don't seem threatened by disturbances in the first place, and even some overt actions may produce uncontrolled results while remaining intentional. Cases of this might include intentional tryings, or actions that do not reach their intended goal or produce consistent results because of some disturbance, yet still remain intentional. Thus, not all controlled behavior is action, and not all action is controlled behavior. This solution gets it wrong on both counts.

Further, even disregarding these problems with the control systems view itself, it will not save EMT from the agency problem. Because our basic intentions depend noninstrumentally on the environment external to the acting agent, the actions caused by these intentions are therefore caused by things locationally external to her. The kind of context dependence being appealed to here seems to divest the actor of her intentional sovereignty, as her intentions are context dependent as opposed to agent-dependent. If the intentions of a control system are context-dependent, then the control of the behavior of that system also depends non-instrumentally on its context.

§2.7.2: Dynamical Systems Theories of Action

Traditional conceptions of the mind and action characterize the cause of our actions as being internal mental states, largely composed of representations, which make up our causally relevant intentions. Hurley gives one alternative to this approach; Chemero and Silberstein (2011) give another, which relies on dynamic systems and their complex interactive feedback loops. These authors argue that agents are extended phenomenological-cognitive systems whose

behavior cannot be explained in virtue of these internal, detached, causal mental states. These systems are extended between the brain, body, and environment, and their behavior involves all of these components.

> In dynamical systems explanation, one adopts the mathematical methods of non-linear dynamical systems theory, thus employing differential equations rather than computation as the primary explanatory tool. Dynamical systems theory is especially appropriate for explaining extended cognition because single dynamical systems can have parameters on each side of the skin.[102]

Likewise, it might seem natural to think that these dynamic systems, and their parameters, are especially poised to explain the behavior of coupled entities comprising organisms and parts of their environments.

Dynamical systems theory (DST) has recently been put forward as a way to solve problems of mental causation.[103] "Some advocates of the dynamical approach to cognition think that DST can provide a more satisfying model of mental causation. They claim that some dynamical systems exhibit a form of global to local determination or *downward causation* in that the large-scale, global activity of the system *governs* or *constrains* local interactions."[104] The theory seems poised to apply equally well to the causation of behavior by extended mental states.

According to dynamical systems theories of action, many of our activities and behaviors involve external feedback loops. The wider causal network of feedback connections and ongoing dynamics plays a crucial role not only in our experiences, but also in the production of our intentional behavior. The explanation of our actions is dynamical as opposed to mechanistic. This is unlike the Standard View, which is highly mechanistic in nature. According to the

---

[102] Chemero, Anthony and Silberstein, Michael (2011). Dynamics, Agency, and Intentional Action. *Humana Mente* 15: 1-20. 3.
[103] See Van de Laar, Tjeerd (2006). Dynamical Systems Theory as an Approach to Mental Causation. *The Journal for General Philosophy of Science* 37: 307-332.
[104] Ibid., 308.

Standard View, reasons cause our actions in a direct, linear fashion. According to the dynamic systems model, it is a complex interplay of many different sources of feedback that structures the parameters of the dynamic agential system, whose behavior is determined in part by these parameters. These parameters act as constraints on the components of extended cognitive systems. These constraints are simply "features of the system that can impact the behavior of the system."[105] They constrain the activity of the system's components by specifying how the dynamical system acts.

Dynamical systems are not specially constituted physical systems in their own right. Rather, they are mathematical models enlisted to predict the way in which observable properties will change over time. A dynamical system itself is a mathematical description that includes a description of the parameters that structure the behavior of a nonlinear system over time. As such, dynamic systems are not themselves agents. A dynamical system itself is simply a set of rules, equations, etc. that describe how variables change over time. The physical, coupled, embodied cognitive systems that are modeled by these mathematical descriptions are more correctly called agents. These systems are not constrained by rigid spatial boundaries, such as the boundaries of a biological organism. They also behave differently than our traditional, commonsense intuitions about causation and constitution.[106] Thus, cognitive systems are a kind of natural dynamical system.

The parameters of dynamical systems place *downward* constraints on the entities that compose those systems. This is supposed to account for how cognitive/mental systems are causally efficacious when it comes to human behavior and action.

[105] Chemero and Silberstein (2011), 15.
[106] Madary (2012), 69.

> The thesis is that the (chaotic) activity of the thousands of neurons that make up the complex dynamical system that is the brain often results in global, large-scale pattern-formation that is associated with certain kinds of mental activity. It is the relation between the neural activity at the micro-level and the overall macro-activity that can be described using DST. The global activity, again, would ideally be described by an order parameter and is the result of local to global determination or 'upward causation'. However, there allegedly also is some form of global to local determination or 'downward causation' in that large-scale, global activity of the system *governs* or *constrains* local interactions.[107]

Oftentimes those invoking dynamical systems theories to explain mental behavior focus on the brain. There is no apparent in principle reason, however, for which a dynamical system could not extend beyond the bounds of the brain, or even the body.

Extended systems are one kind of system that is poised to receive dynamical explanations. These cognitive systems are best characterized as extended, non-linear, coupled, brain-body-environment systems, and are well suited to receive a dynamical explanation. The constituents of these systems are coupled such that their components are highly integrated and their behavior is a nonlinear function of the parts taken together, without boundary. In these cases, the system cannot be treated as a collection of uncoupled individual parts. And, the activity of the system cannot be explained by reference to subsystems and background conditions.[108] There is no concrete separation between the organism and the environment in which it is embedded. An appeal to the entire system, rather than its parts, is necessary for the explanation of the agent's behavior.

A dynamical systems model will not be able to accommodate the Standard View of action, as it is inconsistent with the linear, mechanistic model of primary reasons for action (belief/desire states or the mental events that produce them) being causally efficacious in a

---

[107] Van de Laar (2006), 319. See also Thompson, Evan and Varella, Fransisco (2001). Radical Embodiment: Neural Dynamics and Consciousness. *Trends in Cognitive Sciences* 5, 10: 418-425. 418.

[108] Chemero and Silberstein (2011), 4.

straightforwardly event-causal manner within the internal boundaries that encapsulate cognitive activity. Our everyday understanding of action might be very accustomed to this Standard View, but dynamical systems models of action are in part meant to motivate a rejection of these intuitions. In a complex nonlinear dynamic system, we are unable to assign precise causal roles to individual parts of the system. We are also often unable to isolate a part of the system as being a sufficient condition for producing some state of the system as a whole. This is at odds with the Standard View of agency, as well as our everyday commonsense understanding of behavior and causation. It does, however, offer a novel way of understanding the behavior of an extended system.

Intentions, according to DST, are constrained and partially determined by the higher-order parameters of the dynamical system itself. They are not determined vertically, or linearly, by reasons (constituted by mental states). Rather, they are downwardly determined by the organization of the dynamics of the entire system. The way the system, as a whole, is configured constrains what can be, and what is, intended by the agent. Intention is a higher-order natural pattern that downwardly constrains the behavior of the constituents of the system.[109]

According to Chemero and Silberstein,[110] components of the environment are 'enslaved' by the organism-system, making way for its behavior and active effects in the world. Features of an organism, such as its self-maintenance, development, and survival, act as higher-order constraints on the system. But, the internal features of the organism itself do not determine the activity of the system from within. Rather, the system enslaves the components necessary to maintain its dynamics. Because of this, a developing system will have highly flexible boundaries, and will be composed of different enslaved components over time that function to maintain the

[109] Van de Laar (2006), 321.
[110] See also Madary (2012) and Van de Laar (2006)

dynamic activity of the system.[111] The greater coupled brain-body-environment system is opportunistic in that it makes use of different resources at different times, exemplifying a kind of autonomy from parts of the world with which it is not coupled.

Intentions, rather than being composed of beliefs, desires, or other mental states, are ordered parameters of self-organizing extended phenomenological-cognitive systems that act as dynamical constraints on the enslaved components of these systems.[112] It is not feasible, in this context, to separate the individual contributions of different parts of the system, and it is thus impossible to say which states or events of the system are responsible for which behavioral processes. The only thing that we can appeal to is the greater dynamical structure and parameters of the extended system itself.

Chemero and Silberstein argue that agents are best understood as these systems, which exhibit behavior that is a function of dynamical interactions that are not explainable in terms of a reduction to the behavior of the components of a system. The total system, not the tools used within it, is responsible for action.[113] So, agency, like the system, is extended beyond the bounds of the organism.

Phenomenological-cognitive systems are just one type of dynamical system, one which agents are supposed to embody. There are three necessary components of agency in a phenomenological-cognitive system. The agent must be an identifiable individual, it (the agent) must do something, and there must be norms governing what the agent does.[114] In this case, the extended yet autonomous dynamical system is itself the agent. It behaves as an individual. And, these extended dynamical systems do things that aim at a goal state. If the goal state were not

[111] Chemero and Silberstein (2011), 9.
[112] Ibid., 15.
[113] Ibid., 13.
[114] Ibid.

attained (or, if the behavior was unsuccessful at achieving it), the behavior of the system would change in order to better function to achieve this goal. Intentions, in this sense, correspond with the intentional structuring of an action, without being something over an above the action. Intentions constrain the behavior of a system so that the normative goal of the system is met. Intentions are therefore best understood as boundaries that define the system. They are composed of the system's components and also act as constraints upon the behavior of these components.

Dynamical theories rely on a layered worldview that may be unappealing to some. The system as a whole emerges from its enslaved components, and the system as a whole has causal powers that are distinct from these components. Aside from the possible problems that this account of agency might have in its own right, it is far from avoiding the agency problem. Though the individual contributions of different parts or components of a dynamical extended phenomenological-cognitive system are impossible to separate or understand individually, it is clear that some of those contributions (individualizable or not) will extend beyond the boundaries of the organism into the softly-assembled collection of enslaved, extended, dynamical components.

Agency, in this context, is clearly outsourced at least in part to the external environment in which an organism is embedded. Though the system has some autonomy as a whole, the behavior of the system is not autonomous with respect to its parts. The enslaved components of the system make contributions to its behavior, whether or not these contributions can be singled out. On Chemero and Silberstein's account (unlike Hurley's), the agent herself is clearly extended throughout the phenomenological-cognitive system. These extended accounts of agency have problems in their own right, as well, which I will examine in more detail later on.

Dynamical systems theories of action (control systems and phenomenological-cognitive systems included) are descriptions that extended mind theorists themselves give. These are potentially insightful in that they reject the Standard View that causes problems for EMT. I argue that there is no satisfactory accounting of how extended minds can act intentionally. Though there is a plurality of theories of human action, only so many of these will be theories that the extended theorist ought to be tempted to employ. Non-reductive views of action (teleological and agent-causal) are at odds with the central naturalistic tenets of EMT. The ontologically irreducible relation between an agent that is a substance and the actions that this agent causes seems at odds with the reducible relationship between even extended mental states at the entities that constitute them. A naturalistic, reductive, theory of agency seems poised to be most appropriate for the extended theorist, but both straightforward causal and dynamic systems theories of action seems inadequate when it comes to solving the agency problem, and they may even add force to it.

Chapter 3: The Extended Mind and the Individualized Self

§3.1: Introduction

There are a few ways in which a proponent of EMT might attempt to preserve an internal locus of control for action while still maintaining that mental states are extended. One idea is that extended mental states might satisfy the criterion of having some privileged inner control mechanism. The agency problem rests in part on the existence of external environmental factors that directly contribute to the causation of our actions. However, this problem might be avoided if an agent can control her actions from within. If behavior is controlled from within a bounded agent, the causation of her actions is no longer problematically outsourced to the external environment.

The agency problem involves the claim that, when our reasons states are extended, the causal responsibility for our actions becomes outsourced to external features of an agent's environment. This, in turn, deprives the agent of causal control over her actions. This problem might be solved if the extended mind theorist can adequately argue that there is some privileged control mechanism, internal to the agent, which causes her action. If agents themselves are bounded individuals who are the causal loci of the pushes and pulls that lead to their actions, then this might help to preserve intentional agency, even if an agent's mental states are extended. If an internal proper part of an agent with an extended mind controls and guides her actions, it is possible that these actions are intentional.

In this chapter, I will address views put forth by some authors that might support this position. I find these views to be problematic for several reasons. The main problem is this: there seems to be little principled reason to favor the extension of mental states while at the same time keep the individual agent bounded, and to likewise keep a bounded locus for the control of one's

actions. It seems arbitrary to argue that the boundaries of the mind extend beyond the boundaries of a thinking agent while the boundaries of an agent, and agential control, remain internal. In addition, I will evaluate the possibility that we might simply abandon the notion that there is any central locus of control for action whatsoever. I argue that this possibility is unpalatable, and that the agency problem is even more salient if this is the case. In the next chapter, I will evaluate another option for extended mind theory – that agents themselves are extended. This I find to be inherently problematic as well. However, I will limit the scope of this chapter merely to assessing the possibility of privileging, or outright eliminating, an inner locus of control for our actions while keeping the mind substantively extended.

Overall, the motivation for accepting such a bounded view of agency is inadequate. There is insufficient reason for thinking that the thinking agent, if bounded, can adequately control her actions if the mind is extended. In addition, there is a further question as to whether it is even plausible to view the self as narrow and individualized while viewing the mind as metaphysically extended. The other logical option for the extended mind theorist is to argue that the self, itself, is extended, and, therefore, that extended mental states are proper parts of the acting agent. I will examine this possibility in the next chapter, finding it too to be inadequate, though for different reasons.

§3.2: Bounded Views of Agency

§3.2.1: Wilson, Wide Realization, and the Narrow Individual

There are several proponents of EMT that offer bounded views of the individual. I will survey some of these views in this chapter. And, though this list may not be exhaustive, it

includes a variety of different ways one could defend a bounded view of the subject, or individual, while still extending the mind.

Rob Wilson (2004) specifically asserts that the mind can extend beyond the bounds of the body, while the subject does not and in fact cannot extend beyond these boundaries. Wilson defends a view according to which the individual remains the subject or bearer of psychological states, even if she no longer serves as a boundary demarcating psychological states and processes.[115] The realization base of our mental states, Wilson argues, is wide. Some mental properties are deeply context sensitive, and some mental states are widely realized.[116] Environmental features external to the acting subject constitute the realization base of these mental states. Wilson (2014) reiterates this. When it comes to the question of whether extended cognition implies extended agents, Wilson answers negatively.[117] He states that extended cognition does not imply extended agency, which is just as well since extended agency would lead to a host of problems (a selection of which I will examine in the following chapter).

Wilson admits that the individualistic view is right when it comes to some mental states. Things like fear, pain, motor imagery, etc. involve cognitive processes that are highly individualistic. But this view of cognition as individualistic does not include all of our mental attributes, as other mental states and processes can extend beyond the skin. So, some mental states have entity-bound realizations, while others have wide realizations that include the

---

[115] Wilson (2004), 213.

[116] According to Wilson (2004), a wide realization is a total realization of a property whose noncore part is not located entirely within the individual who has that property. A radical wide realization is a wide realization whose core part is not located within the individual who has that property. A core part of a property is a state of a specific part of the system that is more readily identifiable as playing a crucial causal role in producing or sustaining the property. See Wilson (2004), 12-16.

[117] Wilson, Robert (2014). Ten Questions Concerning Extended Cognition. *Philosophical Psychology*, 27, 1: 1-15. 8.

subject's environment.[118] However, cases in which subjectivity (or, subjecthood) is not actually located in the personal self are very rare, according to Wilson.

Wilson views agency as the control of output, similarly in some ways to the control theories of action that I reviewed in Chapter 2. And, he asserts this is one respect in which the individual is a focal point even for the externalist, as the individual is the agent and controller of a cognitive system. Simply, the cognitive agent herself is individualized. Wilson suggests that the internal locus of control that characterizes symbolic capacities is compatible with the rejection of individualism.

> That is, organisms that clearly have an internal, cranial, locus of control for the core of their mental life may also possess what we might think of as a cognitive loop extending into the world beyond its own boundaries…In short, creatures like us who possess cognitive systems with an internal locus of control can instantiate internal, bodily, and world-involving cognitive capacities.[119]

The individual has certain causal powers in the world, and also a certain independence from the environment. However, what is inside the head and what is in the world are related as parts of an integrated cognitive whole, with information flowing between those parts.

Wilson states that we retain agential control by controlling the extended cognitive system of which we are a part. He admits, "control is critical to cognition being ours and to bearing on our lives as agents." Each of us forms a core part of a specific widely realized cognitive system; this is one in which we, as agents, serve as a locus of control.[120] This is why it is the individual that has psychological states, that has memories, that acts and behaves, even if these psychological states are neither taxonomically nor locationally individualistic. Wilson states, "my belief that Paris is the capital of France remains *my* belief even though it has a wide

---

[118] Wilson (2004), 140.
[119] Wilson, Robert (2000). The Mind Beyond Itself. In D. Sperber (ed.), *Metarepresentations: A Multidisciplinary Perspective* (pp. 31-52). New York: Oxford University Press. 39.
[120] Wilson (2004), 198.

realization…this is also true even in cases of radically wide realization."[121] Wide realization occurs when our psychological states are realized by extended environmental features. So, the individual must be bounded in order to remain a concrete bearer of these psychological states and in order to be an intentional agent who controls her actions. However this does not preclude some of our mental states from extending out into the world. There is thus an asymmetry between the agent, who is bounded, and her mental states, which can extend.

Higher cognitive capacities, Wilson asserts, have an *internal* locus of control. These cognitive processes are governed by cognitive pieces that are internal to the agent.

> What directs them, what leads them ultimately to govern our behavior, lies within us. This control is not always conscious but it is neither environmental nor bodily and is mediated by "what is in the head." Symbolic representation makes for truly voluntary behavior, free action, genuine choice, and rationality.[122]

Wilson grants this symbolic nature of cognition and the idea that cognition has an internal locus of control; he merely denies individualism. We control our cognitive systems, which is critical to these cognitive processes being ours and to their bearing on our lives as agents who can act voluntarily and intentionally. Each individual forms a core part of a wide cognitive system, for which she (the individual, the agent, the cognizer) serves as the locus of control. This is the main reason for which the individual remains the subject of her mental states, even if these states are extended.[123]

So, Wilson does not see a reason to privilege the boundaries of skin and skull when it comes to cognition, but he views this as compatible with privileging these boundaries with regard to the individual. These cognitive systems have an internal locus of control, though they extend out into the world. Two reasons that Wilson provides for thinking this are that organisms

[121] Wilson (2004), 141.
[122] Ibid., 186-187.
[123] Ibid., 198.

have certain causal powers, and that they exhibit a certain independence from the environment. I do not take these to be good reasons to accept the claim that the locus of control for extended cognition is, in fact, individualistic. If the individual is defined by its causal powers, then it seems as if the term 'individual' is being used to describe the biological organism, and not the self. So, this alone is not a good reason for claiming that the self is individualistic. In addition, Wilson himself suggests that the individual organism can no longer be viewed as a self-contained cluster of causal powers. It is true that the individual exhibits a certain independence from the environment, but Wilson also suggests that individuals are deeply reliant on their environments. And, if these are good reasons for privileging organismic boundaries for the individual, it also seems as though they might present a reasonable boundary for the extension of mental states, which govern the behavior of the organism. The same reasons for privileging a narrow, individualized person also seem to apply as reasons to prefer a view of the mind that is likewise narrow and individualized.

Wilson responds to a similar problem for his view. This is the challenge that if we grant to the subject widely realized (even radically widely realized) mental states, we should also grant that there could be (or, that there are) wide subjects that extend beyond the bounds of the organism. Wilson responds that, being the kinds of creatures that we actually are, there is a good basis for marking out individuals as the bounded subjects of properties, even those properties with wide realizations.[124] Wilson briefly puts forth some other criteria for defining the narrow individual in the face of wide realization. Wilson notes that "individuals – and here, as always, our paradigms are individual people and individual organisms – are spatio-temporally bounded,

[124] Wilson (2004), 142.

relatively cohesive, unified entities that are continuous across space and time."[125] The individual is part of the system that realizes mental properties, of which she is the bearer, but the individual is still definable as a narrow entity based on these features of boundedness, cohesiveness, unification, and continuousness. These properties, however, are ones that extended mind theorists have rejected and deemed insignificant when it comes to mental states. Perhaps they are significant in defining a biological organism, but Wilson must do more to argue that we ought to identify the individual with the biological organism and not with the realizers of her mental states. He gives us little reason to think that what we truly *are* are bounded organisms and not psychological beings. Wilson therefore must give an argument for why our self-constitution should exclude widely realized mental states.

Wilson also gestures at an argument, which suggests that we ought to classify extended systems similarly to how we classify narrow systems within the body. These narrow systems (the circulatory system, for instance) are not themselves individuals. They merely constitute individuals, and they realize certain properties that the individual bears. Likewise, wide systems outside of the individual are not individuals. Individuals are core parts of these systems, and these systems realize properties that the individual bears.[126] I think what this in fact suggests, however, is that wide systems should be considered individuals, and the purported individuals that form their core realizers should be considered to be narrow systems, similar to the circulatory or visual systems, that partially realize certain properties of the wide individual. It is at least tempting to think that the two cases are parallel, rather than the inverse of one another. Thus, in parallel, narrow systems in the body are not themselves individuals; they are proper

[125] Wilson (2004), 142.
[126] Ibid.

parts of individuals. And, likewise, the core parts of widely realized systems are not individuals, as they are proper parts of a greater, extended, cognitive agent.

One potential difference between narrow systems like the circulatory system and a core part of a wide system is that systems like the circulatory system would not play the privileged role of controlling the formation of transient extended systems. If individuals were narrow systems like the circulatory system, they would not play this privileged role.

There must be a difference, though, between how the organism-bound individual controls the formation of a transient extended cognitive system and the way in which the circulatory system controls the inner workings of an organism. The circulatory system certainly has wide reaching control over much of the proper functioning of the organismic body. Such control does not make the circulatory system privileged when it comes to its status as a sub-system of the greater individual. Likewise, it is unclear how the privileged status of the organism (or, perhaps, its brain) when it comes to cognitive control ought to give it the status of an individual. Certainly, cognition is important to the identity of an individual person. However, according to Wilson, its importance does not serve to demarcate the individual whose identity it is important for.

Overall, I find Wilson's defense of the narrow individual to be insufficient. Wilson does not provide enough support for the claim that the individual could be the narrowly defined bearer of mental states while those mental states outstrip the bounds of the individual. There seems to be little in the way of appropriate justification for privileging the individual in this manner. Wilson relies on a bounded locus of control to reign in his externalism such that it does not eliminate the agential subject and agential control. In doing so, he seeks to privilege the bounded individual in order to attribute to it the ultimate locus of control for cognition and action, but he

cannot do this without other pre-existing reasons to privilege the status of bounded individuals, which he rejects by way of rejecting individualism about the mind.[127]

This line of reasoning may be problematic for the following reasons: our main *motivation* for accepting the conclusion of the argument for the narrow individual is that we desire to keep the individual as the focal point for mindedness, cognition, and action. But, the satisfaction of this goal requires independent argumentation, and Wilson has not given sufficient independent reasons to motivate the acceptance of the narrow individual. He uses our prior acceptance of certain individualized traits such as agency and control in order to argue that the individual is narrow. But, we need independent reasons for thinking that we ought to acknowledge these subjective features once the mind is extended. In other words, Wilson seems to be arguing that the individual is a necessary bounded locus for cognition and action, because without this locus of control we face a problem of over-decentralization. So much, I have argued, is true. But it is not enough to stipulate that the individual is narrow in order to guard against this problem. So, in order to preserve this central control we need to posit a narrow individual, but the main motivation we have to accept this individualized view of the self is that it preserves a central locus of cognitive and behavioral control.

§3.2.2: Hurley, Dynamics, and Dynamic Singularities

Once the mind is extended, we need principled reasons to think that the minded individual is an unextended, central controller. Susan Hurley offers some reasons in favor of this

[127] While I am not directly asserting a principle of mental state internalism here, according to which the individual is wherever her mental states are, this might be what we are left with once the narrow individual with widely realized mental states is rejected. There are, however, arguments directly in favor of mental state internalism that would likewise serve to deny Wilson's thesis. See, for instance, Adams and Aizawa (2009) and (2010).

view as part of her extended view of the mind. Hurley's own view rejects the input-output picture, including the traditional view of action production. According to this picture of the mind, perceptions are the input from the world into our minds, causing certain mental states, which, after some cognitive processes, cause our actions, which are outputs from the mind to the world. Intentions, according to Hurley, depend noninstrumentally (or, constitutively) not only on these internal states but also on our relations to the environment in which we, as agents, are embedded. The content of these intentions can be determined by environmental context. We are not autonomous agents who affect the world from apart. Rather, our intentions are highly contextualized and, in fact, non-autonomous. On such a contextualist view, our intentions do not have their content intrinsically, independently of their worldly contexts and relations.[128] Behavioral output is noninstrumentally dependent on feedback loops that project out into the world and external states and processes.

The idea of a dynamic singularity is used in this manner in order to regain agential control. Hurley (1998) says this:

> We may rather need to explain input-output relations in terms of an entire dynamic singularity: a continuous and complex dynamic system centered on an active organism, with feedback loops that may have external as well as internal orbits…The standard view, which has seemed natural to many, ignores dynamic feedback and assumes that we should look at the processes leading from input to output for the explanations of input-output relations. But perhaps we should instead look at whole dynamic singularities to find explanations of input-output relations. Given the way external feedback loops may go though, and bounce off, the external world, such explanatory states may include external factors. Nevertheless, dynamic singularities are centered on and move around with organisms.[129]

This is one way to preserve an organism-centric view, including an organism-centered locus of control for action, while extending an individual's mental states. The dynamic singularity is

[128] Hurley (1998), 272.
[129] Ibid., 333.

centered on the organism, who is the agent, while the feedback loops of the dynamical system of which this singularity is a part have orbits the extend out into the world. Hurley seems to acknowledge the agency problem when she says that "relational states of persons, are after all, intrinsic states of something bigger, of systems that include persons: persons-in-environments."[130] These systems both drive and are noninstrumentally constitutive of our cognitive systems.

Hurley goes on to state the following:

> At the subpersonal level, when we look at dynamic causal flows and patterns, distinct selves show up as distinct singularities, foci of systematic and complex feedback relationships…which cluster in and around and move with the active biological organism but which also look out into the environment.[131]

This passage seems to entail a view according to which at the personal level agents are embedded in a system that involves the world, while at the sub-personal level there exists some kind of dynamic singularity that binds and unifies an agent and provides for a coherent notion of a self that is causally responsible for the behavior of the system.

To start, I think that it is up for debate whether there are such dynamic singularities in the kind of dynamic systems that describe the cognition and behavior of an intelligent organism. Quantities in the mathematical model of a dynamical system are put forth as parameters for the evolution and behavior of such a system over time. It is unclear how dynamic singularities fit smoothly into this picture. Hurley does little to define what, exactly, such a dynamic singularity would be, nor does she provide evidence that such things exist. Further, a dynamic singularity centered on an organism, which moves around with it, is not itself an individual, nor does it seem to constitute an individual. The notion of a dynamic singularity seems to fulfill a kind of

[130] Hurley (1998), 335.
[131] Ibid., 336.

centralizing function necessary for an agent-centered view of action. However, Hurley herself has argued for the *decentralization* of cognition. According to Hurley, cognition emerges from dynamic sensorimotor systems and the environmental structure of these systems.[132] The environment is part of this dynamic system.

Even if the idea of a dynamic singularity is plausible, it does not save this view from the agency problem. Hurley's contextualization of agency still takes away an internal locus of control for action. She states specifically that there is "nothing specially oomphy about causal relations inside the skin, or inside the head, nothing specially capable of pushing and shoving."[133] And, provided that this is true, it is precisely the problem. Dynamic systems or singularities cannot save agency if intentions remain extended beyond the boundaries of the agent.

Interestingly, in her other writing Hurley has been hesitant to dispute the Davidsonian picture of action. Hurley (2001b) states that a creature who acts intentionally acts for reasons, but these reasons are not the sort of individualized mental states that we traditionally take them to be. Reasons can be bound to their contexts. They are not autonomous from our extended environment.[134] Given that these are the reasons states that cause our actions, and that these reasons states are extended, then we do not cause our actions autonomously. They are caused by our context, our environment, and externalities.

Hurley also uses dynamic relationships to try to differentiate agential behavior from behavior of the environment. Action, she states, is the control of input into the dynamical system that generates behavior. It is neither purely stimulus driven nor purely autonomous. "Intentional

[132] Hurley (2001b), 4.
[133] Hurley (1998), p. 336.
[134] Hurley, Susan (2001a). Overintellectualizing the Mind. *Philosophy and Phenomenological Research* 63, 2: 423-431.

action, on [control systems theories], implies consistent results in the face of disturbance, or control." The resulting action is under control if it is protected from disturbances in the surrounding environment. In short, intended results are controlled results.[135] The agent in this case is the protector and controller. Intentional action simply is consistent behavior in the face of environmental disturbances.[136]

There are various problems with this view, which I brought up in the preceding chapter. The first is that there are many forms of behavior that such a theory would deem to be intentional that are intuitively non-intentionally performed (digestion, the beating of one's heart, etc.). Second, the idea that intentional action is controlled input seems to presuppose an input-output picture of perception and action, which Hurley is loath to support.[137] Finally, while intended results may be controlled results, not all intentional action is controlled (at least not in the sense being used here). Actions may produce uncontrolled results while remaining intentional. Thus, not all controlled behavior is action, and not all action is controlled behavior.

§3.2.3: Clark, Ecological Control, and Agency

Clark and Wilson (2009) state that "we properly expect our individual agents to be mobile, more-or-less reliable, bundles of stored knowledge and computational, emotional, and inferential capacities," and so we need to be persuaded that the new capacities enabled by the addition of the notebook in a case such as Otto's, or some other extended mental state or cognitive process in a different case, "are likewise sufficiently robust and enduring as to

[135] Hurley (2001b), 25.
[136] For more on this, refer to Chapter 2, Section 2.7.1 of this work.
[137] See Hurley (1998) for an explication of the input-output picture and her refutation of it.

contribute to the persisting cognitive profile we identify as Otto-the-agent."[138] So, there is some persisting cognitive profile that endures with, and perhaps constitutes, the living, breathing, acting, cognitive agent. More recently, Clark seems to endorse a view of the agent, or the self, such that agents are actually extended conglomerates instead of individualized creatures. I will evaluate this possibility in the next chapter. Here, I will focus on earlier views that are informative in this context.

Clark has put forth multiple views of cognitive agents. I will survey some these views here, as I think it is useful to compare them to the other views addressed in this chapter. Clark (2005) suggests that the brain is the bounded locus of control for action. He states that biological brains are by nature open-ended controllers. "To deal fluently with bodily change and growth, they have developed ways of computing, pretty much on a moment-to-moment basis, what resources are readily available and under direct control."[139] The biological brain is a master of what Clark calls 'ecological control,' which is the kind of control that allows substantial devolvement of power and (functional) responsibility. It allows much of our thought and reason to depend upon the robust and reliable operation, in dense brain-involving loops, of a variety of non-biological epistemic devices (such as pen, paper and sketchbooks). Ecological control systems take these devices and build them into problem-solving routines. They exhibit a management style "delicately poised midway between anarchy and enslavement." It is our biological nature, Clark claims, to be open to many forms of physical and cognitive hybridization, and some of these may be so intimate as to effectively extend the thinking

[138] Clark, Andy and Wilson, Robert (2009). How to Situate Cognition: Letting Nature Take Its Course. In M. Aydede and P. Robbins (eds.), *The Cambridge Handbook of Situated Cognition* (pp. 55-77). New York: Cambridge University Press. 67.
[139] Clark, Andy (2005). Intrinsic Content, Active Memory and the Extended Mind. *Analysis* 65: 1-11. 5.

agent.[140] Clark (1999b) states that "biological brains, are, at root, controllers of embodied action."[141] The brain is still the controller of embodied, situated action in the way that the environment (in which the agent is situated) is not. Ecological context, while it is important for cognition, is not the controller of action in the way the brain is.

At the same time, however, EMT claims that the external states and processes that constitute our extended cognitive states can and do cause our behavior. Thus, it seems as if external context needs to be a controller of action. However, one might maintain that EMT can accommodate these claims at a deeper level. For, it may be the case that extended mental states can cause our behavior in unexpected ways, but do not control that behavior. Or, perhaps external mental states cause our behavior by being part of feedback loops that 'loop' through the brain. The brain could be the center of ecological control for these feedback loops. The brain may constitute a kind of dynamic singularity that these feedback loops, as part of a greater dynamical system, move through. There must be some further explanation to ensure the compatibility of these claims.

I think these are the only options for making such a view compatible with a robust form of the extended mind thesis. And, I have given reasons to reject such views in the previous sections. These are probably sufficient reasons alone to reject this particular view. There are, however, other independent theoretical reasons for rejecting the view that the brain might be a central controller for an extended system.

Once the thinking agent is extended, it seems as if the brain is no longer the center of control, ecologically speaking or otherwise. Clark (1999a) states both that the brain is a central

[140] Clark (2005), 9.
[141] Clark, Andy (1999b). Where Brain, Body, and World Collide. *Journal of Cognitive Systems Research* 1, 1: 5-17. 14.

controller for actions and that the environment is a fundamental component of our problem-solving behavior.[142] An agent cannot at once have a brain that is the locus of action-oriented techniques and control while at the same time being a minded agent whose cognitive states extend across ecological boundaries. Once a system is a hybrid one, it is not clear in what sense it is the brain exploiting external tools rather than these tools participating in systemic activity that involves the brain. For a system to be a genuine hybrid it seems as if control must be shared across the brain-environment boundary.

Clark (2008a) states that there is still a sense in which the body is an important boundary, even for the extended mind theorist (and, apparently, even when the agent is extended, a thesis that the author defends in the same work). Clark states that "the body is the point at which willed action, if successful, first impacts the wider world…these points of willed action include all our voluntary sensor movements."[143] This yields an understanding of the body as a persisting locus of voluntary, intentional, action. The body is also a gateway to intelligent offloading, as it is the primary tool that enables the (cognitive) use of environmental structures. The body is a key physical structure whose cognitive role is to allow differing resources (neural, bodily, and environmental) to become highly integrated, enhancing the problem solving capabilities of the human agent.

This, however, is a cognitive role that the body plays. This is not sufficient evidence, considering Clark's prior claims that agents are extended systemic wholes, to support the claim that the body is a boundary for action production, or the limit where action impacts the wider world. While it may be the case that the body is a point where some intelligent action impacts the wider world, we have little reason to think that it is the only interface between willed action and

[142] Clark (1999b), 13.
[143] Clark (2008b), 206-207.

the environment. Consider a case where an agent develops the capacity to use a tool fluidly, without conscious thought. A blind person using a white cane to detect objects in her path could be an instance of this phenomenon. If the cane is part of her cognitive system, functioning as a perceptual vehicle, then it seems as if it is the end of the cane, and not the end of the agent's body (her hand that holds the cane), which constitutes the interface between the agent and the world.

Clark (2007a) states that "it is important that there is no need, in taking extended cognition seriously, to lose our grip on the more-or-less stable, more-or-less persisting biological bundle that lies at the heart of each episode of cognitive soft-assembly." [144] It is only under strict and rare conditions that this more-or-less stable and persisting core genuinely extends out into the environment. This would only occur in cases where the persisting core is augmented in a potentially permanent manner with external resources playing a stable role. However, in almost all cases, "we confront only temporary medleys of information-processing resources comprising an integrated subset of neural activity and bodily and environmental augmentations."[145] The mere fact that such medleys are temporary, however, does not give us good reason to deem them non-cognitive.

Clark notes that we don't identify individual agents by looking for their cognitive mechanisms. "Instead, we find an agent by finding (roughly speaking) a reliable, easily identifiable physical nexus of perception and action, apparently driven by a persisting and modestly integrated body of goals and knowledge."[146] When it comes to the process of recruiting

[144] Clark, Andy (2007a). Curing the Cognitive Hiccups: A Defense of the Extended Mind. *Journal of Philosophy* 104, 4: 163-192.169.
[145] Clark (2007a), 169.
[146] Ibid., 170.

external resources, it is the biological brain "that is in the driving seat."[147] This core of the system, which lies within the organismic boundary, plays a privileged role in recruiting external props to help fulfill the creature's cognitive aims. While, in adopting an extended view of cognition, we reject a view of cognitive processing that is *organism-bound*, we should not feel forced to deny that cognition is *organism-centered*.[148] Once we adopt this extended view, "the outer and inner operations are free to emerge as well-turned co-active participants in the construction of thought and reason."[149]

Clark thereby advocates HOC (Hypothesis of Organism-Centered Cognition), which states 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.[150]

The human organism is the "handy container of persisting recruitment processes and of a batch of core data, information, and body-involving skills."[151]

Clark acknowledges that much of the opposition to his extended theory may come from the fear that by recognizing extended systems as genuinely cognitive we thereby lose sight of the "crucial cognitive core" of the extended system. "The fear would be that to embrace hybrid cognitive forms is to lose sight of the unique importance of the core systems upon whose successful operation *the very possibility of such extended forms depend*."[152] Along with this, Clark allows that it is the biological brain that is "in the driver's seat" of the recruitment

---

[147] Clark (2007a), 175.
[148] Ibid., 175-176.
[149] Ibid., 191.
[150] Ibid., 192.
[151] Clark (2008b), 138.
[152] Ibid., 108.

process.[153] And it is predominantly (though not solely) the biological organism that "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."[154]

Since there is an internal locus of control identifiable with the agent, or her subsystems, one could argue that the agent is responsible for her actions. She controls the external components of the temporary extended system, which in turn controls her actions. Our mental routines are self-engineered, as we ourselves create the extended system that constitutes our cognitive workings.[155]

There are several things to note when it comes to these passages. First, if this reply is to succeed it cannot be that what is doing the controlling of the soft-assembled system is the system itself; it must be something organismic within the system. Clark acknowledges that the brain surely plays a very special role here.[156] Clark also, however, wants to reject outright the idea of an inner central executive within the extended cognitive system.

> In place of such an all-knowing inner executive, we should consider the possibility of a vast parallel coalition of more or less influential forces, whose largely self-organizing unfolding makes each of us the thinking beings we are.[157]

No one element is the privileged source of thinking. When speaking about gesture and its role in cognition, Clark states, "The wrong image here is that of a central reasoning engine that merely uses gesture to clothe or materialize performed ideas. Instead, gesture and (overt or covert) speech emerge as interacting parts of a distributed, semianarchic cognitive engine…"[158] It seems,

---

[153] Clark (2008b), 122.
[154] Ibid., 123.
[155] Ibid., 60.
[156] Ibid., 122.
[157] Ibid., 131.
[158] Ibid., 133.

though, that Clark's remarks on neural/bodily control of the extended system invite talk of some kind of central executive.

This organism-centered view suggests an inner executive or central controller, since it privileges an internal subsystem that controls the partially external soft-assembled cognitive system. If we take seriously the idea that there is no central controller, then there cannot be a cognitive subsystem or organism-system that controls the extended system in a way that would give the agent control over her activities. The biological organism and/or the brain cannot at once be in control of the process of recruitment and external cognitive structures while at the same time there exists no central controller of the system. It is true that "soft control", as Clark calls it, does not entail an all-knowing inner homunculus, but it does require a bounded source of control.

Clark at once asserts that "neural goings-on are not blessed with some intrinsic property that makes them alone suitable to act as the circuitry of the mind and intelligence"[159] and asserts that the brain plays a special, privileged role in cognitive processing. If the brain is in charge of the soft-assembled system, which it must be to fill this privileged role, then there is an inner executive in the system. It may not be an all-knowing inner homunculus, but it acts as a center for control within the system. Maybe there is no *intrinsic* property that makes neural goings-on suitable for cognizing, but there must be some property of the organism and/or the brain that makes gives it this privileged role in the cognitive system. Clark states that we ought not to elevate the brain/CNS or even the biological organism at the expense of the rest of the components of the system, but that seems to be exactly what he has done in order to resuscitate an inner locus of agential control.

[159] Clark (2008b), 136.

It remains to be explained how an internal executive would cohere with the rest of Clark's assertions. Clark claims that we are profoundly-embodied, natural born cyborgs, promiscuously world-involving hybrids. When speaking of an inner executive, he states that "in the *absence* of any such privileged inner component, the outer and inner operations are free to emerge as well-tuned coactive participants in the construction of thought and reason."[160] If there is no privileged inner component, if cognition really is a hybrid process, then the problem at hand has not been solved. Once external resources are recruited into the system that includes the brain and the organism, while they may not be at the core of the system or its most active elements, they involve at least some control over our behavior. So long as the organismic core and the external world are co-active participants in driving behavior, the problem has not been solved. Cognition being organism-centered or necessarily organism-involving does not solve the problem in and of itself. The locus of control of cognition must be organismic and agential.

If Clark is correct that the organism and/or its brain plays a privileged role in recruiting and, more importantly, controlling external cognitive processes, then there seems to be an important sense in which it is the organism that uses the external tools in play rather than them being cognition-constituting. There also must be a kind of inner executive within the cognitive system in order to render this story plausible. This seems to go against the spirit of the extended mind theory, according to which external resources are active participants in guiding our behavior. No element of the extended system is supposed to be a privileged source of thought or thinking. The flow of control is taken to be fragmented, the flux of recruitment somewhat anarchic, with different elements gaining control at different times. We can view external cognitive resources as neither the products of, nor the servants of, a single executive controller.

[160] Clark (2008b), 137 (emphasis added).

It is one thing to assert that EMT does not force us to abandon the organism as a subject of scientific study and another to assert that the organism and/or its brain is a privileged locus of control for the soft-assembled system of which it is a part. Clark seems to want to make both assertions, while at the same time supporting a view according to which there is no central controller, control is fragmented, and the system is (at least somewhat) anarchic. A true privileged locus of control could not exist in such a system.

§3.2.4: Wheeler, Action-Oriented Representations, and Nontrivial Causal Spread

Wheeler (2005) at times appeals to internal, egocentric, action-oriented representations in order to re-define a bounded locus of control for action. These sub-agential representations are tailored to do the job of producing behavior. These are essentially outcome-directed, behavior controlling, action-specifying, and context dependent.[161] Action-oriented representations (AORs) are not abstract and detached, but instead specifically geared to produce specific kinds of actions. "AORs are assumed states or processes, local or distributed, whose functional role is to indicate the presence of, and sometimes to 'stand in' for, states of affairs in order to guide and drive certain kinds of action."[162] These are context dependent in a rich sense in that extended processes that involve the agent, her behavior, and her environment determine their content and character. They are action-specific in that they are tailored to the job of producing behavior. And, they are egocentric in that they involve agent-centered coordinates, including the situated position of the agent in her environment. They are context-dependent insofar as their character is determined by the fact that they "figure in a process in which what is extracted as contextually relevant is fixed

[161] Wheeler, Michael (2005). *Reconstructing the Cognitive World*. Cambridge, MA: MIT Press. 197.
[162] Hutto, Daniel (2013). Exorcizing Action Oriented Representations: Ridding Cognitive Science of its Nazgul. *Adaptive Behavior* 21, 3: 142-150. 143.

by the needs and previous 'experiences' of a (mildly) intelligent agent."[163] The environment is therefore determinate of the content of these representations in a substantial way.

Wheeler argues for the existence of what he calls *nontrivial causal spread*. Nontrivial causal spread arises "when the newly discovered additional causal factors reveal themselves to be at the root of some distinctive target feature of the phenomenon of interest."[164] We expect a purely neural system to be at the root of our intelligent behavior, but we find a highly distributed system involving contributions from the nonneural body and the environment plays a crucial role in the production of this behavior. These nonneural and extra-bodily factors contribute to the generation of intelligent action directly. In cases of *extreme* nontrivial causal spread, extraneural factors account for the distinctive features of intelligent action, while the neural contribution is minimal. However, in normal cases of nontrivial causal spread, "a key role in promoting behavioral success is reserved for neurally located action-oriented representations."[165]

Wheeler states that there is an important asymmetry between different parts of an extended causal system. These action-oriented representations are causally efficacious *sub-agential* states. Though nontrivial causal spread exists, where external features of the environment are determinative of our behaviors, and though these externalities are responsible to a significant degree for the production of our actions, neural states have a privileged place as the center of this system. They are responsible for coding our action-oriented representations, while the external environment provides an ecological backdrop on which our actions depend. Wheeler states that we must commit ourselves to the view that there exists an asymmetry between the different parts of the extended causal system. So, it must be legitimate, in spite of the presence of

[163] Wheeler (2005), 199.
[164] Ibid., 207.
[165] Ibid., 226.

nontrivial causal spread, for us to "treat the nonneural body and the environment as a normal ecological backdrop against which those putatively representation neural states function as action-specifying codings."[166] We ought therefore to privilege internal states in an asymmetric fashion, while treating the environment as a mere backdrop for our behaviors.

However, Wheeler does endorse the view that intelligent action can be, and is, produced by cognitive mechanisms featuring nontrivial causal spread, where extraneural factors do the important work in accounting for the distinctive features of an agent's intelligent actions.[167] Bodily and environmental dynamics, therefore, may be a significant part of what causes our actions. He states that cases where behavior is generated through extreme nontrivial causal spread will be rare, for if we look at the range of human (or, animal) behaviors we should see that little of this behavior is generated by extreme nontrivial causal spread. Rather, the subagential case, where action-oriented representations generate behavior (in the midst of nontrivial causal spread) is likely to be more common.[168]

Wheeler gives little argument for his assertion that cases where behavior is generated through extreme nontrivial causal spread are rare. And, I think we have little reason to accept this claim. If the mind really is extended, and if our cognitive processes are in part constituted by environmental realizers, then I think we have good reason to believe, as an extension of this, that nontrivial causal spread will be a large part of what produces our actions.

Wheeler seems to be defending agency in the face of the extended mind. He even states that "the naturally occurring intelligent agent is principally a locus of real-time adaptive

[166] Wheeler (2005), 210.
[167] Ibid., 226.
[168] Ibid., 230.

action."[169] At the same time, however, Wheeler defends the idea that our online intelligence, which produces our behavior, depends richly on nontrivial causal spread. The body, he states, is not the vessel for cognition, and the environment is not the stage on which action takes place.[170] Online intelligence is the kind of real-time, action-generating cognition that contributes to our behavior as we navigate through the world. Nontrivial causal spread can have many instantiations, but, in part, it is responsible for our online intelligence and cognitive processes. This is more than the enabling ecological backdrop that contributes, instrumentally, to our cognitive processing. It is an extension of an agent's cognitive system into the world.

This representational defense of agency is either insufficient or counterproductive. Action-oriented representations remain context-dependent. Either they are context-dependent in a rich, constitutive sense such that the environment plays a role in constituting their content. Or, they are context-dependent in a weaker sense such that the environment plays a mere enabling or instrumental role in their production. If the former is true, then, egocentric or not, these representations are constituted by extra-bodily elements, which in turn play a decisive role in the production of intelligent behavior. If these representations are context-dependent in the weaker sense, then it seems as if a vast portion of our online intelligence is constituted by purely egocentric mental states and cognitive processes. This poses a problem for someone arguing that mental states and cognitive processes are extended. If our online intelligence, cognitive architecture, and behavior-producing states are not only essentially egocentric but also importantly individualized, this gives us a narrow view of the agent, but also a seemingly narrow view of the mind. Causal spread at this point appears trivial, as opposed to nontrivial, and

[169] Wheeler (2005), 284.
[170] Ibid.

environmental dynamics seem enabling rather than constitutive factors when it comes to cognition and action.

If anything, Wheeler's defense of agency comes at the cost of privileging not only the individual, but also the individualized *mind* over the extended resources available to cognition. Nontrivial causal spread cannot at once be responsible for mental states and mental processing and at the same time not be fundamentally responsible for the production of our actions. If bounded properties provide the coding for representations, as well as other necessary conditions that lead to behavior, and the environment merely provides background conditions for the production of this behavior, then Wheeler seems to have taken a turn towards a bounded theory of the mind. If, however, the environment is responsible for our mental states in a nontrivial way, then Wheeler cannot claim that the bounded mind is privileged with respect to our actions. The environment cannot at once be nontrivially responsible for our actions while at the same time merely instantiating the background conditions for them.

Clark (1997) endorses a view similar to Wheeler's. The brain, according to Clark is not a locus for inner descriptions, or representations, of external states of affairs. Rather, it is a locus of inner structures that "act as operators upon the world via their role in determining actions."[171] Clark also endorses a view (in the same work) in which the brain is the center of our action-oriented representations. These, similar to Wheeler's use of the term, are representations that "simultaneously describe aspects of the world and prescribe possible actions; and are poised between pure control structures and passive representations of external reality."[172] While internal and representational, these structures are not passive. They prescribe possible actions and are

[171] Clark, Andy (1997). *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT Press. 47.
[172] Ibid., 49.

geared towards action production. Our brains evolved as controllers of our bodies. They have computational and (at least minimal) representational resources that are complementary to the real-world actions and interventions that *they control*. Action-oriented representations remain internal to the individual, they are internal states, but they are "simultaneously encodings of how the world is and specifications for appropriate classes of action."[173]

Clark's view, however, suffers the same problem as Wheeler's. Clark (1997) continues to assert that features of the body and the environment determine the content of these representations. They are therefore context-dependent in a strong sense. Further, while these internal representations may offer some prescriptions for action, they do not control actions. They may be part of the causal chain that leads to action, but they are not sufficient to determine action performance. They therefore do not determine our actions in the way needed to solve the agency problem. The mind continues to escape its natural confines and leak out into the world, and environmental interactions still seem equally poised to govern our behavior.

Simply put, the point about the narrow individual is this: The extended mind theorist, at least according to many versions of EMT, is clearly seeking to eliminate the privileged status of the boundaries of skin and skull when it comes to defining the realizers of our mental states and cognitive processes. If the extended mind theorist wishes to then privilege such a boundary when defining the individual, she must give reasons for which this boundary should be privileged. However, these reasons cannot apply to mental states as well, for if they do then the proponent of EMT would be inhibiting her own philosophical program. There are reasons to think it is simply untenable, given the central commitments of EMT, that some bounded part of an agent may be classified as the uniquely located, causally-sourced locus of control for cognition and action.

[173] Clark (1997), 151.

These solutions privilege the boundaries of skin and skull by putting cognitive control exclusively within those bounds. They seem to covertly support a bounded-mind thesis that runs counter to the goals of the extended program. Limiting actions in various ways by requiring that some integral part of the action sequence be bounded by the rational animal conflicts with the claim that there is nothing special (when it comes to action and a subject's agency) about states bounded by a person's organic body.

§3.2.5: Control and Sourcehood

It might be thought possible that the reasons states (mental states, beliefs and desires) that control actions remain internal but the cognitive machinery that executes actions is partly external. I believe that the same problem arises, however, if this is the case. If the cognitive machinery that executes our actions is at least partly external, then control over one's actions is no longer internal. It is doubtful whether it could truly be that our reasons states control our actions while the cognitive machinery that executes them is separate and external to the thinking agent. It is not clear that this cognitive machinery is separable from these reasons states in the way that the former could be extended while the latter is not. And, even if this is possible, I believe that it deprives the person of the kind of cognitive control necessary for genuine agency.

A related issue is that the initiating components of the action may be located entirely inside of the organism. While external tools may include action-relevant or action-enabling information, or even components necessary for action execution, these are not the direct *initiators* of bodily actions. This objection stresses that it is the initiation of action that we are concerned with when we talk about agential sourcehood. External states are not the sole causes of my actions, as they are not the source of these actions. And, my agency is not undermined by

external states that may be part of the causal chain that leads to behavior. Our behavior, while it may involve external cognitive states, is internally sourced and therefore properly agential.

This is not a response that the proponent of extended dynamical agency can give (see Chemero and Silberstein 2011). If the brain, body, and environment are coupled together in a non-linear, dynamical way, the activity of an agent cannot be ultimately explained by decomposing the extended system into subsystems, one of which initiates and controls behavior.[174] It is the entire extended system that engages in action, not merely the organism. The entire system, not one subsystem, is responsible for action.

These claims are insightful when it comes to other, nondynamical, forms of extended cognition, but they are not entirely compatible with the kind of extended view that many theorists espouse. In order for this solution to work, it must be the case that external cognitive entities are passive in generating of behavior. But, this is in direct competition with what many extended mind theorists claim. For the extended theory, it is the active nature of these external artifacts that is most intriguing. They are able to guide various forms of behavior. Clark (2008b) states these external states can guide, shape, and characterize a person's behavior.[175] One of the things that Clark and Chalmers (1998) originally set out to show was that external memory may sometimes be poised to control action in very much the same way as internal memory.

If reasons for action are the cause of actions (i.e. if they are the source of action), and reasons for action are at least partially extended, then the source of our actions will be extended as well. To deny this would be to deny that extended states are active in the way that active externalism claims that they are. This emphasis on the part of the extended theorist on the action guiding potential of external states may be due in part to the fact that, according to her theory,

[174] Chemero and Silberstein (2011), 4.
[175] Clark (2008b), 160.

what makes something a realizer of a functional state (for instance, a belief state), be it internal or external, is the network of causal relations in which it is embedded and the role that it plays.[176] If it no longer bears any behavior-generating properties, then a state like a belief may no longer seem belief-like at all. In a sense, the external artifact is dysfunctional as a realizer of that state. If states like beliefs that in part cause our actions can only be active when read off internally, such that they are then internal beliefs, then extended cognition no longer seems extended any more, but rather a strictly internal affair.

Further, there is a central misunderstanding at issue with this objection. Intentional action requires not just initiation but also guidance and control (see Section 7.3). The mere initiation of behavior is not sufficient for agential responsibility. Intentional action requires guidance and control of that behavior. Every action is characterized by having a goal, but merely achieving a goal state does not indicate that someone acted intentionally. The goal could have been reached accidentally even if it was what the agent desired. Internal states must function as the source of agential control, which guides actions. But, if external states play an active role in guiding, shaping, and controlling behavior, then we as agents have lost this agential guidance and control. The agency problem is going to rule out any cases as counting as action where an external entity is supposed to play a genuinely guiding role.

§3.3: Can EMT Eliminate a Bounded Locus of Control?

Clark states that there are two issues here – one of the locus cognitive control and one of the nature of cognitive processes. One of Clark's responses to a similar problem rejects the

[176] For more on functionalism, see Sections 4.4 and 6.2 of this work.

bounded locus of control as a criterion for the demarcation of cognitive states. He responds to Butler (1998) who states the following:

> ... there can be no question that the locus of computational and cognitive control resides inside the head of the subject [and involves] internal processes in a way quite distinct from the way external processes are involved. If this feature is indeed the mark of a truly cognitive system, then it is a mark by means of which the external processes Clark and Chalmers point to can be excluded.[177]

The suggestion is that the biological brain has the final say, and that this is a difference that makes a difference, cognitively speaking. The brain of an organism is the controller and chooser of things in a way that external things are not. Because it is the locus of control in a cognitive system, Butler argues, it ought to delineate the boundaries of a cognitive system.[178]

This is clearly not the argument that I am making here, as I am not making an argument about the mark of the cognitive or what is poised to constitute a cognitive system, but Clark's response to it is still, I think, informative in some ways. Clark states that this kind of opposition to the extended mind hypothesis trades on a deeply mistaken view of the thinking agent as being some sort of distinct inner locus of control. Given these comments, it seems as if Clark's response might be just as apt to go against the line of argumentation that I have been putting forth thus far, which does trade on a view according to which the thinking agent is, or ought to be, the controller of action-producing cognitive processes.

Clark asks us to reapply the 'locus of control' criterion *inside of the head*. He then asks the following:

> Do we count as not part of my mind or myself any neural subsystems that are not the ultimate arbiters of action and choice? Suppose only my frontal lobes have the final say –

[177] Butler, Kieth (1998). *Internal Affairs: Making Room for Psychosemantic Internalism*. Dordrecht: Kluwer. 205.
[178] Butler (1998), 212.

> does that shrink the real mind to just the frontal lobes? What if...no subsystem has the final say? Have the mind and the self just disappeared?[179]

Minds do not require central controllers in order to exist. If they did, any cognitive process that was not centrally controlled would not count as part of the mind, and if there is no central controller that we can identify minds seem to disappear entirely. The mind does not expand and shrink depending on whether it is centrally controlled and by the location/size of its central controller. Clark rejects the idea that there is such a central bounded controller, and he further states "even if there were some distinct inner locus of final choosing, there would be no reason at all to identify that with the mind or the 'cognitive agent'."[180] In short, there is little reason to think such a central bounded locus of control exists, and even if there were, we ought not to conclude that this should limit the extension of cognitive states.[181]

Clark appears to be making two differing claims. First, he seems to claim that our view of agency as involving some internal control mechanism is mistaken. For, all an agent *is* is a cognitive agent who behaves in the world. So, if cognitive agents are not bounded centers of control, agents, as I have been using the term, are not either. Second, Clark claims that even if our view of agency is not mistaken, there is no reason to identify the bounded locus of cognitive control with a cognitive agent. In an initial reply to Clark, I want to argue that these claims are flawed. A bounded locus of control is something necessary for genuine action, bounded within an acting agent, and is a *subset* of that agent's mental states. Crucially, this bounded locus of control must be a proper part of an agent and partly constitutive of her identity. It is not, however, *identical* to the agent herself.

---

[179] Clark (2010), 55-56.
[180] Clark (2010), 56.
[181] Wheeler (2005), 65 offers a similar, though no less problematic, perspective on this issue.

If the above view of agency is mistaken in the way that Clark thinks it is, then many, if not most, traditional views of agency are mistaken as well. It is unclear what sort of picture of agency we are left with if we abandon the idea of cognitive control. I think we are left with only teleological accounts of agency (though even some of these will be excluded), which I have previously argued are incompatible with the extended mind thesis. So, it turns out that if we abandon the idea of control as a partial criterion for cognitive agency, the extended mind theorist is left with no view of action whatsoever.

Agency without any form of internal control would be severely impoverished. Clark does not give a positive view of what it would mean to be an agent in the absence of control. It seems that the only way to make sense of agency in this context is to consider it to be a grand illusion. The only view of agency supportable according to his line of thought is one according to which individuals (as biological organisms) are not genuine intentional agents, but merely part of a greater social or situational view of agency. I will argue against the idea that genuine, individual, intentional agency might be merely illusory or a useful heuristic concept in the coming chapters (see, in particular, chapter 5). Individual agency is, in fact, presupposed by EMT and may be necessary for its success as a theory.

Clark concludes, finally, "the argument from ultimate control does not reveal the mark of the mental, or the source of the self."[182] The argument from ultimate control ultimately fails, according to Clark, as it does not help us identify the mark of the mental any better than the argument for the extended mind. The agency argument, however, is not meant to show the mark of the mental. It is not intended to formulate a principled set of boundaries for mental states and cognitive processes. It is true that viewing the bounded locus of control as a proper part of one's

[182] Clark (2010), 56.

total mental states does not allow for the locus of control to delineate the boundary for mental states. However, the argument does serve to illustrate a significant problem with EMT. For if EMT is true we lose the individual as a privileged, individualized agent who can uniquely be the author of intentional actions that affect changes in the world.

Clark states that a cognitive task, even a very ordinary one, will often be addressed by a "soft-assembled coalition of distributed (and often highly heterogeneous) neural components and brain areas."[183] Our inner cognitive economy is fragmented, held together through integration that is mirrored by our engagement with the external world. This, too, can contribute to our cognition in unexpected, and perhaps quite fragmented, ways. Clark states that when we "sack the inner executive, embrace the motley crew of cognitive processes, and the fragmentation of the flow of control, take seriously the brain's own stunning indifference to what gets done where", the familiar boundaries that a bounded view of cognition goes to such lengths to preserve begin to look ad hoc and unrevealing indeed.[184] We can explain cognition without such an inner executive by appealing to diverse and fragmented cognitive processes that occur across the boundaries of the brain, body, and environment.

I think that there are, in fact, reasons for which cognitive functioning ought to be considered unified and controlled, even if it is extended. The unity of consciousness and the unity of directed cognitive functioning would seem inexplicable without some kind of control that is unfragmented, even if it is distributed. This is true even if we can make sense of agency in the absence of a robust, concrete metaphysical notion of cognitive control.

Further, when we sack the inner executive one of two things seem to surface. Either the agency problem becomes even more aggressive, or it seems to evaporate, but our agency itself

[183] Clark (2008b),137.
[184] Clark (2008b), 138.

seems to evaporate along with it. The problem may be fiercer because, without any inner control mechanism it appears that external structures and resources are even more responsible for the production of our actions. There is no glimmer of hope of these cognitive processes looping back through some inner center of control along the path to action production. Actions are caused by extended mental states in the absence of control, and thus produced by a highly heterogeneous environmental resources and structures. In this case, the environment is vastly responsible for the production of our actions.

There is another problem, though, when actions are produced in the *absence* of an executive controller. If this is true, then it seems as if actions are being produced in the absence of agency. Really, they are not actions at all. Similar to the doings of a dishwasher, or a computer, or a robotic vacuum cleaner, they are simply behavior, or mere happenings. At the very least, even if our reasons states are extended, without any central control or central executive, then talk of action becomes fast and loose. We may still have beliefs and pro-attitudes that cause our actions, but we seem to have lost a causer, a chooser, or a doer of the action performed. It is not clear that the events and processes constituting these reasons states are at all agent-involving, or even sufficiently agent-involving to motivate the claim that the behaviors they produce are actions. By denying the centrality of control, Clark seems to be depriving the individual of a centralized action plan. These 'actions' ought therefore to be considered borderline because, though they still seem to be responsive to reasons states like genuine actions are, they lack authorship and control, making them more like systemic behaviors.

Further, even if we do end up rejecting a bounded locus of control as necessary for action, a problem that I will review at length in the next chapter – the problem of action proliferation – appears. Rejecting control as necessary for action seems to lead to an

impoverished view of agency, where agents are not the ones that are causally responsible for their actions. And, cognitive creatures are not really the *agents of* their actions at all. And, if control over our actions is extended and external, then our mental states are involved in, and therefore we are involved in, *too many actions*. Our mental states are then responsible for many actions that we (in the physical sense) do not actually engage in. In a sense, then, we would be actors who do nothing because we can do virtually anything. This is also a problem for the idea that the self *itself* is extended, and one that I will come back to in the next chapter.

Chapter 4: The Possibility of the Extended Self

§4.1: Introduction

One natural extension of the extended mind theory is to postulate that cognitive agents *themselves* are extended, along with their mental states.[185] Clark and Chalmers (1998) suggest that this is an acceptable consequence of the theory. Otto's notebook, for instance, is not merely constitutive of some of his mental states. It is also a part of his self. Clark, more recently, has defended the view that the agential self, itself, extends out into the world, such that parts of the world not only comprise part of our cognitive machinery, but also, more profoundly, our *selves*. If this is true, it has positive implications of the extended mind hypothesis when considering the agency problem. The agency problem states that if our reasons states are extended beyond the bounds of the individual then that individual is not causally responsible for her actions. But, if our selves are extended, then extended reasons states remain internal to us when they cause our actions. We can therefore control our actions even when the reasons states that cause them extend beyond the boundaries of skin and skull.

Clark calls these extended beings "profoundly embodied agents."[186] He rejects the idea that we are "locked-in agents", or agents who are bounded by our bodies. He states that this is a mistaken vision of who we are and of what fundamentally constitutes our humanity. Profoundly

---

[185] Clark and Drayson (forthcoming) state that the extended self may not be a *necessary* extension of the extended mind theory. I disagree with this claim, but I will not argue for the necessity of personal extension given EMT here. Either the self is extended (a thesis that I find to be most plausible given EMT), or it is not. If the self is not extended, we are left with the agency problem, and perhaps no coherent notion of what the self really is. I voiced problems with individualized theories of the agent in the last chapter. Conversely, if the self is extended, we end up with the problems addressed and explicated in this chapter. So, whether or not the extended self thesis is a *necessary* upshot of EMT, the theory will face inherent problems.

[186] Clark, Andy (2007b). Reinventing Ourselves: The Plasticity of Embodiment, Sensing, and Mind. *The Journal of Philosophy and Medicine* 32, 3: 263-282. 277.

embodied agents can re-negotiate the agent-world boundary in order to include parts of the world in the total agent-world circuit that constitutes them. The agent becomes more than just her body in these cases. She is not *part of* a new systemic whole; she *is* this new systemic whole. Our bodies are therefore extended and no longer limited by the boundaries of the skin or our biological, organismic components. These extended selves are subject to growings, shrinkings, and perhaps even constitutional changes at a distance.

The extended self thesis therefore postulates that the self, itself, extends out into the world along with its cognitive states. The self grows and shrinks with the cognitive operations of the mind, which extend out into the world as couplings with various externalities, however long term or temporary these couplings may be. Further, these cognitive operations of the mind help to make me the cognitive agent that I am.[187] And, the natural conclusion to draw from this is that these cognitive operations are a part of me, such that when they extend out into the world I, too, extend out into the world. When non-biological tools and structures become sufficiently well integrated into an agent's cognitive system, they count as proper parts of a new whole, which is the extended self.

In her paper, "Overextending the Mind?", Brie Gertler points to an interesting problem for the extended mind, which she calls the proliferation of actions. She then attempts to diagnose which part of Clark and Chalmers' version of the extended mind theory is mistaken. What she concludes is that, contra C&C, standing beliefs are not part of the mind. Therefore, any behaviors caused by Otto's external standing beliefs are not Otto's actions. To stop the proliferation of actions we must stop the proliferation of beliefs, and therefore deny that external standing beliefs are part of the mind. And, in order to render this plausible, we must deny that *all*

[187] Clark (2007b), 279.

standing beliefs are part of the mind. This is but one problem for the extended self thesis, and I think the resolution that Gertler provides, while a tempting one, goes too far. Rather than shrinking our cognitive system by denying that standing beliefs are part of the mind, we ought to deny that the self extends beyond the organic boundaries of the person.

The extended self thesis, while initially proposed for various reasons, shows some *prima facie* potential for solving the agency problem for the extended mind. It also seems like a natural, and even necessary, extension of the theory. However, it leads to implausible results when fully examined. In the last chapter I argued that the extended mind theory is incompatible with a thesis that the self is narrow. In this chapter I will argue that the most plausible alternative, the extended self thesis, is also irreparably flawed. The only remaining option for the extended mind theorist is to deny the existence of the self entirely. I will spend the following chapter rejecting such an option. The conclusion of these arguments is that extended mind theory leaves us with no plausible metaphysics of the self, agents, or intentional action.

§4.2: The Problem of Action Proliferation

Clark and Chalmers put forth in their original paper what might be called the standard extended mind view. Their strategy is one of extension. We start by acknowledging the roles that our cognitive processes and mental states play. Then, intuitively, we can begin extending the mind out into the world if there are things out there in the world that seem to closely resemble the mind itself in their functional roles and overall engagement with the human person. C&C examine processes that mirror or are somehow functionally equivalent to the cognitive processes that take place inside of the head. These, they argue, are cognitive in their own right. The same

holds for mental states which have functional roles that are very close to or the same as the roles of mental states (in particular standing, nonoccurrent beliefs) inside of the head.

Clark and Chalmers state that a coupled system with two-way interaction is a cognitive system in its own right. External features play an active role in cognition, as well as a causal and explanatory role. The main burden of proof for the extended mind theorist has been to show that some external items are at least partially constitutive of a cognitive system, rather than being mere external causally efficacious or causally relevant instruments used by a cognitive agent. One core argument used to show this asserts that there are parts of the world that are functionally equivalent to admitted parts of the mind. These items have the same interaction with recognized parts of the mind as those parts of the mind have with other recognized mental entities. Therefore, they ought to be considered to be parts of the mind as well. Clark and Chalmers state that if a part of the world functions as a process which, if it happened to be done in the head, we would count as part of a cognitive process, then that part of the world is part of the cognitive process. This includes mental states, such as beliefs, as well. These mental states are constituted by parts of the environment. C&C argue that what makes something count as a (standing) belief is the role it plays. External information (sometimes) plays the same role as an ordinary standing belief. Thus, since standing beliefs are part of the mind, so are some instances of information recorded outside of the body and brain. Here is Brie Gertler's reconstruction of this argument:

(1) "What makes some information count as a [standing] belief is the role it plays."
(2) "The information in the notebook functions just like [that is, it plays the same role as] the information constituting an ordinary non-occurrent belief."
(3) The information in Otto's notebook counts as standing beliefs. (from (1) and (2))
(4) Otto's standing beliefs are part of his mind.
(5) The information in Otto's notebook is part of Otto's mind. (from (3) and (4))
(6) Otto's notebook belongs to the world external to Otto's skin, i.e., the 'external' world.
(7) The mind extends into the world. (from (5) and (6))[188]

[188] Gertler, Brie (2007). Overextending the Mind? In B. Gertler and L. Shapiro (eds.), *Arguing*

Gertler then examines the problems both of the limits of introspection and the proliferation of actions. In examining the latter, she claims that the extended mind theory entails that we are responsible for actions that are not (intuitively) ours. In effect, if we take EMT seriously it seems that we have the ability to perform *too many* actions. She uses a thought experiment to illustrate this problem. First, we are supposed to imagine Otto, the notebook-using-maven, but this time to imagine that he uses an external computing device as a repository for important information. He records not only his beliefs, but also his desires in this device. Otto, for instance, records his desire to bake banana bread on Tuesday. He also records the beliefs that banana bread requires bananas, that the corner store has bananas for sale, etc. Moreover, he allows this device to perform some cognitive operations for him, such as devising an action plan based on the information provided. This is similar, Gertler says, to how a GPS devises a plan to get you from point A to point B given your current location and your goal location.

Now, we are asked to imagine that his computing device can be plugged into a humanoid robot that Otto also owns. This computing device serves as part of the robot 'brain' (along, perhaps, with Otto's organic brain given that it is part of the robot's cognitive system). This robot then uses the information from the computing device to go about baking banana bread on Tuesday, all while Otto (or, Otto's organic body) sleeps.

> It uses inputs from the robot's various detection systems to determine the layout of its environment, and it controls the robot's movements by sending signals to the robot's 'limbs'. Otto spends Monday asleep in bed. (Or, rather, the organic portion of his body does)... The robot is, however, very active: using the information stored within it, it 'realizes' that a trip to the grocery store is in order, since this is the most efficient way to execute the desire to make banana bread on Tuesday. Drawing on various other bits of

---

about the Mind (192-206). New York: Routledge.193.

> information, it goes to the grocery store, purchases bananas, and returns home. Alas, the organism's sleep is very deep, and he (it?) does not awake until late on Tuesday. When he does, he is roused by the tantalizing scent of freshly-baked banana bread.[189]

Now we are faced with a question: Did *Otto* make the bread? It seems that the extended mind theorist should say that he did, at least if we agree with C&C. After all, his cognitive system performed the actions that led to the bread's being made. The problem here is that Otto's mental states are extended such that he is responsible for too many actions, or the locus of control for too many actions is somehow owned by him. By *being* an extended cognitive system, Otto is then responsible for all of the actions of that system; they are his actions.

If this is correct, Gertler goes on to say, there seems to be no limit to the actions that one single agent can perform. Otto could program an enormous fleet of robots, linking them so that they are in constant communication. These robots could then engage in widespread, multifarious activities. "Some take a slow boat to China; others descend on a neighborhood in Texas and ring all the doorbells at once; others compete in karaoke contests in Tokyo." And, we must say that Otto is actually performing each and every one of these actions, not merely causally contributing to them in some way or other.[190]

Gertler states that we might question whether this behavior truly qualifies as action, but that there is no way to disqualify these as actions without conflicting with C&C's claim that there is nothing special, having to do with a subject's agency, about states internal to her organic body. So, we cannot discount the causal contributions of external vehicles and states to our behaviors.

The proliferation of actions, according to Gertler, is the most pressing problem for Clark and Chalmers. However, I do not think this problem is adequately explained by this thought experiment alone. If Otto is to be responsible for these actions, it must be because he himself is

[189] Gertler (2007), 196-197.
[190] Ibid., 197.

extended to include all of his mentally states. After all, how could there be mental states that are Otto's, but that exist where Otto is not? Otto himself must extend to the boundaries of his mental states in order for the actions that they are generating to be Otto's actions. These are part of Otto's cognitive identity and therefore part of his self.

There is therefore something more precise that needs to be said about Otto and the robot. For, there can be no problem of Otto sleeping while the robot makes bread. Rather, Otto is sleeping, and Otto is making the bread – at one and the same time. *Extended-Otto*, as we might call him, comprises and encompasses all of those entities that are seemingly acting on his behalf. While we might ordinarily think of these actions as being mutually exclusive, Extended-Otto is enhanced in ways that normal body-bound organisms are not.

Gertler suggests that this itself is problematic. However, instead of action proliferation itself being a problem, I believe it is the extended self thesis that leads to it that is truly problematic. Further, the standard extended mind view leads to this thesis and may even necessitate it. The proliferation of actions brings to light a problem with EMT itself. The extension of our mental states leads to the extension of our selves. Clark and Chalmers are right insofar as they assert this.[191] The problem is that this view of the self is problematic enough that it clearly ought to be rejected. If the extended mind leads us directly to a view of the self that is extended, then there is something wrong with the extended mind view, as it then fundamentally supports these implausible conclusions. I address problems with the extended self thesis below, and evaluate whether these can be avoided by extended mind theorists.

[191] Clark (2003) also gives reasons for which the extended mind thesis and the extended self thesis are theoretically inseparable. See Clark, Andy (2003). *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. New York: Oxford University Press.

§4.3: The Extended Self

Clark and Chalmers put forth an extended self view in their original paper "The Extended Mind."[192] Clark has focused on developing the view in more detail. Information in the external world, they argue, is central to the identity of an agent, and therefore part of this agent in the sense of being part of her constitution. The clear implication here is that selves do not supervene on, or metaphysically track, biological organisms, and that we are, in fact, non-organisms insofar as we have extended properties that no organism has.[193] Because a psychological being must be wherever its mental states are, psychological beings must therefore be extended. The realizers of these mental states must be part of the agent whose mental states they realize. If this is true, then all of the above actions are Otto's; Otto is simply not bounded by his body. According to this view, the self exists, but is extended beyond the bounds of the biological organism. When the mind is extended, the division between the agent and the rest of the world is extended outwards. This boundary is extended by the objects that are then incorporated into the agent's own self-constitution. This new systemic whole is not only an extended mind; it is also an extended agent.

Clark, in his more recent work, suggests that human minds and bodies are essentially open to episodes of deep and transformative restructuring in which new equipment or technology, whether physical or mental, can become quite literally incorporated into the thinking and acting systems that we identify as our minds and bodies.[194] The whole human being (not merely the mind) is open to extension. The boundaries between brain, body, and environment are negotiable not only for mental states, but also for the self.

---

[192] See Clark and Chalmers (1998).
[193] See Olson, Eric (2011). The Extended Self. *Minds and Machines* 21, 4: 481-495.
[194] Clark (2008b), 30-37.

We are therefore profoundly embodied creatures, who are "highly engineered so as to be able to learn to make maximal problem-simplifying use of an open-ended variety of internal, bodily, or external sources of order."[195] These resources become included in our bodily schema, or our underlying concept of ourselves used in cognitive processing. Clark states that we, as these kinds of agents, are promiscuously body and world exploiting. We are 'Natural-Born Cyborgs,' or "systems continuously re-negotiating their own limits, components, data-stores and interfaces."[196] Minds, agents, and persons are plastic and adaptable. Minds are not locked into brains, and agents are not locked into their organismic bodies. "The human self emerges as a 'soft self' (Clark (2003)), a constantly negotiable collection of resources easily able to straddle and criss-cross the boundaries between biology and artifact."[197] These tools and artifacts literally become part of us, or part of our self-constitution, changing who and what we are.

Clark (2008a) explains that, since we are *cognitively* permeable agents, our minds extend to include various environmental tools and resources. And, "It must be possible, at least in principle, for new nonbiological tools and structures to likewise become sufficiently well integrated into our problem-solving activity as to yield new agent-constituting wholes."[198] This is not only in principle possible; when cognition extends into the world it actually happens that we become profoundly embodied creatures with constantly negotiable boundaries. *We* do not stop at the boundaries of our skin. Given the kinds of thinking creatures that we are, our boundaries and components are constantly and forever negotiable.

[195] Clark (2007b), 277.
[196] Ibid., 278.
[197] Ibid. Also see Clark (2003).
[198] Clark (2008b), 40.

If Otto is extended, then it seems plausible that his agency is extended as well. Clark and Chalmers gesture at such a view in "The Extended Mind."[199] This view of the self seems to be a natural extension of the extended mind theory. Further, it would go a long ways towards solving the agency problem. This is because the agency problem rests on the idea that, when the mind is extended, our reasons states are external to our selves. They are external to the thinking agent. When these reasons states cause our actions, it is something external to the agent herself that causes them. She therefore lacks requisite control over her behavior, and it does not count as intentional action. If the thinking agent extends along with these mental states and cognitive processes, thereby encompassing them, then these states and processes are no longer external to the thinking agent, and she controls her actions rather than her actions being controlled from apart.

There are good reasons in favor of the extended self thesis given the extended mind theory. Information in the external world, Clark and Chalmers assert, is central to the identity of an agent, and therefore part of this agent in sense of being part of his constitution. This information is central to his identity because it is part of his *belief system*. We are left with this kind of reasoning: extended states and processes constitute part of the mind. The agent must be wherever her mind is, as minds require possession by minded beings. There are not simply minds floating freely in the universe. Minds *belong* to cognitive agents.[200] Because a psychological being must be wherever its mental states are, psychological beings must therefore be extended. In a way this seems inherently plausible; if something is part of a subject's identity, it must

[199] See Clark and Chalmers (1998).
[200] This will come up again in the next chapter. For arguments in favor of this, see Rowlands, Mark (2009b). Extended Cognition and the Mark of the Cognitive. *Philosophical Psychology* 22, 1: 1-19.

therefore be part of that subject. If something is part of a subject, it must be located where the subject is located. The subject must extend to include it as one of her proper parts. I take these reasons to be the main ones for which an extended view of the self seems appropriate. And, I do not dispute that, given the outward extension of the mind, the extended self seems like a plausible, and possibly a necessary supplement to an extended mind theory.

§4.4: Problems with the Extended Self

The problem arises, then, that there are several strong independent reasons to reject the extended self view. Because of these problems no such view can help the extended theorist when it comes to agency. And, in effect, such a view does the extended mind theory a sizeable disservice. First, this view entails that living organisms do not, and cannot, have mental states. Second, it entails that living organisms are not agents. Third, and perhaps most importantly, this view entails a proliferation of mental states, and overlapping mental states between persons (what I will at times call cross-referencing problems). And fourth, it leads to Gertler's original problem itself. I will address each of these problems in turn.

First, extendedness entails that organisms are not the bearers of mental states. We (minded thinkers) are not organisms. Eric Olson describes this problem in this manner:

> Otto is not a biological organism. If he is any organism at all, he is the one we might call his body, O. There is no other organism he could be; O is the only one in the story. But Otto could not be O, because O does not get bigger when Otto takes up the notebook or smaller when he stops using it. Even if the notebook becomes a part of Otto, it always belongs to O's surroundings. You can't make a biological organism bigger by getting it to use a notebook as Otto does, or smaller by taking the notebook away. So although the notebook comes to be a part of Otto (on the extended self), it never becomes a part of any organism. It follows that Otto is not an organism. He was not an organism even before he began using the notebook. He then had the property having a notebook as a part at some time–a property that no organism ever has. And nothing can both be an organism and have a property that no organism ever has. Not only is Otto not essentially an organism. He is not even accidentally or temporarily an organism, for nothing that is even

> accidentally or temporarily an organism has the property of having a notebook as a part at some time.[201]

If the extended-self thesis is true, and Otto is not an organism, then no one who has external mental states at any time is *ever* an organism. He has the property having a notebook as a part at some time, and this is a property that no organism ever has, according to Olson. Having external mental states at some time and being an organism are incompatible properties. These comments rest on some heavy metaphysical assumptions about the self. But, at the very least, no creature who ever has external states is *always* an organism or *essentially* an organism. Having external mental states at some time and being an organism *at that time* are incompatible. Organisms are contained by the biological boundaries of the skin. They are self-contained, encapsulated, entities. Minds, however, have the opposite feature. Rather than being encapsulated, they can grow and shrink. They have constantly negotiable boundaries that biological organisms do not. Because of this, human minds and organisms have incompatible properties.

These problems may be less challenging if the self is only *sometimes* extended. Persons could sometimes be organisms, if this is true. However, if the self extends to include all of its mental states, then whenever the mind is extended, the self is extended as well. And, according to Olson, even if the self is sometimes extended it has the property of potential extension, which no organism has. Some might argue that there is no such property as that of potential extension, but even if this is true being an organism is incompatible with having extended mental states at some time.[202]

Similarly, organisms are not thinkers, for if they were we would have a problem of too many thinkers. Both systems *and* individual organisms would be bearers of the same (token)

[201] Olson (2011), 486.
[202] Ibid.

mental states. Though we would expect that advanced organisms would be capable of thought, the opposite is in fact the case. It ought to be the case, one might think, that my organismic body would be capable of thought. However, if this were the case it would cause an ontological problem. For, if my body can think, it ought to be the case that it (this body) has the same psychological states (at least in some cases) that I (the extended self) do. More precisely, the mind possessed by my body ought to be psychologically indistinguishable from the mind possessed by my extended self in at least many ordinary cases. Our mental properties would in these cases be indistinguishable.

> …so that whatever mental property I have, the organism has too, and vice versa...But if I am not an organism and my animal body is something other than me, and that organism thinks just as I do...there are two thinking beings wherever we thought there was just one: a person who is not an organism and an organism that is (presumably) not a person.[203]

Therefore, it must be the case that human organisms do not have any mental life of their own. An organism may constitute the body of a thinking being, but it is not a thinking being itself, nor (metaphysically) can it ever be.

The second problem, that living organisms are not agents, is similar to the above problems given by Olson. It is agents who perform intentional actions. The behaviors of non-agents do not count as intentional, as they are merely behaviors or happenings in the causal history of the world. An agent's intentional actions are caused by her mental states, namely her beliefs and pro-attitudes toward some given result state. If the extended self thesis is true, organisms are not the bearers of mental states, extended cognitive systems are. Olson's arguments (above) give good reasons to think that this is true. Organisms and extended selves are non-identical, and it cannot be that both organisms *and* extended selves are the bearers of these states. Therefore, it is extended systems that perform actions, and not organisms. This comes in

[203] Olson (2011), 487.

part from a kind of exclusion principle: that for any token intentional action, only one agent must cause it.[204] Thus, for any action that is performed by an extended system, it is not performed by an organism. And, because extended systems are the bearers of mental states (as opposed to organisms), every action is performed by an extended system. It is therefore the case that organisms can perform no actions. Though we might expect organisms to be capable of intentional agency, they can only perform (at most), unintentional behaviors – similar to the behaviors of an unconscious being. These behaviors, like the operations of a computer or an ATM machine, do not constitute intentional action.

Third, this view entails a proliferation of mental states and a set of overlapping mental state problems that I will call cross-referencing problems. An example will help to illustrate the proliferation of mental states. Let's say that my iPhone is an extended part of my self. It contains certain beliefs, dispositions, abilities, etc. in conjunction with my brain and perhaps other extended parts of me. Now let's say I lose my iPhone while traveling one day. So, lacking a mobile smartphone I go out and buy a new iPhone (we will call it iPhone* in order to adequately keep track). Once I get this new phone, the magic of the Cloud imports into it all of my information so that it functions exactly like my (currently lost) original iPhone. Now, iPhone* makes up a part of myself. It is part of my cognitive identity in exactly the same way as the old iPhone.

However, it turns out that a man on the New York subway has found my original iPhone. I have never met this man, nor ever been in the same room as him, but all of the sudden he is carrying something which until recently was a vehicle of many of my beliefs and functional

[204] This is barring so-called group actions, which I do not have room to fully discuss here. Even if this statement is more broadly construed to include either one agent or one group, however, it is unlikely that the organism and the extended self could be considered one agential group, as the extended system is sufficient for the action to be performed.

abilities, something that was a part of me. Let's say he then uses my iPhone to perform a cognitive function (like finding a telephone number or performing a math problem). Now the iPhone is a part of his extended cognitive system. It has suddenly switched the cognitive system and the person whom it in part constitutes. All of this, while perhaps a bit counterintuitive, is fine in and of itself.

Let's say this man decides to return my original iPhone to me, and he is successful in doing so (he is also, of course, quite a good Samaritan). Now I have iPhone and iPhone*. I have the same psychological parts of me, but twice over. I have twice as many of the same, identical token mental states. These mental states are identical qualitatively, contentfully, and functionally. All of the sudden my cognitive bubble has expanded;[205] yet I have merely been handed a piece of metal and glass. I have gained no new functional abilities, no new beliefs with novel content, nor any new dispositions, etc. Yet I am cognitively different – I have *more* mental states. This is the problem of proliferation of mental states given the extended self thesis and the extended mind.[206]

What of my self when I am reunited with my original iPhone? If a part of me is now duplicated there is a clear problem. I could do this again and again until I have 100 identical duplicate tokens of the same mental states in my cognitive system. This would lead to an incredible, and problematic, proliferation of mental states. It is problematic because my cognitive bubble can grow seemingly endlessly, but to no real end when it comes to my abilities or

[205] It is possible that this expansion is significant. How significant it is will depend upon how cognitively reliant I am on these pieces of technology.

[206] This is not the same as what is often called the "Cognitive Bloat" objection. The Cognitive Bloat objection states that our concept of the cognitive will become too permissive if we accept EMT. What I am beginning to address here is the (admittedly similar) problem that our personal ontology is too expandable given the extended self thesis. See below for more on Cognitive Bloat as well as Section 5.3.

capacities. The same goes for my self, which can grow exponentially without, in effect, changing in these respects. It can grow without gaining and shrink without seemingly losing anything. These parts, while inert, must continue to make up part of myself despite their lifelessness. They are dead weight, but they cannot be pruned away from my self. They cannot be gotten rid of in favor of a seemingly more meaningful core. For, if a part of me is *not* duplicated in the above case, there is an obvious question: which iPhone is a part of my self, and why that one? The two items are functionally equivalent. I could use either at any time. One has temporarily left my direct possession, but this consideration in and of itself cannot mean that it has left my constitution. I could have instead left it on my desk for an hour and come back to get it, and it would not have left my self then. Likewise, just because the new iPhone* has been with me for a shorter period of time does not mean it is not part of my self yet. There can be no concrete time limitations on self-constitution, for there are no reasons for which something could not become quickly coupled to me and in a very short period of time become part of my self. I see no promising criteria for ruling one of these devices out one as self-constituting without ruling out the other and without risk of a challenge of pure arbitrariness.

The problem of Cognitive Bloat has strong connections to the proliferation of mental states and the extended self. This problem alleges that the extended view is counterintuitive because its position on what counts as cognitive is excessively liberal. Without certain conditions that demarcate cognitive states, mobile access to Google could constitute part of my cognitive system. A man with access to a telephone directory will "know" all of the numbers in the directory. He will have true beliefs about all of the phone numbers listed therein.[207] A person living in a library knows (at least in a generic sense) everything in all of the books within the

[207] Rupert, Robert (2004). Challenges to the Hypothesis of Extended Cognition. *Journal of Philosophy* 101, 8: 389-428. 402.

library.[208] Clark acknowledges this issue.[209] Rupert (2004) and Allen-Hermanson (2013) examine it at some length. The Cognitive Bloat objection, though, has implications for the extended self thesis, as any states that constitute part of my cognitive system will constitute part of my self. This is a slightly different kind of mental state proliferation, and it may be even more threatening. Further, there is reason to believe that the conditions for cognitive status and therefore person-constituting ability that Clark and Chalmers offer are deeply flawed, and consequently that they cannot function to reign in the extension of our selves.

Clark and Chalmers (1998) put forth some conditions for extended states to be considered part of an agent and/or to qualify as her mental states. In discussing Otto's notebook, they state the following:

> First, the notebook is a constant in Otto's life – in cases where the information in the notebook would be relevant, he will rarely take action without consulting it. Second, the information in the notebook is directly available without difficulty. Third, upon retrieving information from the notebook he automatically endorses it. Fourth, the information in the notebook has been consciously endorsed at some point in the past, and indeed is there as a consequence of this endorsement.[210]

However, there are some clear problems with these standards. Most importantly, we cannot set standards for extended mental states and cognitive processes that aren't met by internal ones.

Sprevak (2009) argues that the parity principle, what he calls the "fair-treatment" principle, ought to guarantee equal treatment between internal and external cases.[211] The parity principle states that "if an extended process is relevantly similar to an internal cognitive processes (save for having external parts), then that extended process should have an equal claim

[208] Rupert (2004), 421.
[209] See Clark (1997), 217 and Clark (2008b), 80.
[210] Clark and Chalmers (1998), 17.
[211] This is the same as the Parity Thesis referred to in Chapter 1. See Section 1.3.

to be cognitive."[212] Extended processes should not have to meet a higher standard for being considered cognitive merely because they are outside of the head. As stated by Clark and Chalmers (1998):

> 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* (so we claim) part of the cognitive process.[213]

Three of the standards that Clark and Chalmers set, and which Sprevak sets out to dispute, are these: that the external resource be reliably available and typically invoked, that any information retrieved from it be automatically endorsed, and that the information in the external resource be easily accessible as and when required.[214]

Starting with the first condition, Sprevak states that it is not the case that a resource needs to be reliably available or typically invoked in order to be cognitive.

> One could imagine a Martian with internal cognitive resources that are neither reliably available nor typically invoked. The Martian might have cognitive resources that are only available after it gets a good night's sleep, and it does not reliably or often get a good night's sleep. However, that does not stop, on those occasions when the Martian does get a good night's sleep, those resources from counting as genuinely cognitive.[215]

Sprevak, here, is addressing what is called the *Martian intuition*. The Martian intuition claims that it is possible for creatures with a very different physical and biological makeup than ourselves to have mental states.

> An intelligent organism might contain green slime instead of neurons; it might be made out of silicon rather than carbon; it might have different kinds of connections in its "nervous" system. There seems to be no reason why mentality has to involve blood, neural tissue, or DNA.[216]

---

[212] Sprevak, Mark (2009). Extended Cognition and Functionalism. *Journal of Philosophy* 106, 9: 503-527. 505.
[213] Clark and Chalmers (1998), 8.
[214] Sprevak (2009), 514.
[215] Ibid.
[216] Ibid., 507.

The simple fact that the Martian has mental states that are stored differently than ours does not mean that it lacks genuine mental states and processes.

Inherent in preserving the Martian intuition is functionalism. "Functionalism preserves the Martian intuition by claiming that what makes an organism have a mental state is the organism's *functional organization.*"[217] This functional organization is typically understood in terms of the state's causal role or its typical pattern of causes and effects. There is no reason why many states, with their associated typical patterns of causes and effects, could not be realized in a Martian with, say, a silicon-based physiology instead of a human with carbon-based physiology.

Going back to the first condition for an external state to be considered cognitive, Sprevak gives a human example. Take a case of outstanding human creativity. The cognitive resources in such creative acts are often not reliably available or typically invoked. But, if someone does employ these resources, then they count as being part of her cognitive processing. Sprevak concludes that the fact that a cognitive process is not reliably available or typically invoked does not entail that it isn't a cognitive process. This fact does not exclude a process or state from counting as cognitive or mental.[218]

Further, if internal resources need not be reliably available or typically invoked, then requirements for cognitive status should allow that external resources not be reliably available or typically invoked, either. "If the functional roles of cognitive states and processes are specified broadly enough to allow for *internal* resources not to be reliably available or typically invoked, then they should allow *external* resources not to be reliably available or typically invoked

[217] Sprevak (2009), 509.
[218] Ibid., 514.

either."[219] This is what the fair-treatment (parity) principle tells us. We cannot have special constraint on external resources that we do not apply to internal resources.

With regard to the second condition, that of automatic endorsement, Sprevak states that it "requires that information retrieved from an external cognitive resource be: (i) more-or-less automatically endorsed; (ii) not subject to critical scrutiny; (iii) deemed as trustworthy as that from biological memory."[220] All of these conditions, Sprevak argues, are violated by actual and possible cases of *internal* cognition.[221] Some human cognitive processes clearly violate the first condition here. Imagining, supposing, desiring, for example, are not automatically endorsed. The second condition fails for similar reasons. A cautious Martian (or even a cautious human) might routinely subject the output of its cognitive resources to critical scrutiny, but its cognitive resources are not made noncognitive or nonmental as a result. And the third condition suffers as well; there are actual human internal cognitive resources that contain information that is not deemed trustworthy at all.

The third condition (easy accessibility) suffers as well. Humans have plenty of mental information that cannot be easily and immediately accessed. Sprevak gives some examples:

> The visual system contains information about current eye position that cannot be easily accessed. Conscious beliefs can also be difficult to access. A nervous student might cram information into her cognitive resources before an exam that she is too nervous to recall during the exam, and subsequently forgets, but although the information was never *easily* accessible, that did not stop it from counting as genuinely cognitive while she had it.[222]

There are certainly other cases of human cognition that we could use as examples here. The point remains that this condition suffers just like the others since the condition does not manifest itself internally.

---

[219] Sprevak (2009), 514.
[220] Ibid., 515.
[221] For further refutation of these conditions for cognitive status, see Rupert (2004), 424.
[222] Sprevak (2009), 516.

Further, if a belief requires past conscious endorsement in order to have the status of a belief, and conscious endorsement is an internal process, then this privileges the bounded subject in a sense that seems to go against the spirit of the extended mind theory.[223] Clark and Chalmers themselves state that the past conscious endorsement criterion is arguable.[224] So, it seems clear that this final condition fares no better than the first three as a condition that a state has to pass in order to count as cognitive.

Farkas (2012) briefly touches on a problem that is similarly relevant to the extended self. Farkas asks us to think of a person named Lotte. Lotte always carries an electronic reading device. It is constantly accessible to her. Lotte decides one day to download a 37-volume history of Europe, which includes a quick search function. The source of this manuscript is one that Lotte trusts completely. After this download, if any question about the history of Europe comes up, Lotte consults her device and looks up the answer. According to Clark and Chalmers, Lotte must have very quickly and suddenly increased her knowledge of the history of Europe. While Lotte's knowledge may seem in some ways very much *less* impressive than an actual historian's, perhaps we can overlook this fault in order to proceed to some more serious problems.[225]

Lotte then decides to turn her 'knowledge' acquisition to philosophy. She enrolls in a philosophy program, where she passes her exams with excellent grades. A problem then arises, though. For, Lotte has hired a philosophy consultant who is consistently available to aid her in her philosophical studies.

> When her professor congratulates her on the results, she explains that her success is due to the fact that she hired a consultant, with whom she can confer 24/7 about any

[223] Rupert (2004), 404.
[224] Clark and Chalmers (1998), 17.
[225] Also see Rupert (2004), 402-403 and 421 for similar examples of this problem.

> philosophical issue through a radio device. She passed her exams and prepared her papers by writing down what her consultant said.[226]

Lotte is an intelligent person, and she can comprehend and appreciate the information that her consultant feeds her. She is not a mere automaton in the sense that information is necessarily merely flowing through her with no recognition on her part. However, it is still understandable that the professor turns out to be not very happy with this arrangement. He explains to Lotte that she cannot use *another person's* help to pass her exams. This, after all, is not meant to be a collaborative project.

At this point, Lotte explains that Lotte "didn't have another person's help, because *she herself* has all the requisite skills and learning, in virtue of her relation to her consultant. She has easy access to her consultant, the consultant is reliably available, and she, Lotte, automatically endorses everything her consultant says about philosophy." Because the two are properly coupled together, Lotte's mind extends partly into her consultant's mind. If she answers a question correctly it is because *she* knows the answer. The professor might think that this clever reply is coming from someone else, because Lotte repeats the words of her consultant, who has been listening to the conversation for the whole time. "But this would be the wrong way of looking at things: since Lotte's philosophical views are constituted in a large part by her consultant's mental states, this very reply is a product of her own mind."[227] Her mind and her consultant's mind overlap in an intimate way. Her mental states are constituted by those of her consultant. This is what I will call a cross-referencing problem, as two persons' selves are crossed so as to reference the same cognitive states simultaneously. This is a counterintuitive result for those of us who are inclined to think that selves, as well as their psychological states,

[226] Farkas, Katalin (2012). Two Versions of the Extended Mind Thesis. *Philosophia* 40, 3: 435-447. 445.
[227] Farkas (2012), 445.

are unique and individualized (even if they are not bounded by the body). Lotte's mental states, her beliefs, overlap with that of her consultant. They share the same belief states and some mental processes connected with those states.

The problem here is not just that the consultant's mind is realizing her belief states. The problem is that it is realizing his own as well. They are cognitively and metaphysically overlapping. Lotte's mind extends partly into her consultant's mind. This has a few implications, some more detrimental than others. The first is that minds are no longer private things. I think many (though perhaps not all) of us can live with that. The second is that *selves* are no longer private or individualized. This implication is harder to accept. Lotte and the consultant do no merely share some mental states, they share these parts of their *selves*. This seems to me to be an unfortunate result. By extending our selves it seems as if we lose our subjecthood because we lose any distinction between persons.

There is also a question of who is responsible for Lotte's actions – specifically the action of taking (and, under a courser-grained description, passing) her philosophy exams. This is probably what the professor is really worried about. It is not Lotte alone (unextended), for Lotte could certainly not have done this without her consultant. She would not have had the appropriate beliefs. It is not the consultant, because he would not have had the appropriate desires. So, it must be the temporary fusion that exists, ontologically, when the two are temporarily coupled. But, that fusion isn't anyone. Subjecthood requires more than a temporary mental coupling. Though subjects may change both physically and mentally, there is a sense in which they endure through time in a way that a temporary coupling does not. They have a past, a present, and a future. In this case, there is no subject there at all.

One might be tempted to assert that the taking and passing of these exams is some kind of social or group action. Perhaps, one might say, both Lotte and her consultant performed this action as a group. But, if this is the case, it is still true that neither Lotte nor her consultant took and passed the exams. And, the professor is certainly right to be cross with Lotte and unaccepting of her behavior, as she did not take and pass her exams alone. Lotte, herself, is not doing anything. For, it cannot be the case that Lotte and the group are performing the very same action. Rather, it is only the group that acts. I will further address group actions briefly in the section below. But, overall, I find the idea of group or social action to be insufficient to address this problem. Agents and selves continue to disappear into nebulous groupings, stripping us of our ontological sovereignty as individual persons.

This brings us back, lastly, the Gertler's problem. Though this is not necessarily a problem for the extended mind theory alone, it is a problem for the extended mind in combination with the extended self. As Gertler explicates the problem, she seems to presuppose (at times) that Otto is not an extended self. This is because she refers to Otto the organism as an independent person. However, assuming that Otto is an extended self, the problem rears its ugly head in earnest. The proliferation of actions follows not merely from mental states being extended, but rather from the self being extended. The robot cooking banana bread is Otto, or a part of Otto just like an arm or leg of Otto's organic body. Otto does in fact become extraordinarily active if his self is extended. Otto "is enormously busy (though tireless) and, unlike Superman, he can be in two places at once."[228] The question is not whether Otto is responsible for the robot's actions. He might be responsible in some sense because he programmed it. Rather, the question is whether it is Otto who is actually performing these

[228] Gertler (2007), 197.

actions. There seems to be no limit to the number of actions that Otto could perform. This extraordinarily active self seems more like a fiction than how we actually take ourselves to be.

While it is possible that selves can have inorganic parts, the proliferation of these parts leads to intractable problems when it comes to agential responsibility and the concept of the self, itself. So, either the self does not extend out into the world, and the mind does, or both the self and the mind are individualized. I have examined problems with the preceding view in the previous chapter. The latter view is antithetical to the extended mind theory.

§4.5: Extended Minds and Group Minds

Clark and Chalmers, at the end of their original paper, not only suggest their acceptance of an extension of the self to parallel the extension of mental states, they also suggest that it might be possible to have truly coupled "group minds." Several of the above problems are related to the hypothesis that group minds could exist. Clark and Chalmers (1998) state that they see no reason for which an individual's mental states could not be partly constituted by the states of other thinkers. I have raised some objections to this above as this idea pertains to the self. Their ideas seem to suggest, however, that group minds might be possible. As stated above, this might help to solve the problems associated with the example of Lotte, explained above.

Perhaps, it could be argued, that individuals are not the kinds of things that we should focus on when we talk about mindedness. Maybe coupled systems, or groups themselves, are the minded entities that we ought to be focused on. This is often called the *group mind hypothesis*.

> First, there is the idea that groups have properties, including mental properties, which their individual members do not have, and which are not reducible to the properties of those members. This emergentist view of group properties, together with the further

> assumption that some of these properties are psychological, entails a version of the group mind hypothesis that postulates group psychological traits.[229]

This is different from the view that individuals manifest certain traits when in groups, or in social settings. It is a view about the ontology of minds themselves and their bearers.

This view might be put forth in order to help to ameliorate problems like that of the proliferation of mental states and cross-referencing problems voiced in the preceding section. I do not think, however, that it can solve these problems. And, the group mind hypothesis leads to an even more robust version of the agency problem.

First, it will be useful to see how such a hypothesis might help to solve the above problems. Some of the problems listed above might be labeled 'overlapping' or 'cross-referencing problems'. They are problems specifically because two people seem to have identical (functionally, and token locationally) mental states. The main problem is that it seems as if these persons share their mental states in a very close and counterintuitive way. And, because of this, they share a part of their selves. If these psychological and mental states belong instead to a 'group mind' that is the mind of a group of people, then this ceases to become a problem. The minded individual is no longer the individual person who bears these states (and thus, who must then share a part of themselves if someone else bears them as well), but rather a group who bears them. The self is not the individual self but the group. There is no longer a problem of metaphysically overlapping selves, because the entity that emerges is a self that is made up of multiple persons.

The group mind hypothesis, importantly, will not solve all of the above listed problems. In particular, it will not solve the dualism problems voiced by Olson. Rather, it will only make these worse. Groups, couplings, etc. are nebulous entities. If these are the things which possess

[229] Wilson (2004), 281.

minds and which are the bearers of mental states, then all of these problems still exist. Individuals are not thinkers, thinkers are not organisms, and organisms cannot think. These conclusions remain problematic.

Thus, if the group mind hypothesis is correct, and if groups are the ontological things that possess psychological properties and mental states, then any actions brought about by those mental states will be caused by those groups. However, provided that individuals can still perform intentional behaviors, the cause of these behaviors is not merely external; our behaviors are caused by a completely different ontological entity, that entity being the group that is minded. Perhaps environmental context does not directly bring about these behaviors, but the group, constituted by other individuals, will cause them instead. This is just as bad, if not worse, when it comes to the production and explanation of our actions.

If groups realize the mental states that cause our actions, then these groups cause our behaviors, as well. Further, we are not even the bearers of the mental states that cause our actions. The group is. There are other, independent problems with the group mind hypothesis, as selves seem to completely disappear into these nebulous grouping. I will examine the possibility that agents and selves cease to exist in the next chapter, and I will present substantial reasons to reject it at that point.

§4.6: Intermediate Conclusions

One interesting upshot of the extended self view is that it turns out that the mind is not actually extended at all. At least, it cannot be extended into the world beyond the person whose mind it is. If the self is extended, then the mind is internal after all. Clark and Chalmers, and other extended mind theorists must merely be speaking in terms of convenience when they say

the mind is extended. What they must mean is that the mind is comprised of things not actually in the external world, but rather in what we intuitively, pre-reflectively took to be the external world. For, something could clearly not be at once part of myself and at the same time part of the world that is external to me. So, if the extended self thesis is true, mental state internalism is true after all.

I think that it would be going too far to accept the consequences examined above. For, if we do, we are no longer embodied minds as we once thought, we are not the kind of actors that we take ourselves to be, and we are not individualized selves. Embodiment is a key commitment for many extended mind theorists, but if this kind of dualism holds we cannot be embodied. The mind cannot bear the kind of intimate connection with the body that we once thought necessary. Rather, we are nebulous, free-floating, disembodied, extended mental conglomerates. We are not what we take ourselves to be, and we are not how we most fundamentally describe ourselves. The self continues to disappear, reappearing only as a glimmer of hope or a conceptual chimera. There really are no selves, and no persons, if this is true. I will continue to address this possibility in the forthcoming chapter.

§4.7: Can We Reject the Extended Self?

C&C themselves propose that the self extends beyond the bounds of the body, but what if we reject this extension of their extended mind theory? Can we simply have extended minds without extended selves? As Clark and Chalmers themselves suggest, it is a very natural extension of the standard extended mind theory to extend the self as well. However, perhaps this extension is unnecessary. Perhaps it can be resisted. I think that there are good reasons to think

otherwise, and good reasons for which this view of the self is essentially what we are left with (if we are left with any view of the self at all) given the extended mind theory.

The problem with rejecting such a view of the self is that we again need a principled reason to assert the boundaries of skin and skull as valid markers when it comes to self-hood, but deny them when it comes to mental states. Some authors have attempted to set such boundaries for the self. I examined several such theories in the previous chapter.

The main problem with all of these theories is that any boundary that applies to the self ought also potentially apply to the mind. We are still left grappling with the original impetus for extending the self: that the cognitive agent must be wherever her cognitive states are, especially those that constitute part of her identity as an agent. The trick is to provide boundaries for the agent that only apply to the agent *qua* agent, locating the agent within the bounds of the organism, and not agent *qua* cognitive being. I view this task to be inherently fruitless. Any privileged boundary that applies to a cognitive *agent* will also apply to that agent's cognitive *states*.

Furthermore, even if we find a way to individualize the self, and thereby reject the extended self thesis, we have a different problem, as it appears that Otto, or any extended-minded being, cannot perform actions at all. Agents are such only if they *possess* the causal antecedents for their actions internally and if they guide and control those actions. I will examine the possibility that there is no agency in the next chapter. The conclusion is that in rejecting the extended self, we either reject the extended mind along with it, or we get rid of intentional behavior entirely. For, as examined previously, the agency problem unavoidably arises with any individualized view of the self when the mind is extended.

Gertler, after examining the problem of the proliferation of actions, concludes, mainly though not exclusively from her argument about action proliferation that we ought to deny that standing beliefs are mental states. In doing this, we can deny Clark and Chalmers' claim that Otto's externally scribed 'beliefs' are mental states at all and therefore that mental states are extended. Gertler states that if we are hesitant to describe (per her example) making banana bread as Otto's action, then we have serious grounds for doubt that extended states have the feature that qualifies them as truly mental. They are not truly mental because, though they are functional similar to standing beliefs, ordinary standing beliefs are not mental. This stops the extension of mental states and therefore stops the proliferation of actions.

If the above arguments are correct, however, we need not deny standing beliefs the status of mental statehood simply in order to deny that mental states are extended and therefore to stop the proliferation of actions. Clark and Chalmers are correct insofar as they assert that shrinking the self to a bundle of occurrent states severely threatens its deep psychological continuity. My dispositional beliefs constitute a part of who I am in a deep sense.[230] If mental states are extended and the self is individualized, then there are no genuine actions. If there are actions, then the self must be extended in order to accommodate this. Thus, we can deny that mental states are extended not by denying that (what we ordinarily think are) mental states really are mental, but rather by denying that the self is extended and therefore that our mental states extend further than our organismic bodies. Simply, if the mind is extended, the self is extended beyond the boundaries of a living organism. The self does not extend beyond the bounds of the living organism, and therefore neither do mental states.

[230] Clark and Chalmers (1998), 18.

There is, I think, something sacred about skin and skull when it comes to the bounds of the self. Either the self is extended, or it is narrow. If the self is extended, we are faced with many ensuing implausible implications. There are some reasons to believe that if the mind is extended, then the self must be as well. However, if there is a principled way to place boundaries on the self without those boundaries applying to the mind, we are faced with a remaining problem that these bounded agents cannot perform genuine intentional actions because their actions are caused by components external to them. Neither of these options is plausible. And, as these are natural extensions of the extended mind theory, the hypothesis of extended mind itself is rendered increasingly implausible.

This is more than merely an argument that individuals are an appropriate object of scientific, cognitive scientific, biological, and psychological study. It could be the case that the self is metaphysically extended, but some individualized part of this system is an appropriate object for certain kinds of (broadly) scientific study. This is similar to how the heart and circulatory system is the proper object of study for a cardiologist, or how the nervous system is a proper object of study for a neurologist. These subsystems are appropriate areas of study even though they are parts of greater personal level systems. There are plenty of appropriate objects for scientific study that are not 'selves'. So, perhaps there are individualized, organismic systems that are still proper objects of study – even psychological study. This, however, does not prohibit the extension of the personal self beyond these boundaries.

So, any argument for the individualized self that rests on the claim that individualized systems (organisms, perhaps) are the proper object of scientific study will not, on its own, be successful. It is not a good way to make a case for the importance of privileging neural states, because privileging neural states whilst in the midst of studying them is far different them giving

them a privileged importance ontologically and metaphysically. The above arguments, though, are very separate from such a claim. The claim being made here is not that the self is special because of its explanatory role. The claim is that a lack of individualized self leads to consequences that are at best implausible. In addition, these arguments alone do not hinge on any specific concept on the self. They merely suggest an individualized self, however such a self might be constituted.

The arguments in the next chapter are meant to support the claim that, ontologically, individualized agents exist. And, conceptually, we ought to retain our fundamental concepts of agency being an individualized phenomenon. While there are problems with the extended self thesis, the extended mind theory depends too much on personal subjects and agency for it to give up the self as an entity altogether. Doing so would not only be counterintuitive, but also counterproductive for the theory.

Chapter 5: The No-Self Alternative?

§5.1: Introduction

Previously, I have argued that the extended mind theory cannot accommodate an internalist view of the agent, nor can it justify an extended view of agency. But, if agents do not exist, the problem at hand ceases to be a challenge to the extended mind theory. At this point it seems apt to motivate the view that selves, persons, and agents are important parts of our personal ontology that ought not to be given up in order to preserve the view that the mind is extended. A minority of authors focusing on personal identity has argued that there is no self, or nothing that we can rightly call a self. If there are no selves, and no agents, then the agency problem can ultimately be eliminated. One reply to the agency argument would be to maintain that we ought to give up these metaphysical artifacts. If there are no agents who act, no selves, and no persons, then the argument will not extend far. There are, however, reasons for which it seems that the extended mind theorist cannot give up the agential self.

In this chapter I will address the presupposition of selfhood that seems inherent in many authors' renditions of the extended mind theory. I will also address the deep counterintuitiveness of the no-self position, which I believe shifts the burden of proof (when it comes to personal identity) towards the extended mind theorist should they attempt to give up the self entirely. Further, I will argue that the self that is presupposed by the extended mind theory is essentially agential, and thus the agency problem cannot be avoided. Giving up the agential self would seem to undermine integral parts of the view.

§5.2: The Counterintuitiveness of the No-Self Position

The no-self position is certainly the less intuitive of the two alternatives (self and no-self). When we think of others, we think of them as subjects of experience who, as individual agents, carry out actions in the world. Wilson (2014) states that cognition without a cognizer is deeply counterintuitive.[231] I believe this to be correct. The no-self view, clearly, lacks an intuitive basis. The theory simply does not compute for us as thinkers, and it therefore contradicts our deepest intuitions. One might object to any view that posits the existence of a self, period. What if, we might ask, were we to give all of the defining characteristics of the self, such a thing surely did not exist?

Rupert (2009) states, "we might talk of persisting persons, selves, or systems, but from the standpoint of cognitive science, it is convenient talk only. Really, there is only a distribution of various instantiations of cognitive properties, pertaining, for example, to computation and representation."[232] This is the no-self view. On this view, "cognitive scientific explanation appeals only to the distributed instantiation of cognitive properties: information, computation, dynamic flow, activation patterns, and the like."[233] There is no appeal to a subject who owns these processes.

Some may think that a view that involves the self needs to include a central executive, a subject within the Cartesian Theater, or even a separate substance that is the self. This need not be the case. Hume rejected a substantial version of the self in favor of a view according to which the self is a bundle of perceptions. No empirical evidence, Hume thought, can support a substantialist metaphysics of selfhood. Hume sought in vain to find an unchanging, simple

[231] Wilson (2014), 30.
[232] Rupert, Robert (2009). *Cognitive Systems and the Extended Mind*. New York: Oxford University Press. 51.
[233] Ibid., 52.

substance that bears one's experiences, ideas, and impressions. When we look at our experience there is nothing to substantiate our belief in the self. We are never aware of any constant invariable impression that could answer to the name of self. We simply experience a continuous flow of perceptions in rapid succession. And, since ideas must be derived from impressions, we cannot have any idea of the self.[234] Like many present-day philosophers, Hume saw no reason to believe that any unchanging, simple substance has anything to do with our experiences. So, Hume argued, there are no substantial selves or simple substances that are selves. There is no independent entity, aside from our experiences, that is autonomous with regard to these experiences.

According to Hume, the self is not an existent thing, though it is undeniably experienced by us as such. Personal identity is a construction of the mind, as the human imagination imposes identity on oneself. While we experience our selves as somewhat unified, the self is no more than a bundle of properties, without any substratum or substance to possess those properties. A person is a bundle of perceptions, and nothing else. This set of perceptions has no substance that underlies it. Sensory evidence can give us nothing more. Hume ends up concluding that there is no such thing as the self apart from the stream of experiences that we ordinarily think of the self as owning. The self simply is some series of thoughts, actions, and experiences. The self is no more than a bundle of perceptions and experiences in perpetual change.[235]

One's self is simply one's mental life, a bundle of one's thoughts and experiences. Our evidence for this kind of self is abundant, while our evidence for a substantial self is missing. The content of our perceptual worlds is forever changing, and the self is forever in flux.

---

[234] Giles, James (1993). The No-Self Theory: Hume, Buddhism, and Personal Identity. *Philosophy East and West* 43, 2: 175-200. 177

[235] Santos, Ferdinand and Santiago Sia (2007). *Personal Identity, the Self, and Ethics*. New York: Pelgrave Macmillan. 46.

Perceptions make their appearance and then pass, making way for other perceptions to appear. We are therefore constituted by a set of changing perceptions, and the self is to be identified with a set of these. The identification of the mind with its perceptions is absolute. If there is an idea of the self, it is not a substantial self.

There are many other models of the self that we might entertain. Personal identity is too large a topic to tackle fully in this work. The self as a substantial object, separate from the body and the rest of the natural world, must be rejected in this physicalist context. Different authors give different definitions of the self. I propose that we understand the self as (at least partially) a mental thing,[236] while it may be reducible to or emergent from other things. The self is also a subject of experience, a conscious thinker, an agent, and it can be considered both synchronically and diachronically.[237] According to this definition, the self could be emergent and enactive,[238] it could be dynamical,[239] or it could appear under some other description.

The self could be purely inactive if agency (including mental agency) were not part of this list. The self would just be a continuous flow of experiences, sensations, and psychological and possibly physical states. Surely the existence of a concrete individual does not necessitate that that individual be active. I will address this further later in this chapter.

---

[236] This is only insofar as what is mental can be considered apart from what is physical. On the materialist picture of the world the mental may be considered a mere subset of the category of the physical.

[237] This list is in large part taken from: Strawson, Galen (1999). The Self. In S. Gallagher and J. Shear (eds.), *Models of the Self* (pp. 1-24). Exeter, UK: Imprint Academic. 3.

[238] See Mackenzie, Matthew (2011). Enacting the Self: Buddhist and Enactivist Approaches. In M. Siderits, E. Thompson, and D. Zahavi (eds.), *Self, No Self?: Perspectives from an Analytical, Phenomenological, and Indian Traditions* (pp. 239-273). New York: Oxford University Press.

[239] See Tani, Jun. (1999). An Interpretation of the 'Self' from the Dynamical Systems Perspective: A Constructivist Approach. In S. Gallagher and J. Shear (eds.), *Models of the Self* (pp. 149-176). Exeter: Imprint Academic (1999).

This brings us to the next point, about whether an agentive cognitive subject is necessary for an extended theory of mind to be successful. For reasons outlined below, I believe that it is. Not only does the extended approach rely on talk of the self and subject extensively, the no-self view does it a disservice.

The no-self theory seems to represent an impossible scenario. It cannot be simulated because the cognitive agency that it takes to imagine it would have to be dissolved.[240] Any attempt to mentally simulate the theory generates cognitive agency and therefore selfhood. For, in thinking of the absence of the self, one must assert oneself as a cognitive agent. The Cartesian cogito seems hauntingly familiar here. The necessity of the notion that I must exist in order to think of my own absence makes the notion of selflessness seem incoherent. And, this incoherence creates an unshakable counterintuitive flavor that is inherent in the theory.

We might think that someone denying the existence of the self cannot possibly mean what she is saying. Perhaps she is only denying the existence of a certain type of self, or a certain definition of how the self is constituted. But she cannot possibly be denying the existence of the self itself. The position seems self-defeating in that it seems as if stating the proposition behind it undercuts that very proposition, and thus it cannot be stated without being undermined. The statement that the self does not exist seems to entail that one could not utter that very statement. A self, being necessary for genuine agency, also seems necessary for meaningful utterances of this type.

Counterintuitiveness alone may not be sufficient to ground a rejection of the no-self view. However, I do believe that it shifts the burden of proof towards the extended mind theorist. While the lack of an intuitive basis is not alone enough to defeat a theory of mind, this kind of

[240] Metzinger, Thomas (2011). "The No-Self Alternative". In S. Gallagher (ed.), *The Oxford Handbook of the Self* (pp. 279-296.). Oxford, UK: Oxford University Press. 289.

deep counterintuitiveness present is a mark against it. If the extended mind theory must sacrifice the self in order to lend itself plausibility, it has traded one implausibility for another. On the one hand, if there is no room in the theory for the self or agent, the theory seems impoverished. It not only loses its luster, but also all of the support that action gives to extended cognition, a topic which I will address later in this chapter. On the other hand, if the agential self exists, and the theory is incompatible with the concept of genuine agency, then the theory seems deeply implausible.

Importantly, there would be something very strange about a theory of mind that eliminates selves from our personal ontology. Our ordinary conception of the mind involves a being who is minded, whose actions are informed by the cognitive processing and mental states constituting this mind. Whether or not selves are psychologically constituted or physically constituted, they are the things that have minds, the things that have cognitive processes and the things that have mental states. Minds without minded beings, and therefore without mindedness, would not be what we ordinarily take minds to be. Thoughts and thinking presuppose thinkers, while intentionality presupposes intentional action.

In addition, the extended mind theory ought to be cautious of the conclusion that there is no self for the simple reason that a theory of mind without minded selves is deeply counterintuitive. The aim of the theory is to extend the mind from an egocentric position to a world-involving one, or to motivate the view that minds are not simply egocentric things. If the only way to make such a theory tenable is to reject that there are minded beings, and therefore mindedness itself, then we have started with a theory of the mind and ended up with a theory of something very different. If minds require minded creatures (of some kind or other), then a theory of mind that eliminates these will seem incomprehensible.

§5.3: Presupposition by Extended Mind Theorists

However, there is one other initial problem with the extended mind theorist denying that the self exists. The problem is simply that many extended mind theorists presuppose an existent self who is the subject, controller, or agent of cognitive processes. Extended mind theories clearly presuppose a self that exists and is ontologically prior to the existence of its mental states. These theories of mind refer consistently to persons and agents, which do not exist if there is no self. The extended mind theory itself is a theory about the relationship between cognition, individuals, and environments. It is not simply a theory of environmental cognition. The individual must be part of this make-up. Specifically, though, the extended mind theory has inherent to it the idea that there is a *cognizer* behind cognition. Authors who adopt this model of the mind include a robust conception of the self in their theory.

Wilson and Clark state the following in their explication of the extended theory: "Natural extended cognitive systems are those cognitive systems containing natural resources from the cognizer's environment that have been functionally integrated into the cognitive repertoire of that cognizer."[241] The relevant idea here is that the resources employed by extended cognition are the resources that are utilized by the self, by the person, or by the creature who cognizes. There are natural cognitive resources free for the exploitation of cognizers that bind with their cognitive systems in a coupled system of two-way interaction. Wilson and Clark fully admit that cognitive augmentation is the augmentation of the cognitive system of an individual. Going back to Clark and Chalmers' now-very-popular Otto example, Wilson and Clark state the following of the coupling conditions that they endorse:

[241] Clark and Wilson (2009), 62.

> Such coupling conditions are meant to ensure that the capacities of the hybrid system—the biological organism plus augmentation—are plausibly seen as the capacities of a specific individual (e.g. Otto). For we properly expect our individual agents to be mobile, more-or-less reliable, bundles of stored knowledge and computational, emotional, and inferential capacities, and so we need to be persuaded that the new capacities enabled by the addition of the notebook are likewise sufficiently robust and enduring as to contribute to the persisting cognitive profile we identify as Otto-the-agent.[242]

These coupling conditions establish the reliable presence of the new (extended) cognitive capacities when needed and make sure that the component, being automatic and trusted, is more or less a part of the agent. Here, Clark and Wilson seem to be taking up a reductive view of the self, but it is a view that posits a self nonetheless. There is a specific individual, like Otto with his notebook, being pointed to with a persisting cognitive profile, with certain capacities that exist not in a vacuum but for the usage of that individual. Wilson and Clark also state that extended cognitive systems are those systems where natural resources have been functionally integrated into the cognitive repertoire of a cognizer.[243] These are natural resources for a creature, they augment the system of cognition done by that creature, and they augment the capacities that that creature possesses.

While Clark rejects the idea of a central executive that plans and executes actions, he does not reject the idea of agents straightforwardly. Ultimately, all there is is cognitive circuitry and processing all the way down. Persons just are these various cognitive systems and subsystems, internal and external.[244] Perhaps a separate and permanent self cannot be found in these subsystems and agent-world circuits, but that does not mean that there is no self to be found at all.

---

[242] Clark and Wilson (2009), 67.
[243] Ibid., 62.
[244] DeMoss, David (2011). Empty and Extended Craving: An Application of the Extended Mind Thesis to the Four Noble Truths. *Contemporary Buddhism* 12, 2: 309-325. 312.

Hurley (2001, and 1998) talks about the personal level and sub-personal level in her writing. The interesting thing about introducing any such talk into a theory of mind is that it is a prerequisite for that theory that there be a person to whom that levels apply. Hurley likewise talks about a rational agent.[245] In particular, she speaks of a self-world distinction even given extended cognitive functioning. These two distinctions require a self, or a person, in order to substantiate them. If there is no self, then there is no self-world distinction whatsoever. If there is no person, then there are no personal or sub-personal levels.

Rowlands (2009b) additionally speaks of a personal/sub-personal level distinction. The distinction, Rowlands says, roughly corresponds to the distinction between processes that make information available to the agent and those that do not, or at least do not do so without subsequent personal level processing.

> Processes that make information available to a subject are personal level processes, and this is true whether they also make information available to subsequent processing operations. However, processes that make information available only to subsequent processing operations are sub-personal cognitive processes.[246]

Thus, processes that make information available to a subject are personal level processes. Processes that do so only after subsequent processing are sub-personal. But, it is the availability of information to the subject that separates these two levels. Personal and sub-personal level processes therefore only make sense in virtue of a subject that substantiates their distinction. Personal-level processes make information available to the subject, while sub-personal processes do so only with intermediary processing.

Further, Rowlands explicitly pronounces that there cannot be cognition without a cognizer. He agrees (2009b) that there can be no cognitive processes that are not owned by a

[245] Hurley (2001a), 423-431.
[246] Rowlands (2009b), 9.

subject. Further, he states specifically that a process only counts as cognitive if it belongs to a subject. He argues, "anything that is to count as a cognitive process must belong to some or other representational subject…There are no un-owned cognitive processes."[247] There can be no subject-less cognitive processes. There is no thought without a thinker. And, every cognitive process must have an individual owner. Whatever the form and function of cognitive processing, this processing always belongs to someone.[248] In fact, the purpose of cognition simply is to make previously unavailable information available to the subject. Cognition cannot exist in a vacuum, for it would neither fulfill its purpose nor its definition.

Rowlands thus explicitly acknowledges that cognition requires ownership, and therefore a subject or owner of cognitive states. His mark of the cognitive even includes a subject who owns cognitive processes. If cognition requires a subject – i.e. if there are no subjectless cognitive processes – then extended cognition will require a subject from which it is extended.

Rowlands (2009b) states four conditions for a process to be considered cognitive. First, the process must involve information processing, or the manipulation or transformation of information-bearing structures. Second, this information processing has as its proper function making available either to the subject or to subsequent processing operations information that was unavailable prior to this processing. Third, the information must be made available by way of the production, in a subject, of a representational state. And, fourth, the process must belong to the subject of that representational state.[249]

Here, I am most interested in the fourth condition for a process to be considered cognitive, though conditions two and three are also of interest in light of talk of a cognitive

[247] Rowlands (2009b), 10.
[248] Ibid., 15.
[249] Ibid., 8 (emphasis omitted).

subject or agent. Rowlands states that he means to construe the idea of a cognitive subject broadly.[250] Cognitive processes must be owned by a representational subject (broadly construed). The fourth condition is meant to rule out cases of cognitive bloat- or cases in which distant information processing would otherwise count as cognitive according to conditions 1-3. The processes within my computer, for instance, satisfy conditions 1-3, but intuitively they are not my cognitive processes. The intuition is "that there are no *subjectless* cognitive processes: they always have an *owner*, and this owner is an *individual* of some form." And, the extended mind theory needs to explain the appropriate way in which a cognitive subject can own informational processes.[251] Rowlands states "whatever else is true of cognitive processes, whatever the specific details of their form and function, such processes always belong to someone or something." This is why condition (4) is necessary.[252] From this we can conclude that cognitive processes have a subject, though what kind of subject they are owned by is a topic that I will address in more depth later on.

Rowlands argues that explaining the ownership condition is just as much of a problem for the internalist as it is for the extended mind theorist. Ownership is not determined by spatial position (inside vs. outside of the body) but rather by a certain kind of integration. The criterion for appropriate integration will be determined by the proper function of the process. One process is appropriately integrated into another when it is fulfilling its proper function with respect to that process.[253] "Ownership is to be understood in terms of the appropriate sort of integration into the life—and in particular, the psychological life—of a subject." However, in the case of cognitive processes, specification of the appropriate sort of integration will be complicated by

[250] Rowlands (2009b), 10.
[251] Ibid., 15.
[252] Ibid.
[253] Ibid., 16-17.

the possibility of a distinction between personal and sub-personal cognitive processes. Rowlands states that it is his suspicion that ownership of our sub-personal cognitive processes will be derivative upon ownership of personal level cognitive processes. He states that "for cognitive processes, integration is ultimately integration into the conscious life of a subject."[254] Personal level cognitive processes that are available for conscious processing can therefore be integrated into this conscious life, while sub-personal processes that are unconscious are integrated insofar as they are appropriately joined with those conscious processes. The mark of the cognitive therefore includes in one of its conditions (condition four) the idea that a cognitive process must be integrated into the conscious life of a subject, whether directly or derivatively.

Further, whether or not we want to spell out ownership and integration in this particular manner, explaining the relevant notion of integration, Rowlands claims, should be no more problematic for the extended mind theorist than for the traditional internalist.

> That is, there is no reason for supposing that a subject's use of external resources—its manipulation, transformation, and exploitation of external information bearing-structures—will be any more difficult to integrate into that subject's overall psychological economy than will be cognitive processes in the traditional internal sense.[255]

Though there are those who may dispute this claim, I will not take issue with it here. The important point is that however we spell out the notion of integration, and however difficult it may be to do so, ownership requires not only integration itself but integration into the cognitive economy of a cognitive agent and integration, whether directly or derivatively, into the conscious life of a subject.

So, ownership requires integration, because a process cannot be owned unless it is properly integrated into the cognitive economy of a cognitive subject. Integration requires not

[254] Rowlands (2009b), 17.
[255] Ibid.

merely integration with other cognitive processes, but integration into the conscious life of a subject herself. Thus, condition four of Rowlands' mark of the cognitive requires a cognizing subject, not merely cognitive integration itself.

When it comes to condition two, Rowlands stresses that the information must be made available to a subject, and not to subsequent processing operations.[256] Cognition makes information available *to a subject* – not merely to other cognitive processes. It is not enough for a process to be cognitive that it be properly integrated with other cognitive processes within a cognitive system. In order for a process to so much as count as cognitive, there must be a subject present to whom cognition makes information available.

One might disagree with Rowlands and state that a process can be cognitive if it is integrated with other cognitive processes within a system. Rowlands argues, or at least stipulates, that in order for a process to count as cognitive, there must be a subject to whom cognition makes information available. What if an extended theorist trying to take up a no-self position denies this claim? Can there be subjectless, ownerless, cognitive processes?

Usually, we have a sense of ownership when it comes to our thoughts. "The idea of mental processes belonging to no subject is eccentric, but not unheard of."[257] This is not sufficient to show that cognition only counts as such when it is owned by a subject. However, the extended mind theorist may need an ownership condition in order to avoid the Cognitive Bloat objection. If it is enough for a process to be cognitive that it is integrated with other processes within a cognitive system, cognitive bloat seems to ensue. The only conditions that Rowlands gives for integration are that there be a meshing of disparate types of process that, "precisely

[256] Rowlands (2009b), 9.
[257] Allen-Hermanson, Sean (2013). Superdupersizing the Mind: Extended Cognition and the Persistence of Cognitive Bloat. *Philosophical Studies* 164, 3: 791-806. 795.

because of their disparate character can enable a cognizing organism to accomplish tasks that it would not be able to achieve by way of either type of process alone."[258] When it comes to cognitive processes, integration is ultimately integration into a subject's conscious life.[259]

The Cognitive Bloat objection is relatively simple. In essence, it states that in extending the boundaries of the mind, the extended mind theorist gives an overly permissive demarcation of the cognitive, which leads to the *overextending* of these boundaries. We do not want to admit into the category of the cognitive all kinds of processes that are clearly not cognitive processes. This problem arises because, without seemingly arbitrary caveats to EMT, like a clause that conveniently excludes specific resources, many people would have minds that, due to an efficient and consistent ability to access things like the Internet and other electronic resources, consist of an extraordinarily large body of knowledge. People connected to hi-speed Internet devices (like my iPhone, for example) would have a mind that encompassed large parts of the Internet. I, with my iPhone, have no less ability to access an encyclopedia online than Otto would to access his notebook. A theory that asserts that our minds really are that large seems to have gotten it wrong. But, further, there may be other resources that contribute to cognition and our cognitive functioning that would end up being part of our minds, like the sunlight that allows me to see my keyboard and write this right now.[260]

If ownership is not part of the mark of the mental, the threat of cognitive bloat looms large.[261] I do not want to dwell on the general problem of cognitive bloat here, though I do find it

[258] Rowlands (2009b), 3.
[259] Ibid., 17.
[260] For more on this, see Ludwig, David (2015). Extended Cognition and the Explosion of Knowledge. *Philosophical Psychology* 28, 3: 355-368.
[261] See Allen-Hermanson (2013) for more on this. Arguably, even with the ownership condition this problem will still threaten the extended mind.

to be dampening for the extended mind thesis. I merely want to mention that without an ownership condition the problem of cognitive bloat because even more demanding.

Clark admits that belief states are not just beliefs in a vacuum, or even in a system. They are someone's beliefs. They form part of the psychology of a subject. The notebook does not merely count as part of *a* mind, it counts as part of *Otto's* mind. The notebook is Otto's notebook, and it is part of Otto's cognitive processing. It is not part of cognitive processing *simpliciter*. It is required, if mental states or cognitive states are states of a person, that there is a person or self whose states they are.

The theory of extended cognition, Clark acknowledges, is at center an attempt to best model not only minds, but agents and persons.[262] The extended mind theory is not only a theory of the mind, but also a theory about who and what we are as persons. Getting rid of selves and persons makes the theory itself an empty one as it fails to explain the roots of mindedness and fails to explain our personal aspects. Clark himself states that the theory is a vision of our own humanity and of the fundamental nature of the human self.[263] A theory about the fundamental nature of the human self cannot eliminate selves from our personal ontology. If the extended mind theory is a theory of mindedness and mindfulness, this requires as a prerequisite minded selves.

The extended mind theory involves discovering the basis for mindfulness – whether it is internal or external. It is an attempt to "grasp the roots of mindfulness."[264] And, mindfulness requires a minded being. Clark does not wish to dispense with minded beings entirely. His

[262] Clark (2007b), 277.
[263] Ibid., 278.
[264] Clark (1999b), 5.

project, rather, is to reinterpret what the limits of the self are. This changing image does not entail that the self disappears. Rather, the self is still at the core of the extended mind thesis.

Even if the extended mind thesis attempts to shift the boundaries of the observer and the world, or the self and the other, it remains the case that there is an observer, a thinker, a perceiver, and an actor that either supervenes on, is reducible to, or is something over an above an extended mind. For those, like Clark, who think that the self extends, it is obvious that there is a self. But, given that we must deny this feature of the extended mind theory, we can only deny that the self extends. We cannot deny that there is any self at all. Such a view would be questionable with any theory of the mind, and it is inconsistent with a fully explicated extended view.

Chemero and Silberstein (2011) also focus on actions and agency. They state that cognitive systems engage in purposeful actions, and even that the dynamic activity of an extended phenomenological-cognitive system *is* purposeful action.[265] But, there cannot be purposeful action without an agent behind that purposeful action. In fact, Chemero and Silberstein seem to take up a robust view of agency when they describe their view as one that does not reduce agents simply to a sequence of decision making conscious states.[266] They state that the agent must be an identifiable individual, even on their extended phenomenological-cognitive theory.[267] Agents are prerequisites for intentional actions because it is their intentions that cause actions.

These distinctions and requirements are meant to apply to real metaphysical entities – the person and the self. They are used to answer specific questions about the mind and its

[265] Chemero and Silberstein (2011), 10.
[266] Ibid., 11.
[267] Ibid., 13.

enhancements and restrictions. By abandoning a minded subject, these theories would also be abandoning many of their core principles.

§5.4: What Kind of Subject Does the Extended Mind Theorist Require?

Even if EMT requires a subject, what kind of subject does it require? If EMT requires a bare, non-agential, subject, the problem at hand may slip away. Ordinarily, we conceive of the self as an agent, but agency of the kind I have been talking about (it might be argued) need not be essential to our concept of the self. If the cognitive subject exists, but is passive when it comes to intentional action in the world, then the agency problem will not apply for the extended mind theorist. The extended mind theorist could admit that the self exists, and thus avoid the deep counterintuitive implications addressed above, but deny that the self is agential, thus side-stepping the problem that the self cannot act if the extended mind theory is true. The extended mind theorist could *admit* that the self is not agential, but still hold that the self exists, and therefore avoid the problem. A deflationary view of the self will keep the agency objection from obtaining purchase. Even if minds are owned by subjects, one could presuppose a vision of the self that does not ground agency. However, there are reasons to think that the subject that the extended mind theorist presupposes is essentially agential. I will explore these reasons in this section.

Clark (1997) states that real embodied intelligence is fundamentally a means of engaging with the world. We use active strategies that leave much of the information that we need for cognitive processing out in the world, and we then astutely use interactions between body and world in order to solve problems in a robust and flexible way. The agent and the environment are

two complex systems coupled together in a joint problem-solving venture.[268] Embodied intelligence is fundamentally connected to active engagement with the world, or action. Body-world interactions would be impossible without action. The agent, while coupled to her environment, is a necessary part of this pair. And action, at a fundamental level, allows for the kind of problem solving that is essential to this kind of intelligence. Clark speaks extensively about autonomous agents, or creatures capable of survival, action, and motion in real time in a complex environment. Clark also focuses on how cognition exploits real-world action so as to reduce one's computational load. He is skeptical of action-neutral models of cognition generally. Cognition without action would be impoverished if not impossible.

Clark does not seem to be discussing bare selves here who can exist without agency. Rather, he seems to be discussing real, ontologically significant *agents*. The interaction of the brain, body, and environment is not simply for the creation of thought, or information processing. Rather, the creation of action is a fundamental element. All of the components in a coupled system need to play an active causal role, and they need to jointly govern behavior. The subject of cognition therefore must exert causal influence within the system in order to jointly govern behavior (along with external environmental entities).

It seems to be an essential part of extension that a subject engages with the world. Menary (2007) states, "it is through our bodies that we primarily engage with the world and through this engagement the body is constantly integrating with the environment."[269] Bodily action is essential to cognitive integration. Cognitive processes are in part constituted by bodily

[268] Clark (1997), 98.
[269] Menary, Richard (2007). *Cognitive Integration: Mind and Cognition Unbounded*. New York: Pelgrave Macmillan. 77-78.

manipulation of structures in the environment.[270] Continuous reciprocal coupling between the environment, body, and brain is necessary for cognitive extension, and active engagement of the subject is necessary for reciprocal coupling to take place.

Clark (2008a) states that we are "cognitively permeable agents." Profoundly embodied creatures are "forever testing and exploring the possibilities for incorporating new resources and structures deep into their embodied acting and problem-solving regimes."[271] Human minds are the minds of *profoundly embodied agents*. Simply put, it seems that there can be no extension without action. This implies that action is fundamentally tied to the project of extended cognition. The self, it appears, must be active in order to accommodate these claims.

Clark's continual talk of agents and agency, however, does not by itself entail that agency is an essential part of the cognitive subject. It might be possible for a subject to be essentially inactive while cognition is essentially active. The problem with this is not simply that cognition is separated from subject. The problem is that it is hard to envision a subject that doesn't have mental or cognitive capacities. What would constitute such a subject? The subject may be simply a bodily vessel for cognition, in this case, but this physical self would be an empty vessel. In this case the mind and the self become radically disconnected.

Cognitive states would no longer be important to the identity of an individual, if this view is put forth. This view is somewhat implausible for the reason that we tend to identify at least part of the self with our mental states and cognitive processes. If the self were essentially inactive, cognition would be something the individual passively has as one of its inessential properties. Cognition would not be something that lends itself to our identity as subjects. The body and the self would be a mere passive puppet for cognition.

---

[270] Menary (2007), 83.
[271] Clark (2008b), 42.

While Eric Olson[272] holds an animalist view of personal identity that does not make reference to psychology, the extended self theorist cannot rely on such a view. This view states that we are animals, and that no kind of psychological continuity is necessary for us to exist. Psychological facts would be completely irrelevant for determining personal identity. This undermines the idea that the mental can be part of our identity as persons. Adopting animalism, further, might pose a problem for the view that the self is essentially passive. According to this view, you are a body, or a human animal, you are not a corpse. You are a living organism. And that living organism is not essentially passive. Something that is alive in the biological sense cannot be essentially passive. "The stability of a living thing…is dynamic."[273] The living thing only maintains itself through engagement in constant activity. A living organism has a life, which is a complex self-sustaining process. It is essentially dynamic and not essentially passive.

The question is, is intentional activity (necessary for extension) an essential or integral *part* of the cognitive subject? I stated above that it appears that the cognitive subject must be active, but must it be essentially active? I think that the extended mind theorist must say that it is. The very concept of a cognitive subject described by extended mind theorists seems to include agency. In brief, agency must be part of the cognitive subject because action is part of cognition, and taking action out of the picture would be akin to taking cognition away from the cognizer.

Action itself is a means by which cognition is implemented. A whole, extended, embodied system requires action in order for cognition to take place. The cognizer herself, it seems, must be essentially active if action is essential to cognition. Clark (2008a) supports this, stating that bodily actions are among the means by which certain computations and

---

[272] See Olson, Eric T. (2004). Animalism and the Corpse Problem. *Australasian Journal of Philosophy* 82, 2: 265-274.
[273] Ibid., 269.

representational operations are implemented within a cognitive system. The only difference between extended cognition and internal cognition is that extended cognitive operations are realized not only in the neural system of the body, but in an entire embodied system that incorporates both the body and the world.[274] Cognition is therefore realized through bodily action.

If the mind is intimately connected to the body (a topic that I will address in more depth later on), it seems that it is also closely tied to bodily action. If the subject is essentially inactive and cognition is essentially active, then the mind and the subject become remarkably metaphysically disconnected. The extended mind theorist claims that the division between body and mind is impossible to maintain such that the mind permeates the body. The mind and the body cannot be separately individuated or described as two separate components with an interface between them.[275] If the mind is integrated throughout the body, it must be intimately tied to bodily action. Thus, it seems necessary that a subject be minded, and it seems necessary for mindedness that a subject be active. While the mind may not be identical to the self, the mind (or at least parts of it) seem like they must be part of our self-constitution. What all of this entails is that the minded subject must be essentially active.

§5.5: Revisiting the No-Self View

Even if there is no-self, the extended view probably still runs aground. Rupert (2009) argues that the embedded view of the mind better supports a no-self view than an extended one.

[274] Clark (2008b), 14.
[275] Shapiro, Lawrence (2004). *The Mind Incarnate*. Cambridge, MA: MIT Press. 183.

Rupert argues that the commitments of the no-self position "dovetail nicely" with his primary arguments against the extended approach.[276]

> On the no-self view, cognitive-scientific explanation appeals only to the distributed instantiation of cognitive properties: information, computation, dynamical flow, activation patterns, and the like. These quantities appear far and wide, however, throughout the universe. Thus, a plausible so-self view must endorse a principle of demarcation, one that differentiates between, for example, the ubiquitous flow of information and dynamical evolution and, in contrast, processes that are genuinely cognitive.[277]

Assuming for the sake of argument that human cognitive systems can and do include external resources (and, therefore, are extended), the extended mind theorist must provide a principle of demarcation between causal contributors to the cognitive system and those external contributors that are constitutive of it. "Even if we provisionally accept an extended view of cognitive systems, the need for a principle of demarcation remains."[278] The extended view cannot provide this while consistently holding that the self does not exist.

What principle of demarcation might those who accept the no-self view appeal to? One way to go might be to appeal to the connectionist view of the human cognitive architecture. Connectionist models might be thought to ground the no-self view:

> …in a connectionist network, there would seem to be no single self, no agent making decisions or controlling the transformation of activation vectors. Nevertheless, there is a clear distinction between the causal role of the connectionist network and the causal role of materials that are outside the connectionist network but that contribute causally to cognitive outcomes.[279]

An embedded, systems-based, view can account for this distinction. Rupert puts forth one plausible criteria for demarcation: "A process is cognitive if and only if it is the activity of part of

[276] Rupert (2009), 52.
[277] Ibid.
[278] Ibid.
[279] Ibid., 53.

the persisting connectionist network (the subject's cognitive architecture)."[280] Given that we do not see extrabodily/extended connectionist networks, or even extra bodily networks that are sufficiently integrated with the typically acknowledged, neurally realized ones, Rupert's connectionist-inspired no-self view supports an anti-extended conclusion. External cognitive aids are not integrated with the internal mechanisms of the system in the way that those internal mechanisms are integrated with each other.[281]

Any other potential demarcation criteria put forth cannot be too broad, either.

> To claim that just any instance of information-carrying, computation, or dynamical flow constitutes a cognitive state or process trivializes the extended view…The extension-friendly proponent of the no-self view might have some further principle of demarcation in mind, but it is incumbent on her to present and argue for it.[282]

So, if the extended theorist wants to take up the no-self view in spite of the objections presented above, she needs to give a plausible principle of demarcation that separates genuinely cognitive processes from non-cognitive causal influencers.

> The advocate of the no-self view begins with the claim that cognition is nothing more than a fluid, ongoing interaction of instantiations of cognitively relevant properties, all on a par. In response, I have argues both that there are privileged collections of such property-instantiations and that they appear entirely in the organism…Even here, perhaps especially here, internally instantiated cognitively relevant properties and externally instantiated one exhibit theoretically important asymmetries.[283]

Rupert concludes that his arguments against the extended approach go through on any plausible version of the no-self view.

What we can take away from these arguments is that a plausible no-self view must offer a principle of demarcation for cognitive properties/processes. And, the view must not appeal to any self-centered or personal qualities. Though I take the no-self position to be an implausible

[280] Rupert (2009), 53.
[281] Ibid., 57.
[282] Ibid., 53.
[283] Ibid., 58.

one for the extended theorist to take up based on the persistence of the self in the extended mind literature, a revision of the theory will not conceivably save the extended mind theorist, who is then faced with a demarcation problem that cannot be solved given the central tenets of the extended view.

Truly, little attention has been paid to personal identity and the nature of the self in this literature. These authors simply haven't directly addressed issues of personal identity thoroughly enough. Clark (2008b) states that it must be readily conceded that arguments in favor of the extended mind "interact with vexing questions about personal identity and the nature of the self."[284] And, a full defense of the extended mind would, indeed, need to resolve these issues. Clark further states that "as things stand, the best we can say is that arguments for the extended mind provide further motivation, and perhaps some new kinds of scaffolding, for this important project."[285] Arguments in favor of the extended mind intersect with questions about the self, and the extended mind thesis provides some motivation for investigating these questions, which have yet to be answered. Therefore, this examination is admittedly somewhat impoverished. However, once we begin to examine questions of personal identity in conjunction with the extended mind, it appears that rather than helping to fully defend the extended theory, these considerations only serve to undermine it. Clark is right insofar as he asserts that a full defense of the extended mind would need to resolve these matters. However, it appears that no resolution is readily available.

[284] Clark (2008b), 161.
[285] Ibid., 162.

Chapter 6: Embodied, Embedded Agents

§6.1: Introduction

I believe that, even without extended cognition, we are left with a strongly embodied form of mind that is embedded in an environmental context. By acknowledging that the mind is an embodied system and that it is affected by a highly distributed network of causal connections, we can retain many of the benefits of an extended view of cognition while retaining our fundamental concepts of action and agency.

Historically, cognitive scientists have functioned under the assumption that cognition is driven by computations involving amodal symbols or representations. These mental symbols are manipulated solely on the basis of their syntactic properties.[286] According to this historical assumption, our mental states and processes could therefore be studied without reference to the body.[287] Even theorists not specifically focused on embodiment now consider this to be untenable. This assumption, however, is consistent with two theses. The first is the separability thesis, which states that the human mind could exist without a human-like body. The second is the thesis of envatment. According to the separability thesis properties of the mind and properties of the body are separable.[288] While the mind and the body interact causally, the mind is somewhat autonomous with respect to its body. Separability, then, implies *body neutrality*. This is the doctrine that the kind of mind one has is not determined by characteristics of the body.[289] Rather, the mind is neutral with respect to the body and an explication of the mind can abstract

---

[286] Dove, Guy (2011). On the Need for Embodied and Dis-Embodied Cognition. *Frontiers in Psychology* 1: 1-13.1.
[287] See Spackman, Jonathan and Yanchar, Stephen (2013). Embodied Cognition, Representationalism, and Mechanism: A Review and Analysis. *Journal for the Theory of Social Behavior* 44, 1: 46-79. 49.
[288] Shapiro, Lawrence (2011). *Embodied Cognition*. New York: Routledge. 167.
[289] Dempsey, Liam and Shani, Itay (2013). Stressing the Flesh: A Defense of Strong Embodied Cognition. *Philosophy and Phenomenological Research* 86, 3: 590-615. 593.

away from the physical details of the body. The second view, envatment, holds that the mind, like a brain in a vat, is an entity distinct from the body "and has as its function the processing of signals *from* the body and the computation and transmission of signals back *to* the body."[290] This is a thesis about the mind's overall functional role.

According to the envatment thesis, the mind is a kind of way station. "The mind is a function to which sensory inputs are sent and then processed. From this junction emerge instructions for bodily movements."[291] This view also presupposes the sandwich view, or the input-output picture of perception and action.[292] According to this picture, inputs from the world enter the mind via perception, and then outputs emanate from the mind in the form of action. This is a picture of the mind that invites us to speak of brains in vats, because it is consistent with the view that the mind can be described without reference to the body. All that matters is that the brain receives sensory signals – whether from bodily sensory organs or artificial sources makes no difference.[293] Envatment suggests that the mind and body are separable, or that they can be construed as separate components that communicate only through some interface.[294]

The extended mind theory is one way in which we might reject the hypotheses of separability and envatment. If the mind is extended to include the body and/or external objects, then there is a sense in which it is not, as a matter of fact, separable from these. It is constituted by them such that without these objects the mind would be substantially different. Clark states that our human experience does depend (at least in part) on many idiosyncratic aspects of our embodiment. While it does not follow from this that only creatures embodied like ourselves

[290] Shapiro (2011), 169.
[291] Ibid., 181.
[292] See Hurley (1998).
[293] Shapiro (2011), 181.
[294] Ibid., 215.

could have those same experiences, *we* would not have the experiences that we have were it not for certain idiosyncratic facts about our embodiment.[295]

There is now some general agreement that representations are not simply amodal structures but are tied to sensory and motor modalities. Many researchers and philosophers argue that cognition needs to be viewed as fundamentally based in our bodies and our bodily interactions with the world.[296] Work in embodied cognition in part suggests that bodies permeate minds. One cannot disregard facts about the body when trying to understand how the mind works.

The embodied cognition program advocates for two positions. The first is that we cannot study the mind without also studying the body. The two are not separable as entities of scientific study. The second thing that this program suggests is that the body permeates cognition in such a way that cognition strongly depends on the body, whether constitutively or through causal interactions. This dependence may represent a causal integration of the mind and the body, or a realization of the mind as it is constitutively integrated throughout the body. I will focus mainly on this second claim.

Clark distinguishes between three kinds or grades of embodiment – mere embodiment, basic embodiment, and profound embodiment. A merely embodied creature is one for whom the body is nothing but a highly controllable means to implement practical solutions arrived at by pure reason. Envatment and body neutrality could be true, and a mind could still be merely embodied. This is a creature with a body, but who is not embodied in the above sense of there being an intimate connection between the mind and body. A basically embodied creature is one

[295] Clark, Andy (2008a). Pressing the Flesh: A Tension in the Study of the Embodied, Embedded Mind? *Philosophy and Phenomenological Research* 76, 1: 37-59. 42. See Section 6.2 for more on this point.
[296] Dove (2011), 1.

for whom the body is a resource whose own features and dynamics could be actively exploited allowing for increasingly fluent forms of action selection and control. A basically embodied creature has a body, and has a mind that is highly dependent on the features of that body for its functioning. This is the closest form of embodiment that Clark recognizes to what I defend. A profoundly embodied creature is one that is able to learn to make maximal problem-simplifying use of an open-ended variety of internal bodily, or external sources of order.[297] Only a profoundly embodied creature is one for whom extended cognitive processes are possible.

Merely embodied creatures and basically embodied creatures are organisms whose bodies influence their cognitive capacities. A profoundly embodied creature has a body that is an actual constituent of its cognitive system. This is a creature for which cognition involves a constitutive interplay between brain, body, and world.

Clark, at other times, makes a distinction between simple and radical embodiment. He states that he wants to distinguish two differing ways to treat facts about embodiment and environmental embedding. The first, 'simple embodiment', "treats such facts as, primarily, *constraints upon a theory of inner organization and processing.*" The second, 'radical embodiment' goes further "and treats such facts as *profoundly altering the subject matter and theoretical framework of cognitive science.*"[298] Radical embodiment treats the traditional notions of internal representation and computation as inadequate for explaining cognition, while simple embodiment treats the body as a conduit for information that leads to these computational and representational processes. These are theses about how we ought to explain cognition, or the framework that cognitive science ought to take.

---

[297] Clark (2008b), 42.
[298] Clark, Andy (1999a). An Embodied Cognitive Science? *Trends in Cognitive Science*. 3, 9: 345-351. 348.

§6.2: Embodiment vs. Extension

We can acknowledge that the mind is embodied – or that the body plays a key role in shaping the mind, without granting that the mind extends out into the world. The embodied cognition program claims that cognition is a situated activity.[299] It is not separable from the body and its environment. However, cognition's being situated and embodied does not entail that it must be extended. Cognition can be highly dependent on environmental entities without being extended to constitutively include those entities.

One main claim of embodied cognition is that the body permeates cognition in such a way that the mind is strongly dependent on the body. There are two ways of understanding this claim. The first seems to support extended cognition. "Aspects of body and world can, at times, be proper parts of larger mechanisms whose states and overall operating profile determine (or minimally, help determine) our mental states and properties."[300] Clark calls this the Larger Mechanism Story (LMS). According to LMS, the body is (or minimally, can be) part of the larger mechanism that constitutively determines or helps to determine our mental states. LMS follows naturally from a commitment to functionalism about the mind.

> According to the functionalist, cognitive processes are abstract, in the sense that they cannot be identified with (are distinct from) particular physical parts or processes …something counts as the cognitive process of memory not because it is the hippocampus, but because it stores information in a way that can be accessed for future use. Functionalism as a theory of mind regards cognitive processes such as memory, attention, and perception, as defined by what they do rather than by their physical realization.[301]

---

[299] Wilson, Margaret (2002). Six Views of Embodied Cognition. *Psychonomic Bulletin and Review* 9, 4: 625-636. 626.
[300] Clark (2008a), 39.
[301] Shapiro, Lawrence. Matters of the Flesh: The Role(s) of the Body in Cognition. Unpublished manuscript. March 22, 2016. Microsoft Word File. 4-5.

Because the defining features of the mind are functional, the processes that constitute it might be realized in numerous ways. Functional processes are "locationally unconstrained," and when they are seen as such "the idea that cognition must occur within the confines of a brain betrays nothing more than a kind of brain-based chauvinism."[302] This kind of chauvinism privileges the brain at the expense of external cognitive structures. And, it is exactly this brain-based chauvinism to which the extended mind theory is staunchly opposed.

The second way in which this claim might be interpreted is to understand the mind's inseparability from the details of human embodiment as the body making a special and ineliminable contribution to the mind and its properties.[303] Clark calls this the Special Contribution Story (SC). LMS claims that larger systemic wholes containing brains, bodies, and (perhaps) the environment constitute the mechanistic supervenience base for mental states and processes. SC claims that "specific features of the body (and perhaps the world) make a persistent, non-trivial, and in some sense special contribution to our mental states and processes."[304] This special contribution, however, does not entail the extension of mental states and processes.

Clark points out that there is in fact a tension between accounts that stress features of human embodiment and accounts that stress environmental embedding and intervention. SC-type accounts depict bodily form and sensorimotor patterning as elements that might make a special contribution to human thought and reason. But LMS-type accounts seem to depict bodily action

[302] Shapiro (manuscript), 8-9.
[303] Ibid.
[304] Clark (2008a), 40.

and environmental structuring as merely additional elements in a wider computational, dynamical, and representational nexus.[305]

The claim that our human experience depends in part on aspects of our embodiment does not entail the claim that there is a larger mechanism that constitutes the material supervenience base for cognition. Clearly, LMS lends itself naturally to the idea that the mind is extended. SC, however, does not. The mind can be embodied in the sense of SC without being extended. In other words, bodies can matter without cognition spreading beyond them.

SC is also independent of LMS in the sense that LMS can be true without the body being a particularly special contributor to cognitive processing. This is true because the body, while an important feature that structures cognition, is not unique when compared to other features of the external environment. The body is nothing more than the thing that plays a certain functional role in an information-processing economy. Extended functionalism of the sort that Clark advocates embraces LMS without accepting that the body provides a unique contribution to mental processing.[306]

Both LMS and SC stress that embodiment plays a crucial role in cognition, but they do so in different ways. "The wide computationalism of LMS treats the role of the body, not as making a special or unique contribution, but of simply adding, along with environmental structuring, additional elements to the cognitive processing in question." The body makes no *unique* contribution to cognition on this view. If defenders of LMS are correct, two creatures with different bodies could have identical mental states. The mind can compensate for different bodily structures. According to the compensation thesis, the same mind is realizable in a radically different body just so long as other contributing factors, like the brain and environment, are

[305] Clark (2008a), 49.
[306] Dempsey and Shani (2013), 594.

adjusted such that the entire system realizes the same computational profile.[307] The compensation thesis "suggests that different embodiments may yield the same experience while, conversely, identical embodiments may culminate in distinct experiences." Clark has argued for the compensation thesis.[308] For any particular organism, there is an environment such that the organism's brain, body, and that environment together could form a system capable of completing any information-processing task that some other organism can complete.[309] The character of mental experience depends on the overall character of the larger functional context of agent-environment interaction.[310]

LMS claims that the underlying material base of our mental states and processes can include the brain, the body, sensory organs, and non-biological external props and cognitive aids. LMS is, in this way, an obvious consequence of the functionalist perspective. SC also claims that the body and sensory organs make a special contribution to our mental states and processes, but SC does not admit that the body, nor the external environment, constitute the material supervenience base of our mental states and processes such that these are partially constitutive of these states and processes. If SC defenders are correct, the body grounds, determines, and shapes our cognitive capacities. SC asserts that the body makes a special, unique, and typically irreplaceable causal contribution of the body to cognitive processing. LMS and SC, then, offer

[307] Dempsey, Liam and Shani, Itay (2015). Three Misconceptions Concerning Strong Embodiment. *Phenomenology and the Cognitive Sciences* 14, 4: 827-849. 828.
[308] Some argue that the compensation thesis is false, and thus that the body does make a special and ineliminable contribution to cognitive processing that could not be compensated for. See in particular Dempsey and Shani (2013 and 2015). I will not dwell on this issue here, as to reject the compensation thesis would be to reject functionalism about the mind (See Clark 2008a, 53), and this would be to deny one of the central tenets of many extended views.
[309] Shapiro (manuscript), 11. See also Clark (2008a), 42.
[310] Dempsey and Shani (2013), 598.

one way of drawing a line between an extended approach to cognition and a basically embodied approach.

§6.3: The Body without Profound Embodiment

The boundaries of the organism are significant for multiple reasons. The body is a common, persisting locus of willed action.[311] The body is also an obligate system, rather than a facultative one. Obligate systems are more or less persistent, while facultative systems are temporary.[312] Extended cognitive systems are facultative, and thus cannot be a persisting locus of willed action. This is part of what leads to the agency problem. The body, however, is an obligate system that persists more or less over time. The body, as a system, retains its identity over time in a way that extended cognitive systems do not. This is because extended cognitive systems are constantly changing in their component makeups. The body, however, is not a fully closed system with respect to its environment, as the system perceives input from the external world, and it constantly responds to this information by producing outputs.

The body is also an interface that makes it possible to carve a coupled system into body-bound and world-wide components. Interfaces are points of interactive contact that can be more or less articulated. The more clearly articulated these points are, the more motivation there is for construing the component parts at the point of contact as separable and distinct.[313] While these components may be tightly coupled at times, they are importantly separable. And, the more separable two components are, the less tight their coupling, the more persuasive it is to see them as interfacing rather than as one minded system. Our bodies cause just this kind of separability.

[311] See Clark (2008b), 207.
[312] Wilson (2002), 630.
[313] Shapiro (2011), 179.

The body represents the exact kind of distinct component system separable from the individual features of an agent's environment.

Alsmith and de Vignemont (2012) note that "The amount of information received from the body is quantitatively superior to that received from any environmental object."[314] Not only can we see and touch our bodies, we also receive a continuous flow of proprioceptive and interoceptive inputs from them. The body, therefore, presents a highly special feature compared to any other object in the world. It provides a special contribution that the environment does not provide.

The above gives us reason to privilege the body, but there are additional reasons to discount the cognitive efficacy of the external world. External symbols are causally passive. This is a charge that has been levied against Clark by several authors.[315] The information that external symbols encode doesn't do any work until we perceive them. It is causally passive until then. When we perceive the information it becomes active, but only because it has been re-coded inside of our skulls. When externally recorded information does become causally engaged with parts of the mind it does so only in a causally discrete fashion: "Each separate piece of information, coded by a distinct symbol structure, must be individually accessed and processed."[316] External information needs to be re-coded in order to be accessed and utilized. These differences would seem to mark an important natural boundary, one that makes it more difficult to justify any extension of the mind's representational substrate to include external

[314] Alsmith, Adrian and de Vignemont, Frederique (2012). Embodying the Mind and Representing the Body. *Review in Philosophy and Psychology* 3: 1-13. 2.
[315] See in particular Adams, Frederick and Aizawa, Kenneth (2008). *The Bounds of Cognition*. Malden, MA: Blackwell.
[316] O'Brien, Gerard (1998). The Mind: Embodied, Embedded, but not Extended. *Metascience* 7: 78-83. 82.

states of the environment.[317] The body, while an important interface with the rest of the world, does not provide for an extension of the mind out into the world.

§6.4: A Moderate Approach to Embodiment

If cognition is not profoundly embodied, when and how is it embodied? Alvin Goldman has put forth a view, similarly espoused by Frederique de Vignemont, that cognition is embodied when it is performed using representations that are encoded in bodily formats.[318] Goldman defines embodied cognition in the traditional sense of giving necessary and sufficient conditions for its occurrence. The position that Goldman advocates is this: "Cognition C is a specimen of embodied cognition if and only if C uses some member of a special class of codes or formats for representing and/or processing its content, viz., a body-related code or format (B-code or B-format)."[319] The brain employs various codes in "a neural syntax for the purpose of representing a variety of states of affairs."[320] Different kinds of codes might be used for representing different kinds of facts.

> For instance, some neurons might encode some features of a stimulus (e.g., light intensity) in terms of rate of firing, but rate of firing might be a poor way to encode other features of a stimulus. Spatial location too might play an important coding role, as it seems to in various topographic maps of sensory surfaces.[321]

B-codes are merely one of the hypothesized codes used to encode bodily features.

---

[317] O'Brien (1998), 82.
[318] An earlier version of a similar theory is explored by Prinz, Jesse (2009). Is Consciousness Embodied? In P. Robbins and M. Ayede (eds.), *The Cambridge Handbook of Situated Cognition* (pp. 419-436). New York: Cambridge University Press. 430- 431.
[319] Goldman, Alvin (2012). A Moderate Approach to Embodied Cognitive Science. *Review of Philosophy and Psychology* 3, 1: 71-88. 73.
[320] Shapiro , Lawrence (2014). When is Cognition Embodied? In U. Kriegel (ed.), *Current Controversies in Philosophy of Mind* (pp. 73-90). New York: Routledge. 79.
[321] Shapiro (2014), 79-80.

Thus, embodied cognitions are those that use a B-code or B-format that represents properties of the body. Representations of the body, therefore, are pervasive and important to cognition. Systems of bodily representations are the key to embodied cognition. "Specifically, embodied cognition is the application of *special* systems, systems *dedicated* to inner bodily representation."[322] Given that there are multiple representational systems in the brain, Goldman postulates that a subset of these systems is dedicated (or was originally dedicated) to representing bodily subject matters from an internal point of view.

Goldman states that there are two core elements to his approach. The first element of his approach is the idea of bodily representational codes – codes that are fundamentally utilized in forming representations of one's own bodily states and activities.[323] The second element of the approach is the idea of redeployment of these cognitive processes. This second element "adduces wide-ranging evidence that the brain reuses or redeploys cognitive processes that have different original uses."[324] Goldman proposes to define the philosophical concept of embodied cognition in terms of these B-codes or B-formats. For some form of cognition to be embodied, it is both necessary and sufficient that it use some member of this special class of codes or formats – body-related codes or formats. Further, many cognitions seemingly unrelated to bodily states might be embodied if B-formatted representations are implemented in their processing.

The proprioceptive representational system might be the clearest example of a system that represents bodily subject matters from an internal point of view. Unlike the visual system, which represents parts of the body from an external perspective (and therefore does not utilize B-

[322] Goldman, Alvin (2014). The Bodily Formats Approach to Embodied Cognition. In U. Kriegel (ed.), *Current Controversies in Philosophy of Mind* (pp. 91-108). New York: Routledge. 100.
[323] Goldman, Alvin and de Vignemont, Frederique (2009). Is Social Cognition Embodied? *Trends in Cognitive Sciences* 13, 4: 154-159.
[324] Goldman (2012), 71.

codes), the proprioceptive system is one that has represent the states of an area of the body as being a certain way or undergoing certain changes from an internal perspective.

> Proprioception and kinaesthesis give the brain information – couched, presumably, in distinctive formats – about states of one's own muscles, joints, and limb positions. These interoceptive senses are the basis for B-formats of representation. One's own body, or selected parts thereof, is what they *primarily*, or *fundamentally*, represent.[325]

Proprioception and kinaesthesis provide information in distinctive formats about states of one's body. These senses are the basis for B-formatted representations, which primarily and fundamentally represent the body.[326] However, these are not the only systems whose cognitive tasks make use of B-formats for their processing.

The cognitive tasks that utilize these bodily representations need in no way be involved in bodily tasks, such as intentional actions or reflexes, which overtly lead to bodily motions and may in some cases directly influence objects in one's external environment. Such representations are embodied not because of their use, but because they belong to the representational system whose purpose or function is to represent the body (or, its parts). If I look at my hand, the representation of it is coded in a visual format as opposed to a bodily format. This visual format is not one whose original purpose or function is to represent the body. Because of this, the representation is not embodied even though the content of the representation includes a part of the body.[327]

However, embodied cognitions do go beyond things like proprioception, which is clearly dedicated to representations of one's body. This form of cognition, dedicated to representing bodily orientation, is a clear candidate for cognitions encoded in bodily formats. Embodied cognition, however, is far vaster than this. This is because B-codes can be reused or redeployed

---

[325] Goldman (2012), 73.
[326] Ibid.
[327] Shapiro (2014), 80.

in cognitive tasks that have little obvious connection with the body. Goldman hypothesizes that these B-formats are redeployed or co-opted for representing non-bodily items. These representations of non-body parts and states also qualified as embodied cognitions.[328]

> The underlying idea is that reuse of neural circuitry for a variety of cognitive purposes is a central organizing principle of the brain. In other words, it is common for neural circuits originally established for one purpose to be exapted (exploited, recycled, redeployed) during evolution or normal development and put to different uses, without necessarily losing their original functions… Accordingly, rather than posit a functional architecture for the brain in which regions are dedicated to large-scale cognitive domains like vision, audition, and language, respectively, neural reuse theories posit that low-level neural circuits are used and reused for various purposes in different cognitive tasks and domains.[329]

Research suggests that B-format representations are tacitly at work during the execution of a variety of tasks that we perform on a daily basis. There is wide-ranging evidence that the brain redeploys cognitive processes for different purposes than their original use. If it turns out that cognitive processes in B-formats are redeployed in this manner, such that they represent things other than states of the body, then the representations that result from this redeployment also count as embodied cognitions.[330] A B-code that represents facts about the body may be used for a purpose other than its original purpose.

> The Tin Man, we all know, had a funnel for a hat. This illustrates the obvious point that items made for one purpose might be "reused" for another. An item originally used for pouring liquids into small-mouthed containers gets reused for something else – namely, a hat. Turning from artifacts to organisms, we find that evolution is an expert reuser.[331]

Embodied cognitions do not need to be limited to proprioception and kinaesthesis, for example, because B-codes that represent states of the body may be reused in tasks that have little obvious connection to the body.

---

[328] Goldman (2012), 74.
[329] Ibid., 75.
[330] Ibid., 74.
[331] Shapiro (2014), 81.

De Vignemont (2011) proposes a similar thesis. Cognition is embodied, according to de Vignemont, if it is processed in the same way as the properties of one's body. Embodiment corresponds to a specific mode of presentation of a property, which results from a specific way the property is processed.[332] This is consistent with the thesis that cognition is embodied if B-formats are used for its coding and processing. Embodiment, in both cases, results from a specific way in which a property is processed. This definition of embodiment, in terms of the types of processing that characterize the representations of one's body, does not give any information about the way in which the properties of one's body actually are processed.[333] However, the theory does tell us that the same kind of processing must be in play in order for cognition to be embodied.

The massive redeployment hypothesis, or the neural reuse principle, embodies the underlying idea that reuse of neural circuitry for a variety of cognitive purposes is a central organizing principle of the brain. In other words, it is common for neural circuits originally established for one purpose to be exploited, recycled, or redeployed and put to different uses, without necessarily losing their original functions.[334]

> Accordingly, rather than posit a functional architecture for the brain in which regions are dedicated to large-scale cognitive domains like vision, audition, and language, respectively, neural reuse theories posit that low-level neural circuits are used and reused for various purposes in different cognitive tasks and domains.[335]

This gives us substantial evidence that embodied cognition is in fact a pervasive form of cognition. Embodied cognition occurs when processing involves the use of B-formats, which

---

[332] de Vignemont, Frederique (2011). Embodiment, Ownership and Disownership. *Consciousness and Cognition* 20, 1: 82-93. 84.
[333] de Vignemont (2011), 84.
[334] Goldman (2012), 75.
[335] Ibid., 76.

fundamentally and primarily represent the body, but also when these neural codings are redeployed or reused for different purposes other than pure bodily representation.[336]

Bhalla and Profitt (1999) acquired results consistent with the theory that B-coded representations influence representations of non-bodily objects. In this study, subjects' judgment of the geographical slant of a hill was heavily influenced by the physiological potential of their own bodies. The experimenters found that when subjects were required to wear a heavy backpack while judging the slant of hills, they overestimated the slant of those hills more than did subjects in the control condition. The verbal and visual overestimation of the hill slant increased with a decline in behavioral potential due to the load of the backpack.[337] This experiment is consistent with the idea that B-coded representations, like the representation of the weight on one's back, influence representations of non-bodily objects.

This form of embodied cognitive science "predicts that neural circuits originally evolved for one purpose will be reused in later developing functionality." And, there has been experimental work demonstrating exactly this. "There is, for example, ample evidence that verb retrieval tasks activate brain areas involved in motor control functions, and naming colors and animals (i.e., processing nouns) activates brain regions associated with visual processing."[338] Cognitive abilities are not cleanly separable from one another, according to this hypothesis. Bodily structure and activity are deeply implicated in cognitive processes and cannot be isolated in their own category of study. Anderson et al. (2012) cite studies that point to the influence

---

[336] Goldman (2012), 80. Also see Goldman (2012) for evidence of redeployment.
[337] Bhalla, Mukul and Proffitt, Dennis (1999) Visual-Motor Recalibration in Geographical Slant Perception. *Journal of Experimental Psychology: Human Perception and Performance* 25, 4: 1076-1096.
[338] Anderson, Michael, Chemero, Anthony, and Richardson, Michael (2012). Eroding the Boundaries of Cognition: Implications of Embodiment. *Topics in Cognitive Science* 4, 4: 717-730. 721.

between motor control and biographical memory and studies that demonstrate that abstract planning can activate motor areas of the brain. Motor circuits are also involved in number processing. These are only a few examples of how B-coded processes might be redeployed during the occurrence of cognitive activities with different functions.

In fact, evidence shows that such neural overlaps are the general rule, instead of modularity in neural systems being the norm. A typical neural region is involved in supporting diverse tasks across cognitive domains such as vision, audition, attention, language, memory, abstract reasoning, etc. Even small neural regions typically support tasks across multiple of these domains. "In addition, it appears that the functional complexes supporting tasks in newer—more recently evolved—cognitive domains utilize more and more widely scattered circuitry than do the complexes supporting older functionality like vision and motor control (Anderson, 2007, 2008)."[339] These findings suggest an evolutionary pathway that favors the re-use or redeployment of existing neural circuitry to support new functions. This is instead of generating new, dedicated neural circuits dedicated to each emerging purpose.[340] Our cognitive processing re-uses or recycles existing circuits in order to support newer needs.

The above view is consistent with the thesis that the details of one's embodiment profoundly affect the nature of one's experience, including perceptions and actions.[341] This seems to support, or is at least fully consistent with, SC rather than LMS. For, this moderate approach to embodiment acknowledges that the body is important, and perhaps fundamentally special, when it comes to our neural coding and redeployment of representational formats that

[339] Anderson et al. (2012), 722.
[340] Ibid.
[341] Dempsey and Shani (2013), 599.

guide so much of our cognitive processing.[342] This kind of intracranial processing is not non-bodily simply because it takes place intracranially. Embodied cognition is a significant and pervasive part of human cognition. The studies described above show how the body shapes, determines, and influences cognition. The body is thoroughly implicated in cognitive processing due to that processing utilizing representations in B-formats. The body is therefore implicated in the operations of cognition without cognitive processing being constituted by a larger environmental mechanism.

§6.5: An Objection to B-Format Embodied Cognition

One objection to this view is that it might be consistent with the hypothesis of envatment, as explicated above. "Because activities of the body are screened off by the activation of B-codes, cognition turns out to be an exclusively neural event."[343] Cognition, therefore, ends up being an exclusively brain-bound activity. Stating that representations of the body depend upon the body does little to assuage this worry, as this dependence relation does not seem *necessary* for cognitive processing using B-codes to take place. Explicating embodied cognition in terms of B-formats preserves the idea that cognition is 'pure cognition', where the brain creates models that allow it to discard the world almost in its entirety. Action, according to this paradigm, is a mere means of implementing solutions arrived at by pure cognition. This seems consistent with envatment and the sandwich view, or the input-output picture of perception and action.

Goldman responds to this objection by stating that envatted brain states would not have the same contents as the brain states of embodied minds simply because the content of these

[342] Dempsey and Shani (2013), 604.
[343] Shapiro (2014), 84.

brain states is a function of what they causally interact with.[344] In other words, cognitive processing in the brain (and maybe cognitive artifacts in the world) would not be enough to compensate for different body types, thus allowing for differently embodied creatures to have similar bodily representations. The body and our representation of it are too closely entangled.

Additionally, it might be the case that the body is implicated in cognitive processing through the representations of the body that are involved in cognition. Perhaps "certain types of representation are so closely dependent upon the non-neural body (i.e. the body besides the brain), that their involvement in a cognitive task implicates the non-neural body itself."[345] The idea here is that representations of the body do not automatically screen off cognition from the body. These replies, though, may not be strong enough to satisfy the exponent of cognitive extension, for it remains possible that the contribution of the body could be largely insignificant. The first reply may also deny functionalism about the mental, as it appears to deny the compensation thesis.[346]

I do not think, however, that this objection is enough to refute Goldman's hypothesis. The body, according to Goldman's explanation of embodied cognition, still makes a special contribution to cognition, through its influence on the cognitive system's representations. The body ends up permeating cognition through these representations, on which cognition strongly depends. However, the body is not part of a larger mechanism of distributed cognition that crosses from brain to body and from body to world.

Fred Adams gives one account of how representations (motor representations or M-representations) can be embodied. Adams addresses the important point that the simple fact that

[344] Goldman (2014), 104.
[345] Alsmith and de Vignemont (2012), 2.
[346] Clark (2008a), 53 indicates that a denial of the compensation thesis is not reconcilable with any recognizable form of functionalism.

something is a representation does not entail that it is a pure cognition, envatted and body-neutral. Cognitive states can be situated and still be representational in nature. Motor-intentions (M-intentions), which lead to bodily movements and actions, involve M-representations that are situated. Action theory is committed to there being representational mental states in the initiation, production, sustaining, and terminating behavior.[347]

Dorsal stream processing extracts visual information from the environment about objects and uses this information to build motor representations. And, "the representations of the necessary movement reflect implicit and tacit knowledge of biomechanical constrains and the kinematics and dynamics of movement."[348] Motor representations, however, are relational. They represent relations between the body and states of the environment, or motor patterns that objects afford to the agent.[349] This makes our M-intentions, which they partially constitute, situated cognitions. M-intentions are also clearly context-sensitive, as they depend upon the goal of the bodily movement and the objects with which the body interacts. These are representations that are situated, coupled, and time-pressured. These are cognitive states that have their content as a function of what they causally interact with. This does not mean that environmental objects constitute the cognitive processes leading to and constituting M-representations. The particularity being referred to is secured by causal connections, not constitutive ones. An M-intention is about a particular action, or coded for a particular action, in virtue of being causally connected to the body performing the action. This is the body to which cognition has become finely tuned and to which the brain is causally connected.[350]

[347] Adams (2010a), 235-236.
[348] Ibid., 237.
[349] Ibid., 245.
[350] Ibid.

Motor-intentions seem like the kind of cognitive processes that would need to use B-codes or B-formats, or to redeploy these. Thus, representations using B-codes are not incompatible with some of the central tenets of embodied cognition, such as cognition being on-line, relational, situated, and cognition being for action.

§6.6: Distributed Causality vs. Distributed Cognition

Persons clearly exploit their environments when acting. They use their environments in order to process information. However, this usage, according to this more modest account of embodiment, is causal instead of constitutional. The wider world is still available to the agent as an arena for embodied action, but there is a mediation by the body between the world and the cognizing agent. This mediation comes in the form of the body feeding information into the cognitive system that then translates this information into representations, B-coded or otherwise.

The idea of environmental scaffolding is often used to explain how the environment can support human cognition without being part of human cognitive processing. It is undeniable that human cognition depends upon environmental resources, in particular human cognition that leads to intelligent action. One of the ways in which organisms make use of their environments is to construct niches. Many animals intervene in order to shape their environment to improve the adaptive fit between the animal and the world. These animals not only adapt to their niche, they construct it to fit their own needs.[351] Many examples of purportedly extended cases of cognition can, instead, be seen as special cases of niche construction. These are "cases in which human competences depend intimately on the environment being scaffolded to support adaptive

[351] Sterelny, Kim (2010). Minds: Extended or Scaffolded?. *Phenomenology and the Cognitive Sciences* 9, 4: 465-481. 466.

decision making."[352] This dependence of human cognitive processing on the environment, however, is meant to be understood causally, in terms of environmental support, rather than constitutionally.

Scaffolding can take place in many ways, shapes, and forms. Our digestive system, for instance, is environmentally scaffolded. We use our environment to cook, chop, and otherwise prepare food to take some of the burden (of chewing and otherwise breaking down food) away from our internal digestive system. Our digestion is environmentally, technologically, supported "in profound and pervasive ways."[353] But, the environment plays an indispensible role in supporting and amplifying other human functions as well, including cognition. Agents adapt to their environments, but they also cause their environments to adapt to them. Traditionally, these niche constructions are direct practical physical interventions, like a bird building a nest or a beaver building a dam. However, epistemic actions like altering the informational character of one's environment (as Otto does with his notebook) also fall into the category of niche construction. Cognition is environmentally supported in that skills develop in a world that is both rich in relevant information and adaptively organized to make that information more salient. [354]

> The scaffolded mind hypothesis proposes that human cognitive capacities both depend on and have been transformed by environmental resources. Often these resources have been preserved, built or modified precisely because they enhance cognitive capacity. The extended mind hypothesis proposed that human cognitive systems include external components. These components are coupled to human bodies, but not located within human bodies.[355]

Scaffolding is one way of explaining our embeddedness in the world. Environmental support of one's cognitive system, however, is far different than the environment constituting one's

[352] Sterelny (2010), 466.
[353] Ibid., 467-468.
[354] Ibid., 470-471.
[355] Ibid., 472.

cognitive system. Just as our digestive systems are not constituted by food processors, blenders, etc., our cognitive systems are not constituted by iPhones or notebooks.

This does not deny that we are agents tightly coupled to features of our surrounding environment. This merely denies that these surrounding features are constitutive of our cognitive systems. "Many forms of cognition are embedded in that they are deeply dependent upon features of the individual's physical body" such that the body plays some significant causal role in cognitive processes.[356] Non-extended theorists merely insist that such structures and processes in the environment, while causally active, "influence or act merely as input into intracranial cognitive processing (Adams and Aizawa 2010, 2009, 2001; Rupert 2004). These external structures and processes are never constitutive, that is, components of genuinely cognitive systems."[357] Though many consider embodied and embedded cognition as leading to extended cognition, the two approaches are separable, as embodied/embedded cognition is entirely compatible with the denial of extended cognition. Both are compatible with neural and non-neural components being involved in inside-the-skin cognition, the claim is merely that extra-bodily components do not constitute cognitive processing.[358]

[356] Greenwood, Jennifer (2015). Is the Mind Extended or Scaffolded? Ruminations on Sterelny's (2010) Extended Stomach. *Phenomenology and the Cognitive Sciences* 14, 3: 629-650. 631.
[357] Ibid., 632.
[358] Ibid.

Chapter 7: Embodied Actions

§7.1: Bodily Representations and Action Production

According to the theory of embodiment examined in the previous chapter, cognition is embodied to the extent that it uses or reuses representations in B-formats. Further, representations encoded in B-formats are necessarily implicated in the motor representations that generate and guide actions. The neural underpinnings of these bodily representations are used in action-oriented representations and motor representations. These representations are employed in the initiation, guidance, and overall execution of action. Because bodily representations, or representations in B-formats, are implicated in these representations, the cognition executed by this particular kind of neural deployment is likewise embodied cognition. I will call representations of the body in B-formats basically embodied cognitions. Because a cognitive agent embodies the representations that initiate and guide actions, the actions these representations guide are also embodied. The agent, who controls these actions, is then metaphysically responsible for them.

Action-oriented representations are representations that include in their contents commands for certain behaviors.[359] Action-oriented representations are specifically representations of the body that carry information about the bodily effector (and the bodily goal in reflective actions) that is used to guide basic actions.[360] Reflective actions are simply those actions where the body is represented as the goal of the action and the effector of the action. By their very nature, therefore, these representations implicate B-formatted representations.

---

[359] Mandik, Pete (2005). Action-Oriented Representation. In A. Brook and K. Akins, eds. *Cognition and the Brain* (pp. 284-308). New York: Cambridge University Press. 285.
[360] de Vignemont, Frederique (2010). Body Schema and Body Image – Pros and Cons. *Neuropsychologia* 48, 3: 669-680. 672.

Motor representations include in their contents information about the action to be performed, and they are 'commanding' in so far as the neural activation that underlies them does not dissipate before the completion of the act.

> …representations are built on sustained neuronal discharge arising in structures relevant to the various stages of the preparation of motor acts. The neural layout of the representation can be figured out as a network where the goal information (information about the action to be performed) appears as an array of enhanced activity…This description implies that return to rest does not occur until a signal is given that the goal encoded by each unit [in the network] has been attained.[361]

Motor representations must be related to the preparation of an action, but not necessarily essential to its execution. They must present a sustained neuronal discharge for as long as the action is not completed. And, they are typically not influenced by the presence or absence of an actual target, but rather a bodily goal. The neuronal discharge corresponding to the motor representation encodes an internal goal, instead of an objective target, and relates to a motor intention, and not a movement.[362] Motor representations represent the outcome of actions in terms of a bodily goal, coordinate their execution, and they thereby facilitate the occurrence of the represented outcomes.

From a neural point of view, the underpinnings of a motor representation are a network where the information about the action to be performed, or the goal, corresponds to an array of enhanced activity, which persists until the action is complete.[363] Motor representations must also involve not only mechanisms for steering and directing action, but also mechanisms for monitoring and correcting its course, and checking whether the action has been completed.[364] The system of central monitoring related to motor intentions, the direct antecedents to

---

[361] Jeannerod, Marc (1997). *The Cognitive Neuroscience of Action*. Cambridge, MA: Blackwell. 165.
[362] Ibid.
[363] Ibid., 164.
[364] Ibid., 174.

movements, includes the monitoring of efferent signals related to motor intentions and afferent signals indicating whether the represented goal has been met.

In accordance with B-format embodied cognition, circuits involved in bodily information processing are also used in the generation of action. In particular, B-formatted representations are involved in the production of motor representations. Body schemas, unified representations of the body theoretically encoded in B-formats, are heavily involved in the production of action. These two types of representation – action-oriented representations/motor representations and B-formatted representations – significantly overlap.

It is necessary to represent the body, or a body schema, in order to represent the way in which the body can move given environmental affordances. Encodings of the body must be involved in motor representations in order for cognition to direct the body appropriately. Motor representations such as these must be grounded at least partially in representations of the body, and these are in B-formats.

Proprioceptive attention, in particular, usually at an unconscious level, is necessary for the physical implementation of action. One must be proprioceptively aware of bodily movement and bodily positions across time in order to complete intentional actions.[365] While proprioceptive attention and proprioceptive content is processed at a low level, it still underlies and grounds the motor representations necessary for action execution. Proprioceptive awareness is therefore an important condition of intentional action. And, further, this proprioceptive content is one primary candidate for the content represented in B-formatted representations.

[365] O'Shaughnessy, Brian (1995). Proprioception and the Body Image. In J. Bermudez, A. Marcel, and N. Eilan (eds.), *The Body and the Self* (pp. 175-204). Cambridge, MA: MIT Press. 179.

Intentional action requires knowledge at the very least of the location of one's body parts in space.[366] A defining, though not definitionally sufficient, feature of an action-oriented representation is that it represents the body, and a defining feature of the body schema is its role in the motor hierarchy. The motor system uses the body schema at different stages in its functioning. In computing motor commands that lead to action, the motor system uses both an inverse model and a forward model. Both of these models involve the body schema in their computations.

> The inverse model has the role of computing the motor commands needed to achieve the desired state given the agent's current body state. The inverse model is thus fed by the initial body schema, including long-term information like the size of the limbs, and short-term information like the joint angles and the hand position. In parallel, the motor system anticipates the sensorimotor consequences of the movement through the forward model. The forward model predicts what an action will be like given the specific body that executes the motor commands. It results in the predicted body schema, which carries information only about the bodily parameters that will be altered by the movement like the hand position. It is involved in motor imagery and allows anticipatory control of movements. Finally, there is the actual sensory feedback resulting from the execution of the action that provides the updated body schema, which again carries information only about the bodily parameters that have been altered.[367]

Both the predicted body schema and the updated body schema are online short-term body representations. These are dynamic and constantly updated. Only the initial body schema, used by the inverse model, includes long-term body representations along with short-term ones.

Proprioception (as well as vision and other sensory modalities) is important in coding the on-line position of the body in space and therefore important in planning motor activity of the body. There are, as mentioned above, short-term on-line representations of the body, which include representations of posture at any given time, and which are constantly being updated.

---

[366] Shenton, Jeffrey, Schwoebel, John, and Coslett, H. Branch (2004). Mental Motor Imagery and the Body Schema: Evidence for Proprioceptive Dominance. *Neuroscience Letters* 370: 19-24. 19.

[367] de Vignemont (2010), 673.

There are also long-term off-line representations of the body, which represent long-term bodily properties like the spatial organization of body parts and their respective size, and these are relatively stable over time, without constant updating.[368] The body schema includes both short- and long-term representations of the body. And the production of action involves both of these.

Successful bodily action therefore would simply not be possible without a representation of the parameters of one's body, including the size of one's limbs and their relations to one another. Many bodily representations simply are action-oriented representations because they include in their make up commands for basic actions. Many bodily representations, therefore, will include both a bodily component and action-specifying properties. The body schema is sometimes simply defined as the unified group of sensorimotor representations that guide actions. All the body-schema really is, then, is a set of action-oriented representations.[369]

Proprioceptive input provides both the coding for the body schema and the information for motor representations that lead to overt actions. Motor control is heavily influenced by proprioceptive feedback, as evidenced by subjects who lack such proprioceptive feedback due to abnormalities in the nervous system. Subjects without proprioceptive feedback have difficulties doing even the most basic motor activities such as walking, eating, speaking, etc. It takes intense mental concentration on these subjects' part and constant visual vigilance in order to move with any degree of control. In patients without proprioceptive feedback in their faces, facial expression becomes extremely limited. The visual feedback necessary to move effectively must be used with concentration and intellectual effort.[370] Patients without afferent feedback often

[368] de Vignemont (2010), 672.
[369] Ibid., 672-673.
[370] Cole, Jonathan, and Paillard, Jacques. (1995). Living with Touch and Peripheral Information about Body Position and Movement: Studies with Deafferented Subjects. In J. Bermudez, A.

make up for these deficits by using visual monitoring to build a representation of their bodies in order to guide actions. The proprioceptive information that is lacking in these cases is necessary for updating the postural body schema. These deafferented patients have, in effect, lost their body schema. They have lost the ability to update their body schema, and cannot locate the place of their sensorimotor body in space.[371]

Ideomotor apraxia is another disorder that evidences the motor deficits caused by disruptions in the body schema. The disorder is primarily characterized by difficulties in imitating both meaningful and meaningless gestures, which require matching someone else's body or body part onto one's own body. But, persons with this disorder have multiple motor problems. Patients with ideomotor apraxia often mislocate the goal of their actions, even when they goal is their own body – such as brushing one's chin instead of one's teeth. They are also deficient in locating the body part, e.g. the hand, which will carry out the action. However, their inability to imitate meaningless gestures is of particular interest because this form of imitation relies exclusively on the encoding of bodily attributes of the movements to be imitated, "i.e. encoding the position of one's own body, abstracting the target position from individual differences, and matching one's position onto the other's position."[372] There is no stored information to confer a processing advantage in meaningless imitation. The main deficit for these subjects is the deficit in locating or encoding one's own body parts. The primary sensorimotor body schema that guides one's movements is deficient in cases such as these. These motor deficits are the result of deficits in the body schema, which make it impossible to form the appropriate motor representations to guide one's actions.

Marcel, and N. Eilan, (eds.), *The Body and the Self* (pp. 245-266). Cambridge, MA: MIT Press. 248-250.

[371] Cole and Paillard (1995), 254.

[372] de Vignemont (2010), 675.

It has been proposed by Goldenberg (1995) that apraxia is a deficit in forming a motor intention in relation to a representation of the body. It is therefore related to an impairment of this representation, or an impairment of the body schema. Studies demonstrate that "apraxic patients show a general deficit in the representation of body part configurations and relationships." The proposed model in particular theorizes that imitation of meaningless gestures "requires a coding of gestures with reference to a classification of body parts."[373] Without the ability to represent these body parts and their relations, gestures cannot be coded effectively. So, impairment of the body schema leads to impairment in motor acts, such as the act of imitation.

Personal neglect is an example a different deficit. "Personal neglect is clinically defined by a lack of exploration of half of the body contralateral to the damaged hemisphere."[374] This occurs when cognition fails to *attend to* the body schema. This attentional deficit effects motor imagery, motor representations, and therefore action. This deficit produces a lack of actional movement from one side of the body. There is also a lack of awareness of this side of the body.

The association between the body schema, a prototypical case of a B-formatted representation, and the motor representations that lead to actions, is strong. The body representation is, itself, primarily referenced to action.[375] And, motor representations must refer to the body schema in order to properly move the body towards a goal state. Thus, motor representations are typical cases of embodied cognitions.

The overlap in processing between bodily representations and motor representations is evidenced when the body schema diverges from the norm. In cases of phantom limb, when a

---

[373] Chiminade, Theirry, Meltzoff, Andrew, and Decety, Jean (2013). An fMRI Study of Imitation: Action Representation and Body Schema. *Neuropsychologia* 43, 1: 115-127. 116.
[374] de Vignemont (2010), 676.
[375] Kinsborne, Marcel (1995). Awareness of One's Own Body: An Attentional Theory of Its Nature, Development, and Brain Basis. In J. Bermudez, A. Marcel, and N. Eilan (eds.), *The Body and the Self* (pp. 205-224). Cambridge, MA: MIT Press. 219.

subject experiences the sensation of having a body part that is in fact absent, this phantom will guide the actions of the subject experiencing it. This phantom is part of the subject's body schema. For instance, when reaching for an object, a subject may fall short of the object because she experiences an extended limb that is not there. The patient experiences the limb, and even incorporates it into the movements of her body. "For instance, when a man with an amputated leg stumbled, he felt himself extend the missing phantom leg, as it were, to save him from falling."[376] Persons with phantom limb have a body schema, which guides their actions, but which is faulty and therefore leads to unsuccessful action. There is a persisting motor intention, referenced to the body schema, even though the body schema is misrepresenting the actual body.[377] These motor intentions, erroneous as they may be when it comes to guiding action, are a result of a morphed body schema. The neural encodings of the body that are deployed for the creation of motor intentions and the guidance of action continue to be deployed even when the body schema is morphed in this manner.

This evidence supports the idea that action is embodied, because it is grounded in representations of the body, or the body schema, and these representations are B-formatted representations. Actions are the direct result of motor representations, and motor representations necessarily involve bodily representations. Motor representations could not guide basic bodily action without a representation of the body that is acting. Because many representations of the body are encoded in B-formats, at least part of the content of motor representations must be encoded in these formats as well.

When it comes to neural redeployment, while there is significant evidence that motor representations are involved in processing in other domains (such as language processing, etc.),

[376] Kinsbourne (1995), 213.
[377] Ibid., 214.

less attention has been paid to the neural underpinnings of bodily representations being involved in other domains. It is clear, however, that bodily representations are heavily implicated in motor representations. Thus, it appears likely that where motor representations are redeployed bodily representations must be as well.

§7.2: Bodily Representations and Neural Redeployment

This theory of embodiment and massive neural redeployment allows for our actions to be generated by robustly embodied cognitive resources. This is because motor representations and the actions that these produce are themselves generated in part from basically embodied cognitions, or cognitions that are originally coded in B-formats. This non-extended embodied view therefore preserves a clear vision of agency and a robust view of the agential self. If the immediate mental antecedents to an action are embodied cognitions, then that action is embodied by the agent. And, given that the motor representations that lead to actions must involve basically embodied cognitions, these cognitive states are states that are embodied.

Under this view of embodiment, a cognition is embodied if and only if this cognitive process uses some member of a special class of codes or formats for representing and/or processing its content, viz., a body-related code or format (B-code or B-format). The properties of one's body are represented by these special systems that are dedicated to bodily representation and that encode their contents in B-formats. So, any cognition involving the representations originally or primarily used for the representation of one's body is embodied cognition. Further, these special systems, it is hypothesized, are redeployed for the purposes of representing and/or processing properties that are non-bodily. And, whatever cognitions reuse these systems are also embodied cognitions.

Part of the idea of B-format embodied cognition, therefore, is the idea of neural reuse. Neural reuse bridges the gap between basically embodied cognitions – like representations of the body coming from proprioceptive and kinesthetic feedback – and cognitions that use these representations or the systems that encode bodily representations. The theory conceives of neural reuse or redeployment as a pervasive feature of the functional organization of the brain. Neural reuse entails that we must give up the assumption that the brain is composed of sections that are function specific and domain dedicated. The idea here is that individual circuits in the brain can be put to various cognitive uses.[378] Neural circuitry associated with action and perception is most often reused, but there is certainly no confinement of reuse to these two areas. And, it is essential that the neural circuitry associated with representations of the body will also be used for the guidance of action.

Neural reuse theories suggest that neural circuits established for one purpose are commonly "exapted (exploited, recycled, redeployed) during evolution or normal development, and put to different uses, often without losing their original functions."[379] Cognitive systems, according to this hypothesis, are not decomposable into insular, modular systems. Regions of the brain are not exclusively dedicated to singular tasks. An individual brain region, or even an individual neural network, may perform a similar information-processing operation in different contexts, but this region is not dedicated to one higher-level use. Different higher-level cognitive functions are supported by many of the same neural circuits put together in different arrangements.[380] The mind is not comprised of anatomical, dissociable components that are independently modifiable. While some small neural circuits may be dissociable and

[378] Anderson, Michael (2010). Neural Reuse: A Fundamental Organizing Principle of the Brain. *Behavioral and Brain Sciences* 33, 4: 245-313. 262-263.
[379] Ibid., 246.
[380] Ibid., 248.

independently modifiable on some level, even these are not encapsulated, and neural circuits tend to span different areas of the brain, as well as there being cross-talk between these different areas.

Action execution has been one domain studied with regard to neural reuse. In particular, circuits associated with motor control functions show signs of neural reuse. There is, for instance, an entanglement of language and action. In one study, language comprehension response time was tested when subjects were asked to move in order to respond to the sentences presented to them.

> Participants are asked to judge whether a sentence makes sense or not and to respond by pressing a button, which requires a move either toward or away from their body. In one condition "yes" is away and "no" is toward; another condition reverses this. The sentences of interest describe various actions that would also require movement toward or away, as in "put a grape in your mouth," "close the drawer," or "you gave the paper to him." The main finding is of an interaction between the two conditions, such that it takes longer to respond that the sentence makes sense when the action described runs counter to the required response motion. More striking, this was true even when the sentences described abstract transfers, such as "he sold his house to you," which imply a direction without describing a directional motor action.[381]

This is called the action-sentence compatibility effect (ACE). The participants' comprehension time for the sentence presented to them was affected by the motor response that they had to make in relation to it. The interaction between these two factors implies one or more shared components between the two processes. A likely candidate for this is the neural circuit involved in motor control. So, language comprehension, even the comprehension of abstract sentences, uses motor representations. Other areas of the brain that are strongly associated with language processing are also involved in action-related tasks.

Broca's area, for instance, has long been associated with language processing, but it has now been shown to be involved in many different tasks – including movement preparation,

[381] Anderson (2010), 254.

action sequencing, action recognition, imagery of human motion, and action imitation.[382] The language system, therefore, is not an encapsulated and autonomous system. It is integrated with older systems "with a function inherited from and grounded in our experience of and capacities for perceiving and navigating the world."[383]

Circuits that support action and perception seem to be favored targets for reuse, and these have been the most extensively studied. Areas associated with perception and action are involved in non-perception, non-action domains.[384] Regions associated with both perception and action are also reused in other cognitive domains. "Regions of the brain – even fairly small regions – are typically reused in multiple domains. If a region is involved in perception tasks, action tasks, or both, it is more likely to be reused than if it is not involved in such tasks."[385] The sensorimotor system provides many examples of reuse. However, neural reuse is not limited to neural networks used for action and perception. These may be implicated in many other cognitive processes, and circuits for representing the body may also be exploited in a similar fashion insofar as they are implicated in those networks.

Some instances of embodied cognition will involve one kind of neural reuse, though it is possible that not all instances of neural reuse will be instances of embodied cognition. According to the theory under examination, cognition is embodied in so far as it involves encodings in neural circuitry in bodily formats (B-formats). B-coded representations are simply those that are encoded in bodily formats. This type of cognition represents the most basic form of cognitive

---

[382] Anderson, Michael (2007). Massive Redeployment, Exaptation, and the Functional Integration of Cognitive Systems. *Synthese* 159, 3: 329-345. 338.
[383] Anderson (2007), 14.
[384] Margaret Wilson gives evidence for the recruitment and usage of sensory and motor resources and sensorimotor coding in working memory. See Wilson, Margaret (2001). The Case for Sensorimotor Coding in Working Memory. *Psychonomic Bulletin and Review* 8, 1: 44-57.
[385] Anderson (2010), 258.

embodiment. And, these circuits can be redeployed for other cognitive uses aside from processing information about the body (such as proprioceptive information).

Action production is one form of embodied cognition, though not the only form. It may be the case that all cognitions involving the redeployment of motor representations are also embodied cognition. Embodied cognitions are therefore pervasive in the human cognitive system. Language processing, for instance, may be embodied, as it involves the motor system. What is important with respect to cognitive extension is that embodiment without extension preserves a robust view of agency. Basic embodiment, or B-format embodiment, preserves ownership over the mechanisms that generate and guide an agent's actions. This, in turn, makes the agent responsible for her actions.

Because of the reuse of neural circuits dedicated to action, many different kinds of cognitions are embodied. Because bodily representations are deeply involved in motor circuits, and because these are redeployed for the purposes of other cognitive activities, these other cognitive activities are embodied. Language processing is only one example of how motor circuitry is redeployed, other functions may heavily involve motor circuitry as well.

Goldman and others probably conceive of reuse of occurring *because* of what bodily codes do. Cognitions are embodied when they reuse B-codes *because* they have the function to represent bodily states. But suppose the reuse of B-codes isn't like this. Suppose that their reuse is more like the Tin Man's reuse of a funnel for a hat, which does not implicate its functional properties of a funnel at all. "The properties that makes the funnel a good hat…are not the properties that make it a good funnel."[386] Cognitions are embodied, on Goldman's account, whether they reuse B-codes because they represent conditions of the body, or because they do

[386] Shapiro (2014), 87.

other things totally unrelated to representation of the body. This places a burden on the embodied theorist to justify the claim that any reuse of B-codes counts as embodied cognition.[387]

There is more research to be done in order to address this objection. Further research into cognitive functioning may reveal that B-codes are in fact reused because of their bodily properties. This distinction between reuse *because* and reuse *regardless* of original representational function might also invite a thesis of massive embodiment, according to which any reuse constitutes embodiment. However, such a thesis threatens to make embodiment trivial in the cognitive domain.

§7.3: Embodied Actions: Control, Responsibility, and Ownership

Agential responsibility requires control over one's actions in some substantial way. I will call this the Control Principle. Our ordinary intuitions tell us that voluntary control is a precondition for agential responsibility. An agent must voluntarily perform an action, where the action itself (at least, the basic action) is under her control in order for her to be responsible for it. Further, cases in which an agent is physically coerced, mentally ill, hypnotized, etc. are not cases where an agent can be responsible for her actions. This is the case simply because she was not in control of the states (physical or mental) that lead to her behavior. She cannot be held responsible for an action when it was not under her control.

Much of the literature on control and responsibility is concerned with *moral* responsibility. John Martin Fischer, whose account of responsibility I take as a starting point for a more metaphysical account of responsibility, argues that moral responsibility for actions is associated with *guidance control.* Fischer defines two elements of guidance control: reasons

[387] Shapiro (2014), 88. Rupert (2009), 232-234 also references a similar problem.

sensitivity and mechanism-ownership. In order to be in control, one must be responsive to reasons, and the mechanism that issues the behavior in question must be the *agent's own* mechanism.[388] In short, this mechanism must be part of that very agent. This account of moral responsibility references a kind of control that I believe to be an essential part of an account of metaphysical agential responsibility.

Fischer (2012) argues that guidance control is the particular kind of control necessary for moral responsibility. This kind of control does not require that alternative possibilities be open to the agent, but it does require that the agent be responsible for her actions in a significant sense. Fischer agrees that responsibility requires control, but he insists that it requires control of a particular type – guidance control, as opposed to regulative control. Regulative control involves the freedom to choose otherwise, or access to alternative possibilities. Guidance control does not require freedom to choose or do otherwise. Guidance control is the freedom-relevant or control component of moral responsibility. An agent can only be held responsible for what she has guidance control over. However, an agent need not have regulative control over her behavior in order to be morally responsible for it.[389]

Guidance control of action occurs when the mechanism that issues in a choice and ensuing action is the agent's own mechanism and the agent is suitably reasons-responsive, or sensitive to reasons. This mechanism must be an agent's own in the sense that the agent can see herself as being in possession of it. An agent sees herself as being in possession of a mechanism in virtue of having certain beliefs about one's own agency and the effects that we can have in the world. This relevant constellation of beliefs must be present in order for an agent to have

[388] Fischer, John Martin (2012). *Deep Control: Essays on Free Will and Value*. New York: Oxford University Press. 11.
[389] Ibid., 6.

guidance control. It must also be the case that an agent has the *ability* to recognize reasons, and recognize certain reasons as moral reasons. Thus, in order for an agent to be morally responsible for her actions, it must be the case that those actions issue from the agent's own, moderately reasons-responsive mechanism.[390]

Fischer's own account of mechanism ownership is subjective. He states that an agent acquires control (and moral responsibility) by taking control, or by seeing herself as the source of her behavior (by having certain beliefs about her own agency). I think that the concept of guidance control will function as an account of responsibility whether we understand 'an agent's own mechanisms' in a subjective way or in a metaphysical way. A more metaphysical account, which I believe may be preferable (for some) to Fischer's own account, explains mechanism ownership in terms of what is internal or inherent to the cognitive agent. These mechanisms, in other words, are part of an agent's self-constitution.

The responsible agent, therefore, must appropriately own or posses the locus of control of an action. Mechanistic ownership in particular requires that the mechanism responsible for generating the behavior in question be spatiotemporally coincident with the agent herself – it requires that the locus of control be located exactly where and when the agent is, i.e., bounded by the agent's body. The agent herself needs to be the spatiotemporal and causal locus for the efficacious pushes and pulls that lead to her actions. These initiating and guiding factors must be part of her agency and part of her self. In the temporal moment in which a voluntary action (i.e. intentional bodily movement) occurs, the conditions for that action's initiation or generation need be tied to an essentially bounded state of affairs that is spatiotemporally coincident with and belongs to the organismic body that performs that action.

[390] Fischer (2012),187.

This account thus translates nicely into an account of metaphysical responsibility. An agent is responsible for her actions, in a metaphysical sense, if she is reasons-responsive and the mechanism that issues in her action is her own mechanism in the sense outlined above.

This account of responsibility coheres well with the Standard View of action production. Agents, according to the Standard View, are reasons responsive. Not only do they possess reasons, but they can recognize reasons as well. And, further, if our reasons states that lead to our actions are, or perhaps are generated by, an agent's own mechanisms, then agents acting in accordance with the Standard View also have the mechanism ownership necessary for these two theories to apply in concert. What matters for guidance control and responsibility is that the agent possesses her reasons states and that these reasons states are moderately reasons-responsive. Agents acting according to their reasons states can therefore be responsible according to the latter theory of guidance control and mechanistic ownership.

§7.4: Intermediate Conclusions

The B-format theory of embodiment therefore enables the agent to be truly in control of her actions. The motor representations that are the immediate antecedents to action are grounded in or perhaps partially constituted by representations of the body, or parts of the body. The mechanism that issues in action is clearly the agent's own mechanism in these cases. The representations that lead to action are embodied by the agent herself.

When actions are embodied, the mechanism that generates them is the agent's own mechanism, and the agent is (in the metaphysical sense) responsible for them. According to this theory of embodiment, the mechanism that issues in behavior is not only fully internal to the agent, it is also embodied by the agent. Motor representations include intentions that represent the action, initiate it, guide it, and monitor its effects. These representations issue in motor

commands through causing prior intentions and subsequently intentions in action, which cause bodily movements.[391] Causal efficacy is part of the content of intentions in action, which represent an action as a sequence of specific movements.[392] External context may influence our motor cognitions, but it does not constitute the cognitions that lead to our actions. Action preparation and execution are controlled by the motor system, which is internal to the agent.

This is an individualist theory of embodiment. It does, however, preserve a robustly embodied view of agency that an extended view cannot preserve. This allows for an embodied view of the mind without extended minds. The benefit of this view is that it preserves a robust view of agency and the self. No longer do we need an extended view of the mind in order to preserve an embodied view of cognition. The mind remains embodied, though not extended.

This theory of embodiment is beneficial because it draws a boundary for cognition based on informational feedback loops throughout the body. Depending on where cognitive systems are demarcated, the brain may present the boundaries of cognition, or the body may do so in virtue of being a boundary for proprioceptive and kinesthetic feedback. Whether the mind is extended to the boundaries of the skull or the boundaries of the skin, it remains the fact that the mind is internal to the agent and to a bounded self. If the mind extends throughout the body, then the agent does, as well, but the agent is still bounded by the body and therefore cannot change on a whim, include other agents' cognitive systems, proliferate unexpectedly, etc. These were the main problems with an extended theory of mind as it pertains to the self. If the mind is bounded by the brain, in which B-formatted representations are represented, the mind remains embodied. While the cognitive system and therefore the mind may seem overly restricted when kept within

[391] de Vignemont, Frederique and Haggard, Patrick (2008). Action Observation and Execution: What is Shared? *Social Neuroscience* 3, 3: 421-433. 430.
[392] Ibid., 422.

the boundaries of the skull, feedback from the body plays a unique and influential role in cognition. Even if the mind is bounded by the brain, embodiment plays a crucial role in cognition.

§7.5: Reasons and Causes

At a lower level, the motor representations that lead to one's actions are not only bounded by the agent, they are also clearly owned by the agent as they are embodied by her. Since motor intentions cause our actions, and motor intentions are both internal to and embodied by an agent, we can appropriately view these as being owned by the agent. Therefore, the actions that follow from them are properly described as the agent's own actions, and not mere happenings in the causal history of the world. Motor functions of the brain not only represent action, but also control action and lead to its execution. Motor intentions control actions insofar as they both generate bodily movements and guide these movements. The representations inherent in these are intrinsically related to action control, as the motor system employs these representations in order to control actions.[393] Thus, agents can be responsible, in a metaphysical sense, for their actions according to this theory of embodiment.

However, somewhat problematically, the neural correlates of the mental items that realize reasons for action must also be the appropriate causes of actions. As previously outlined in Chapter 1, an event is an action partly in virtue of its being appropriately caused by certain mental items (or their neural correlates). Under the Davidsonian version of the causal theory of action, reasons – understood as a combination of beliefs and pro-attitudes or desires – are causes of actions. Further, an event counts as an action, as opposed to nonactional behavior, only when

[393] Gallese, Vittorio (2000). The Inner Sense of Action: Agency and Motor Representations. *Journal of Consciousness Studies* 7, 10: 23-40. 23.

it is caused by a reason.[394] Reasons are mental states, which have neural correlates that must play a role in the causal explanation of action in order for these two theories to cohere appropriately.

The above theory of embodiment coheres well with a causal theory of action, like the one that I adopted earlier on. Such a causal view of the nature and explanation of action is typically embraced as part of a naturalistic stand on agency, where the mental items that explain and rationalize actions are in some way dependent on or realized by some physical states or events.[395] These physical states and events are, inevitably, brain or neural states. However, though the B-format theory of embodiment is compatible with a causal theory of action and, further, with a Davidsonian version of the view, the divergence of reasons for action as intentions and motor intentions seemingly creates a difficulty in the explanation of action production.

Intentions, in the more scientific sense, are sometimes described as a plan for movement that specifies the goal of that type of movement. The neural correlates of intentions do not necessarily contain information about the fine-grained details of a movement – such as the joint angles, torques, and muscle activations required to make a movement. They are coded to specify the goals of movements rather than the exact muscle activations required to execute the movement.[396] Intentions encompass some motivation for performing an action, but unlike desires they settle the matter of what you are going to do. While one can have a desire without being settled on acting on it, intending to do something is partially constituted by being settled on doing it.

Proximal motor intentions initiate actions. These intentions are intentions to do something straightaway, as opposed to intentions to do something in the nonimmediate future

[394] Mele, Alfred R. (2009). Action and Mind. In J. Symons and P. Calvo (eds.), *The Routledge Companion to Philosophy of Psychology* (pp. 609-620). New York: Routledge. 609.
[395] Ibid., 610.
[396] Ibid., 610-611.

(distal intentions). A proximal intention to A incorporates a plan for A-ing, and which action an intention generates is a partial function of this intention-embedded plan. In terms of reasons for action, however, an agent's intention in acting is constitutive of her primary reason for action. Intentions are part of reasons for action, and they involve the desirability of the intended action. They are not necessarily plans for immediate action so much as combinations representations of states of affairs or objects and features of the action that are desired.

The problem emerges that intentions are understood differently in philosophical literature on action and in cognitive scientific literature. A reason for action, according to the Davidsonian causal theory of action, is its cause. When an agent performs an action because of a reason, that reason is the cause of the action. Reasons not only explain actions and rationalize them, they also cause them. Reasons must be psychological states – beliefs and desires – as well. But, it seems as if proximal motor intentions are the cause of our actions, at a neural level. And, motor intentions do not seems to be constituted by beliefs and desires in terms of their encoded representational content. The content of beliefs and desires are likewise distinct from the content of proximal intentions. Proximal intentions represent an action plan, while reasons for action represent some feature or aspect of the action that is desired as well as beliefs about states of affairs. The contents of these mental states bear a justificatory relation to action that the contents of proximal intentions do not seems to bear. Fortunately, all of the above (intentions, beliefs, desires, and therefore reasons) are some kind of psychological state, but unfortunately they seem to be distinct states. The question then remains as to whether reasons for action can translate or be reduced to representations at the neural level and whether, once reduced, they can play a role in action production.

Davidson does specify that we want an agent's reasons to play a role in the course of practical reasoning.[397] When an agent acts for a reason, the agent takes certain considerations to settle the questions of whether to act in that way. She therein intends to act, and she executes that intention in action.[398] Reasons for action, then, settle the matter of acting in a similar manner to how intentions settle the matter of how to act. An event is an action done for reasons when the agent of the act takes certain considerations to settle the question of whether to act in that manner, therein intending to so act, and successfully execute that intention in action. Such a usage of the term 'intention' bears some similarities to the term's use in the more cognitive scientific literature.

So, reasons for action admittedly do not translate easily into intentions at the neural level. Certainly, to be the cause of an action a reason state must factor into the causal chain leading to that action's production at some point. However, reasons for action may not be the proximate link in that chain. It may be that states that correlate with beliefs and desires are involved in action production at the pre-proximal-intention period. One might also assert that the beliefs and desires that lead to our actions are constituted or realized by the cognitive processing that *generates* motor intentions and motor representations. The integration of motor intentions with other areas of the brain that support propositional cognitive processing supports the theory, which has recently been put forth, that the constituents of reasons for action – including beliefs and desires – are integrated into intentions that are involved in the production of action prior to motor intentions being fully constructed. These pre-proximal intentions cause our actions by providing motor instructions that guide behavior.

---

[397] Hieronymi, Pamela (2011). Reasons for Action. *Proceedings of the Aristotelian Society* 111: 407-427. 420.
[398] Ibid., 421.

Cross-talk between different systems in the brain shows that the motor system is involved in semantic and propositional processing. Certain types of object recognition and representation, generating action words, and mental imagery all involve activity in the ventral premotor areas of the brain.[399] Chao and Martin (2001) conclude from multiple sources of data that the representation of different object categories is distributed across many regions of the brain. Distributed networks of cortical regions are active during object processing, and the distribution of these regions varies as a function of semantic category.[400]

Martin (2007) finds that information about object properties is stored in the sensory and motor systems in the brain and is represented in sensory- and motor-based neural networks. There is evidence that there is at least partial overlap between neural systems supporting perception and action and those that support the representational storage of object properties.[401] Motor areas of the brain are active when viewing and imagining manipulable objects (such as tools), making semantic decisions about actions related to these objects, and holding these objects in working memory. Object properties are stored throughout the brain, with sensory- and motor-based information being stored in the sensory and motor systems. Beliefs thus involve these areas of the brain, as information stored is represented in distributed networks including the sensory and motor system.

It is consistent with this data both that perceptual-motor activity constitutes mental states and cognitive processing and that it merely causally induces cognitive processing to take place, and beliefs and intentions to arise. When embodied cognition theorists describe cognition as

---

[399] Martin, Alex, and Linda Chao (2001). Semantic Memory and the Brain: Structure and Processes. *Current Opinion in Neurobiology* 11: 194-201. 197.
[400] Ibid., 199.
[401] Martin, Alex (2007). The Representation of Object Concepts in the Brain. *The Annual Review of Psychology* 58: 25-45. 32.

being grounded in or involving perceptual or motor processing it is likely that they have the constitutive thesis in mind. Adams (2010b) argues against this thesis, while Shapiro (manuscript) argues that it doesn't matter which thesis is true.[402] As Shapiro states, regardless of which thesis is true, the body and motor areas contribute something important and previously unexpected to cognition. And, this is what the embodied cognition program aims to show.

Linguistically structured representations in the language processing cortex also show overlap with visual and motor regions of the brain. Representations of words are deeply associated with representations in the more embodied part of the brain, with the motor cortex being tied into these. This kind of propositional processing is also embodied in that it is linked to the embodied representations that guide actions.

Glenberg and Kaschak (2002) have also found that meaning is closely tied to action. These authors found that sentence comprehension can facilitate or interfere with one's physical responses.[403] This interaction, which I briefly mentioned previously (see Section 7.2), the authors call the *action-sentence compatibility effect* (ACE). Language, it seems, is tied to action, in that action can be inhibited or facilitated based on the meaning of a sentence presented at the time of the action. The ACE is evidence that language understanding "taps into an action-based system."[404] The action-sentence compatibility effect is found for sentences that describe concrete actions, like imperative sentences (e.g. "Close the drawer"), as well as sentences that describe more abstract actions (e.g. "Liz told you the story" or "You told Liz the story"). When subjects were asked to judge whether an imperative sentence like "Close the drawer" is sensible, they

[402] Adams, Frederick (2010b). Embodied Cognition. *Phenomenology and the Cognitive Sciences* 10, 4: 619-628.
[403] Glenberg, Arthur, and Michael Kaschak (2002). Grounding Language in Action. *Psychonomic Bulletin and Review* 9, 3: 558-565. 559.
[404] Ibid., 561.

were slower to respond if their responses required a physical motion in the opposite direction entailed by the sentence. This might be the motion of pushing a lever in response to the sentence "Open the drawer", which entails a pulling motion towards oneself. The ACE supports the idea that language understanding is grounded in bodily action.[405] These studies show that motor information is activated during sentence processing. Other studies have shown that information related to the shape and orientation of objects is activated during sentence processing.[406] The data Glenberg and Kaschak obtain support an embodied theory of meaning where the meaning of sentences is related to human action.

Neuroimaging studies have also shown that motor and premotor cortices are activated during processing of action words or sentences.[407] Processing word meanings relies on embodied representations. Boulanger et al. show in their studies, as is shown in the above studies, that processing words related to action can either interfere with or facilitate overt motor behavior depending on whether this processing is done concurrently with or prior to the motor task. When performed concurrently, processing action verbs seems to hinder the execution of a reaching movement, but when performed prior to movement onset, processing action verbs seems to assist motor behavior.[408] Interference between language and motor tasks, whereby processing action words seems to interfere with motor behavior, suggests that action word processing and retrieval engages cortical structures involved in the programming of motor actions.[409] The processing of

[405] Glenberg and Kaschak(2002), 561-562.
[406] Kaschak, M., Aveyard, M., Blanchard, A., Madden, C., Therriault, D., Yaxley, R., and Zwaan, R. (2005). Perception of Motion Affects Language Processing. *Cognition* 94: B79-B89. B80.
[407] Boulanger, Veronique, Alice Roy, Yves Paulignan, Viviane Deprez, Marc Jeannerod, and Tatjana Nazir (2006). Cross-Talk Between Language Processes and Overt Motor Behavior in the First 200 msec of Processing. *Journal of Cognitive Neuroscience* 18, 10: 1607-1615. 1607.
[408] Ibid., 1609.
[409] Ibid., 1612.

action words, therefore, involves the cortical motor regions where embodied representations reside.

Cross-talk between motor control programs and other regions of cognitive control shows that reasons for action may be intimately tied to (and perhaps partially constituted by) motor representations. These are cases in which propositional information clearly interacts with information that informs motor control. Representations associated with reasons for action may thus be causally associated with representations that inform motor control. Below, I sketch an account of how the Standard View of action in neural terms might go.

Beliefs and desires are often taken to be propositional attitudes, and as such make up much of folk psychology. However, beliefs and desires described in propositional terms are not fully naturalized, and thus do not correspond to neural activations as representations. It seems as if beliefs and desires, then, as propositional attitudes cannot be part of the physical mechanism that leads to behavior – including any cross-talk that informs that mechanism.[410] Beliefs and desires need to be construed naturalistically for this to be possible. Understanding beliefs and desires as patterns of neural activity allows this naturalistic turn.

I will presuppose that objects and states of affairs can be represented by neural populations. Paul Thagard (2008) suggests that it is plausible to conclude that neural populations represent objects by patterns of firing. There is some reason to think that neural populations represent, as neurons show highly selective responses to different object categories. "Studies reporting similar patterns of activation when subjects view and imagine objects from different object categories provide further support that the responses in these regions are driven by stored

[410] Thagard, Paul (2008). How Cognition Meets Emotion: Beliefs, Desires, and Feelings as Neural Activity. In G. Brun, U. Doguoglu, and D. Kuenzle (eds.), *Epistemology and Emotions* (pp. 167-184). Burlington, VT: Ashgate. 171-172.

object information."[411] Neurons may be tuned to the features of objects and store information about object properties. Thagard states that a neural population represents an object "when exposure to stimuli adjusts the excitatory and inhibitory connections between its neurons in such a way that the population is tuned to the object, in that its neurons exhibit a specific pattern of firing in response to the object."[412] When a specific pattern of firing in a neural population is tuned to an object, it represents that object.

Objects and concepts are typically thought of as having neural representations. But beliefs and desires can be represented by neural activations as well. Because we can construct sentences out of neural activity, it is legitimate to think that beliefs are neural activity. My belief that something is the case is a pattern of firing in my neural populations. Further, we can still hold beliefs when we aren't consciously thinking about them at a given moment through the dispositional capacities of our neural populations. This occurs when "the synaptic connections between the neurons in the relevant populations are such that they will generate the relevant firing pattern" when such a time comes that the belief is needed to be occurrent.[413] Both occurrent and dispositional beliefs can thus be fitted into this neural representational framework. Occurrent beliefs are actual patterns of neural firing, while dispositional beliefs are structures of neural connections that lead to these patterns of firing.

A neural account of desires can also be constructed within this framework. Desires, according to this model, are patterns of neural activity that connect representations of things or states of affairs with representations of some kind of reward. The things and states of affairs that are thus associated with reward representations are represented by patterns of neural activity, or

[411] Martin and Chao (2001), 196.
[412] Thagard (2008), 173.
[413] Ibid., 174.

patterns of firing neural populations or, if non-occurrent/dispositional, by neural structures that generate those patterns of neural activity. An account of the association between a given representation and a representation of a reward requires a neurocomputational theory of reward processing that describes how certain brain areas (the nucleus accumbens, amygdala, and others) attach positive and negative evaluations to representations.[414] This explanation involves positing patterns of neural activity involving brain areas that perform cognitive appraisal functions.

Cognition involves an array of different representations, some of which we can categorize as beliefs, desires, and emotive states. So, we can reduce beliefs and desires to neural representations in the brain, but how do these connect with motor representations? We have already seen that semantic processing, processing of objects, and motor processing have significant overlap. The processing of states like desires has yet to be thoroughly explored. Though less empirical work has been done on the connection between belief, pro-attitude states, and action production, there is some theory as to how these states play into the intentions that cause our actions from a neural computational point of view. Schroder et al. propose a theory of intention as a brain process binds together information. This information might be about situations, emotional evaluations, actions, or the self.[415] Human intentions are thus neurocomputational processes operating in the brain.[416] This theory coheres well with the Standard View of action in that the Standard View also supposes that intentions involve beliefs and desires.

The authors claims that:

1) Intentions are semantic pointers, which are patterns of activity in populations of spiking

[414] Thagard (2008), 175.
[415] Schroder, Tobias, Terrence Stewart, and Paul Thagard (2014). Intention, Emotion, and Action: A Neural Theory Based on Semantic Pointers. *Cognitive Science* 38: 851-880. 853.
[416] Ibid., 852.

neurons that function as compressed representations by binding together other patterns.
2) Specifically, intentions bind representations of situations, emotional evaluations of situations, the doing of actions, and sometimes the self.
3) Intentions can cause actions because of neural processes that connect semantic pointers with motor instructions.[417]

Semantic pointers, as the authors describe them, "are patterns of neural firing activity whose structure is a consequence of information compression operations implemented in neural connections."[418] Semantic pointers are representations of other representations that are compressed.

> Neural compression operations bind semantic pointers into complex symbol-like structures. Semantic pointers can be decomposed into the underlying representational structures, thereby enabling the cognitive system to control flows of information across different modalities. For understanding how intentions cause actions, the decompression operation is crucial, since it specifies how high-level symbolic representations set off the low-level motor representations that ultimately govern physical actions.[419]

Intentions are higher-level phenomena that model lower-level representations. They are a special instance of the kind of cognitive process that brings forth a representation from binding together other representations.

The representations included in these intentions will include information about the subject's external environment, representations of goal states, representations of other situations, representations of the action itself, and evaluative representations from the emotional system including positive attitudes towards a goal state. According to this view, intentions as semantic pointers can cause our actions by channeling or routing information to motor areas of the brain. This makes intentions akin to reasons for action according to the Standard View. Reasons consist of beliefs and desires, representations of which are bound together in intentions, which cause our actions.

---

[417] Schroder et al. (2014), 853-854.
[418] Ibid., 854.
[419] Ibid.

Schroder et al. note that intentions are treated differently in different disciplines. In social psychology, intentions are treated as "high-level symbolic phenomena involving planning for the future, without caring about the details of implementation in the brain."[420] In cognitive neuroscience, however, intention-related work predominantly deals with low-level cognitive processes and behaviors like the motor control of hands or fingers, adding numbers, etc. The philosophical account of intentions arguably falls between these two groups, but the rift has remained similarly in terms of intention-related work in philosophy and cognitive science. Intentions as semantic pointers, perhaps, can bridge this gap and provide a way to construct computational models that address both high-level and low-level accounts of intention.[421]

So, beliefs and desires fit neatly into the picture of action production, according to this view. Representations of situations, as well as pro and anti evaluations are part of intentions as semantic pointers. The theory fits with a causal belief-desire model of action production, as well. Neural representations of situations and evaluations are required for, and perhaps constitutive of, beliefs and desires. Intentions, according to the Standard View, are part of one's primary reason for action, which is in turn the cause of an action. These intentions also involve the desirability of an intended action, which is also a feature of the above view. Reasons for action can therefore be thought of as constituted by intentions as semantic pointers.

This is a tentative cognitive version of the standard Davidsonian causal view of action production. I will not defend the view here. I merely give it as an example of how one might give a neural account of reasons for action and connect them to action production. The above view of embodied actions and the Standard View of action production are not incompatible, and it is possible to reconcile the two in naturalistic terms.

---

[420] Schroder et al. (2014), 860.
[421] Ibid.

§7.6: A Second Objection to B-Format Embodied Cognition

This form of embodied cognition has been recognized as a form of embodied cognition that resists Radical Embodied/Enactivist Cognition. Clark, in particular, is one proponent of radical embodiment. The strategy to resist radical embodied cognition, an extended form of cognition, and radical enactivism, is to posit internal representations the format of which constitutes a special cognitive vehicle. However, radical enactivism in particular rejects the idea that basic minds are contentful, or that perception encodes contents that are then used in cognitive processing, leading to action. One potential problem for the above theory of embodied cognition is that it relies on representations to do significant explanatory work. This approach posits representations with particular vehicular properties that make them special. The special representational format is associated with a characteristic processing profile.[422] This processing may differ from other neural processing by differing in the firing rate of neurons, their firing synchrony, or their population behaviors.[423] Goldman does not specify how the exact characteristic processing profile of bodily codes and their special representational format. B-formatted representations have a specific kind of content and a distinctive format. The idea is that special neural vehicles – namely vehicles that can only be used by particular kinds of mental machinery – carry special contents in virtue of their formal features (the way in which they encode content). Hutto (2013) calls this the 'format' strategy to resist radical embodied cognition.

The Hard Problem of Content (Hutto 2013) emerges as a potential problem for this strategy. How is it possible for representations to possess their contentful properties? One must

[422] Hutto 2013, 4.
[423] Shapiro (2014), 80.

only appeal to natural phenomena in order to answer this question: causation, nomic dependencies or biological functions.[424] The problem charges that there is no scientific reason for thinking that informational or representational contents exist in nature.

> The Hard Problem of Content is a challenge to anyone hoping to supply a naturalized theory of content by standard means. It is based on the observation that co-variance relations do not—in and of themselves—constitute, entail or suffice for the existence of content. Co-varying states of affairs are utterly ubiquitous. They are not as common as muck; they are, in fact, much more common.[425]

And, because we should be aiming to produce a naturalized theory of mind, we should not posit cognitive systems as trafficking informational contents or as being otherwise content-involving. "Unless the Hard Problem of Content is addressed there is simply no reason to believe that informational contents are picked up and processed – coded and decoded – by cognitive systems."[426] This is a problem for the B-format theory of embodied cognition because B-formatted representations have as their content bodily features that are necessary for encoding the bodily goal of an overt action. This content is necessary for a theory of embodied action.

However, while some enactivist approaches to cognition are content to give up many inner representations (see Hutto and Myin 2012, 2014), extended approaches to cognition have been resiliently representational. While I will not here address the problem of naturalizing content, I will say that I think the project has made significant gains (See Anderson and Rosenberg 2008, Maloney 1994, Fodor 1990, Dretske 1988), and that even if these gains are insufficient, this does not indicate that content is not naturalizable.

Many extended mind theorists require keeping representations in their theory of mind. The basis of some accounts of extended cognition is the concept of an external representation

---

[424] Hutto (2013), 5.
[425] Hutto, Daniel and Myin, Erik (2014). Neural Representations Not Needed – No More Pleas, Please. *Phenomenology and the Cognitive Sciences* 13, 2: 241-256. 251.
[426] Ibid., 252.

and the claim that these are integral to the cognitive processing going on in the system, and without these representations these cognitive processes would either cease to exist or drastically change in nature.[427] Opposed to these are internal representations, which are contained within the cognitive agent, and cognition which is not distributed, but which are necessary for cognition to take place (see Rowlands 2009a).

Clark has argued against the anti-representationalist movement in philosophy of mind. He argues that the post-Cartesian vision of an agent as an individual who "harbors no internal representations and resists analysis in terms of any cognitively important distinctions between 'inner' and 'outer' processes, between perception, cognition and action, or between mind, body, and world" is unconvincing and empirically unsound. [428] The rejection of the notion of internal representation is associated with the idea that cognition is not an isolated enterprise comprised of the manipulation of internal symbols. However, most versions of extended cognition endorse representationalism. For most proponents of these views, the existence of mental representations and the necessity of referring to them when one tries to explain a great variety of cognitive phenomena are not at issue.[429] It seems as though extending cognition to include external representational states must not thereby eliminate internal representational states. Extended cognition, it appears, is representation-hungry, and many of its proponents are self-proclaimed representationalists.[430]

---

[427] Zhang, Jiajie and Patel, Vimla L. (2006). Distributed Cognition, Representation, and Affordance. *Pragmatics and Cognition* 14, 2: 333-341. 335.
[428] Clark, Andy (2004). Embodiment and the Philosophy of Mind. In A. Peruzzi (ed.), *Mind and Causality: Advances in Consciousness Research* (pp. 35-52). Philadelphia: John Benjamins North America. 36.
[429] Steiner, Pierre (2014). Enacting Anti-Representationalism: The Scope and Limits of Enactive Critiques of Representationalism. *Avant* 5, 2: 43- 86. 53.
[430] Clark has likewise rejected anti-representationalism in Clark, Andy and Toribio, Josefa (1994). Doing Without Representing? *Synthese* 101, 3: 401-431.

Clark and Wheeler (1999) also try to stem the tide of anti-representationalism. While the authors think that cognitive science should focus on more than what is internal to the agent and incorporate more than neural states and processes, they defend the notion that there exist internal states and processes that encode representational content. The inclusion of bodily and environmental factors in the explanation of intelligent behavior does not vitiate the need for internal representations. Even an action-centered or action-oriented approach to representation requires internal structures that describe aspects of the world and prescribe possible action. While it is difficult to define exactly what a representation is, or how it is contentful, we cannot eliminate inner representations from our explanatory ontology. Perhaps not all cognitive-scientific understanding needs to involve internal representation, but we cannot abandon the notion of internal representation and a representation-based understanding of cognition. Clark (1997) suggests that the most promising class of neuroscientific models of the mind includes representations as part of the cognitive economy.[431] The notion of inner representation still plays a key role in the extended mind theory, even if attention is shifted away from it towards the environment in which we are ecologically embedded.

Dynamical systems theory is often invoked as a nonrepresentational theory of cognition, but in reality it does not bind us to anti-representationalism. Dynamical systems approaches to cognition are free to retain the idea that cognition takes place over representations, and even that it takes place in the head.[432] Further, Clark (1997) argues that the tools of dynamical systems theory should be treated as complementary to representational understandings of the mind as opposed to replacing these. This is partially because dynamical systems theory gives us a mere

[431] Clark (1997), 130.
[432] See Shapiro, Lawrence (2013). Dynamics and Cognition. *Minds and Machines* 23, 3: 353-375.

abstract description of the mind, rather than a full understanding of it. And, this description becomes less tractable when applied to something as complex and high-dimensional as a coupled system that includes the human brain. [433]

Radical enactivism is one of the only theories that takes an explicitly eliminativist stance when it comes to contentful representations. However, enactivism and the extended mind are only superficially, if at all, similar. Arguments for the extended mind presuppose functionalism- that what is crucial to a mental state or process is its functional role, not its physical realization.[434] Enactivism is often understood to require the rejection of functionalism, so the two are at odds. Radical enactivism as espoused by Hutto and Myin (2013) and Hutto et al. (2014) rejects the extended functionalist approach of Clark, Wheeler, and others. Enactivism also does not straightforwardly lead to an extended account of mental states. While some authors (see Steiner 2015) are tempted to conflate the two groups of theories into one, this is a mistake. Radical enactivism of the variety that Hutto and Myin advocate is externalist in the sense that the wide-reaching (nonrepresentational) sensorimotor interactions that are constitutive of cognition may be contextually embedded but not all versions of enactivism are so externalistically oriented.[435] Wheeler (forthcoming) even argues that radical enactivism does not necessarily deliver externalism, as relational structures of our cognitive machinery, embodied sensorimotor capacities, and environmental affordances are all consistent with the claim that "the genuinely

[433] See Clark (1997), chapter 5.
[434] Rowlands, Mark (2009a). Enactivism and the Extended Mind. *Topoi* 28: 53-62. 56.
[435] For more on this, see Wheeler, Michael (forthcoming). The Revolution Will Not be Optimised: Radical Enactivism, Extended Functionalism and the Extensive Mind. *Topoi*: 1-16.

cognitive elements of the behavior-generating process all remain internal."[436] Even anti-representational enactivism does not secure a robust externalism.

Mainly, however strong or weak the case for representationalism is, the extended mind theorist must stand by it. It remains central to the theory that even extended cognitive processing involves the use of contentful physical structures – both internal and external. The theory routinely counts external representational elements as part of cognitive processing. Representation "provides a perfectly good basis for externalism."[437] The problem solving strategies of coupled systems involve content-rich representational structures that cannot be abandoned without a radical shift in the extended mind theory.

[436] Wheeler (forthcoming), 14.
[437] Ibid., 10.

Conclusions and Prospects for Further Research:

This dissertation has examined one novel criticism of the extended mind thesis (EMT). I hope to have shown that the theory suffers from not only a lack of attention paid to action theory, but also that this lack of attention veils an irreparable problem with the theory. Here, I hope to have not only shed light on this problem, and to have built a case against extended theories of mind and cognition, but also to have illuminated the intersection of some typically isolated philosophical positions.

All prominent extended theories of mind face the same problem – that they are incompatible with a robust view of human agency. If our mental states are extended, and the mental states that constitute our reasons for action are extended, then the causal control of our actions is also extended out into the world, away from the agent and secluded from her. If this is the case, than extended theories of mind are incompatible with agential control of our actions. These theories are thus incompatible with a core feature of ourselves as human persons. This is the agency problem for the extended mind.

This problem cannot be solved in the three ways that I address in Chapters 3, 4, and 5. Neither an internalized self, nor an extended self, nor a no-self view can solve this problem. An internalized self further enhances the problem, and it may be incompatible with the extended view itself. An extended self leads to several problems in its own right. And, a no-self view is both counterintuitive and incompatible with the extended position. I believe that these three options exhaust the possibilities for the extended mind theorist.

Further, as addressed in Chapter 5, extended mind theories face a further problem in that many versions seem to presuppose a robust agential personal profile. There is thus a conceptual

tension within the theory. On the one hand, extended theories of mind are incompatible with a robust view of agency. On the other hand, such theories presuppose that exact view.

Trying to create a bounded subject in the midst of an extended cognitive system cannot mend this problem, nor can it be fixed by extending the subject such that it encompasses the whole extended cognitive system. And, eliminating subjecthood entirely is not a viable option for the extended mind theorist, as it is both counter to our deepest held intuitions about our selves and counter to many deeply ingrained tenets of the extended mind theory.

In the end, I have offered what I believe to be a superior theory of cognition and mind that dovetails nicely with reductive-causal theories of action and agency. While such a theory requires further expansion, it provides a much-needed starting point for theorizing at the conjunction of these two philosophical topics.

Prospects for further research in this area are abundant. These include: First, further research into B-format representations and cognition involving them. Second, research into the physical realization of belief and desire states, and research into the naturalization of intentions (primarily those intentions that are not motor intentions, though further research could be done there as well, as motor intentions are the proximal causes of our actions). Third, additional attention should be paid to the further philosophical consequences of this problem for the extended mind, including ethical consequences. Fourth, further research on the self and its connection to agency may be necessary in order to strengthen the thesis given here.

To start, research into B-formatted representations and cognitions is in its early stages. Goldman and de Vignemont (2009), Goldman (2012), and Goldman (2014), have made progress in supporting the hypothesis of B-format embodied cognition, but work in this particular area is not abundant. The hypothesis of B-formatted representations, or that these special vehicles exist,

is in its more hypothetical stages. Systematicity and further research will be necessary in order to further flesh out these ideas. Determining the extent to which this theory is neurally embedded may also serve to further elucidate its philosophical implications.

More research into the realization of belief and desire states also remains to be done philosophically. More work needs to be done in order to review cognitive scientific literature on neural representations of beliefs and desires, or their constituent states, and to translate this into philosophically relevant material. While there are excellent reviews of cognitive scientific work dealing with the overlap between representations of objects, understanding of linguistic meanings, and motor representations (See Chao and Martin (1999) and (2001), Martin (2007), Boulanger (2006), Glenberg et al. (2003), Glenberg and Kaschak (2002)), there is less work illustrating the philosophically important implications of this. While Schroder et al. (2014), and Thagard (2008) go to some lengths to develop a theory of how belief- and desire- like states can be reduced to neural states (Thagard 2008) and how these neural states can constitute intentions which causally inform our actions (Schroder et al. 2014), there is little else in the literature that links beliefs and desires to neural representations and to action production.

The ethical consequences of this problem for the extended mind theory must also be recognized. Little attention has been paid to moral responsibility in conjunction with this theory of mind. However, the agency problem for the extended mind leads not only to problems for metaphysical agential responsibility, but also to problems for agential moral responsibility.

The actions of agents with extended cognitive systems seem to violate, in a deep sense, the control principle, which states that voluntary control is a necessary precondition for moral responsibility. Because, as is argued in this dissertation, the actions of extended cognitive agents are often influenced heavily by coupled environmental features that are beyond an agent's

control, agents with extended cognitive systems in many cases cannot be morally responsible for their actions. When the behavior being morally assessed is partially produced by external entities, this will in many cases annul an agent's responsibility for that behavior. This problem parallels the agency problem in many ways. Instead of metaphysical responsibility being undermined, however, moral responsibility is. If we keep the standard view of action production argued for here, and if the agency problem cannot be overcome in the metaphysical case, as I argue it cannot, then the moral problem will arise in addition to the metaphysical one. If moral responsibility is associated with voluntary control of some sort, and agents lack this control when their mental states are extended, then cognitive extension will in many cases undermine an agent's moral responsibility for her actions.

The intersection of action theory and philosophy of mind is not the only juncture that deserves further attention. The mind and its relationship to self-constitution also deserve further consideration. The mind, it seems, is inherently connected to our self-constitution, the self, and personal identity generally speaking. Thus, theories of mind must take into account their own implications for personal identity. Thus, this dissertation has additionally aimed to bring this issue to light, along with a global issue for views of extended mind and cognition.

Bibliography:

Adams, F. (2010a). Action Theory Meets Embodied Cognition. In J. Aguilar and A. Buckareff (eds.), *Causing Human Actions: New Perspectives on the Causal Theory of Action* (pp. 229-252). Cambridge, MA: MIT Press

Adams, F. (2010b). Embodied Cognition. *Phenomenology and the Cognitive Sciences* 10, 4: 619-628.

Adams, F., and Aizawa, K. (2001). The Bounds of Cognition. *Philosophical Psychology* 14: 43-64.

Adams, F., and Aizawa, K. (2008). *The Bounds of Cognition*. Malden, MA: Blackwell.

Adams, F., and Aizawa, K. Why the Mind is Still in the Head. In P. Robins and M. Aydede (eds.), *The Cambridge Handbook of Situated Cognition* (pp. 78-95). New York: Cambridge University Press (2009).

Adams, F., and Aizawa, K. (2010). Defending the Bounds of Cognition. In R. Menary, ed. *The Extended Mind* (pp. 67-80). Cambridge, MA: MIT Press.

Adams, F., and Mele, A. (1989). The Role of Intention in Intentional Action. *Canadian Journal of Philosophy* 19, 4: 511-532.

Adams, Z., and Maher, C. (2015). Cognitive Spread: Under What Conditions Does the Mind Extend Beyond the Body? *European Journal of Philosophy* 23, 3: 511-531.

Aguilar, J., and Buckareff, A. (2010). The Causal Theory of Action: Origins and Issues. In J. Aguilar and A. Buckareff (eds.), *Causing Human Actions: New Perspectives on the Causal Theory of Action* (pp. 1-26). Cambridge, MA: MIT Press.

Allen-Hermanson, S. (2013). Superdupersizing the Mind: Extended Cognition and the Persistence of Cognitive Bloat. *Philosophical Studies* 164, 3: 791-806.

Alsmith, A., and de Vignemont, F. (2012). Embodying the Mind and Representing the Body. *Review in Philosophy and Psychology* 3: 1-13.

Alverez, M. (2007). The Causalist/Anti-Causalist Debate in the Theory of Action: What It Is and Why It Matters. In A. Leist (ed.), *Action in Context* (pp.103-123). New York: de Gruyter.

Alverez, M. (2009). Reasons, Desires, and Intentional Action. In C. Sandis (ed.), *New Essays on the Explanation of Action* (pp. 203-219). New York: Pelgrave Macmillan.

Anderson, M. (2007). Massive Redeployment, Exaptation, and the Functional Integration of Cognitive Systems. *Synthese* 159, 3: 329-345.

Anderson, M. (2008). Are Interactive Specialization and Massive Redeployment Compatible? Behavioral and Brain Sciences 31, 3: 331-334.

Anderson, M. (2010). Neural Reuse: A Fundamental Organizing Principle of the Brain. *Behavioral and Brain Sciences* 33, 4: 245-313.

Anderson, M., Chemero, A., and Richardson, M. (2012). Eroding the Boundaries of Cognition: Implications of Embodiment. *Topics in Cognitive Science* 4, 4: 717-730.

Anderson, M., and Rosenberg, G. (2008). Content and Action: A Guidance Theory of Representation. *Journal of Mind and Behavior* 29: 55-86.

Bhalla, M., and Proffitt, D. (1999) Visual-Motor Recalibration in Geographical Slant Perception. *Journal of Experimental Psychology: Human Perception and Performance* 25, 4: 1076-1096.

Bishop, J. (2007). On the Importance of Reconciling Two Models of Agency. In A. Leist (ed.), *Action in Context* (pp. 154-162). New York: de Gruyter.

Bishop, J. (2010). Skepticism About Natural Agency and the Causal Theory of Action. In J. Aguilar and A. Buckareff (eds.), *Causing Human Actions: New Perspectives on the Causal Theory of Action* (pp. 69-83). Cambridge, MA: MIT Press.

Boulanger, V., Deprez, V., Jeannerod, M., Nazir, T., Paulignan, Y., Roy, A. (2006). Cross-Talk Between Language Processes and Overt Motor Behavior in the First 200 msec of Processing. *Journal of Cognitive Neuroscience* 18, 10: 1607-1615.

Buckareff, A. (2011) How Does Agent-Causal Power Work? *The Modern Schoolman: Special Issue on Free Will and Moral Responsibility* 88: 105-121.

Butler, K. (1998). *Internal Affairs: Making Room for Psychosemantic Internalism*. Dordrecht: Kluwer.

Chao, L., and Martin, A. (1999). Cortical Regions Associated with Perceiving, Naming, and Knowing About Colors. *Journal of Cognitive Neuroscience* 11, 1: 25-35.

Chao, L., and Martin, A. (2001). Semantic Memory and the Brain: Structure and Processes. *Current Opinion in Neurobiology* 11, 2: 194-201.

Chemero, A., and Silberstein, M. (2011). Dynamics, Agency, and Intentional Action. *Humana Mente* 15: 1-20.

Chiminade, T., Decety, J., and Meltzoff, A. (2013). An fMRI Study of Imitation: Action Representation and Body Schema. *Neuropsychologia* 43, 1: 115-127.

Clark, A. (1997). *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT Press.

Clark, A (1999a). An Embodied Cognitive Science? *Trends in Cognitive Science*. 3, 9: 345-351.

Clark, A. (1999b). Where Brain, Body, and World Collide. *Journal of Cognitive Systems Research* 1, 1: 5-17.

Clark, A. (2003). *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. New York: Oxford University Press.

Clark, A. (2004). Embodiment and the Philosophy of Mind. In A. Peruzzi (ed.), *Mind and Causality: Advances in Consciousness Research* (pp. 35-52). Philadelphia: John Benjamins North America.

Clark, A. (2005). Intrinsic Content, Active Memory and the Extended Mind. *Analysis* 65: 1-11.

Clark, A. (2007a). Curing the Cognitive Hiccups: A Defense of the Extended Mind. *Journal of Philosophy* 104, 4: 163-192.

Clark, A. (2007b). Reinventing Ourselves: The Plasticity of Embodiment, Sensing, and Mind. *The Journal of Philosophy and Medicine* 32, 3: 263-282.

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

Clark, A. (2008b). *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. New York: Oxford University Press.

Clark, A. (2010). Memento's Revenge: The Extended Mind, Extended. In R. Menary (ed.), *The Extended Mind* (pp. 43-66). Cambridge, MA: MIT Press.

Clark, A, and Chalmers, D. (1998). The Extended Mind. *Analysis* 58: 7-19.

Clark, A., and Drayson, Z. (forthcoming). Augmentation, Agency, and the Spreading of the Mental State. Manuscript forthcoming in *Consciousness and Cognition*.

Clark, A., and Toribio, J. (1994). Doing Without Representing? *Synthese* 101, 3: 401-431.

Clark, A., and Wheeler, M. (1999). Genic Representation: Reconciling Content and Causal Complexity. *The British Journal for the Philosophy of Science* 50, 1: 103-135.

Clark, A., and Wilson, R. (2009). How to Situate Cognition: Letting Nature Take Its Course. In M. Aydede and P. Robbins (eds.), *The Cambridge Handbook of Situated Cognition* (pp. 55-77). New York: Cambridge University Press.

Cole, J., and Paillard, J. (1995). Living with Touch and Peripheral Information about Body Position and Movement: Studies with Deafferented Subjects. In J. Bermudez, A. Marcel, and N. Eilan, (eds.), *The Body and the Self* (pp. 245-266). Cambridge, MA: MIT Press.

Davidson, D. (1963). Actions, Reasons, and Causes. *Journal of Philosophy* 60, 23: 685-700.

Davidson, D. (2001). *Essays on Actions and Events*. 2nd ed. New York: Oxford University Press.

DeMoss, D. (2011). Empty and Extended Craving: An Application of the Extended Mind Thesis to the Four Noble Truths. *Contemporary Buddhism* 12, 2: 309-325.

Dempsey, L., and Shani, I. (2013). Stressing the Flesh: A Defense of Strong Embodied Cognition. *Philosophy and Phenomenological Research* 86, 3: 590-615.

Dempsey, L., and Shani, I. (2015). Three Misconceptions Concerning Strong Embodiment. *Phenomenology and the Cognitive Sciences* 14, 4: 827-849.

Dove, G. (2011). On the Need for Embodied and Dis-Embodied Cognition. *Frontiers in Psychology* 1: 1-13.

Drayson, Z. (2010). Extended Cognition and the Metaphysics of Mind. *Cognitive Systems Research* 11, 4: 367-377.

Dretske, F. (1988). The Explanatory Role of Content. In R. Grimm and D. Merrill (eds.), *Contents of Thought* (pp. 31-61). Tucson, AZ: University of Arizona Press.

Dretske, F. (2009) What Must Actions Be Like for Reasons to Explain Them? In C. Sandis (ed.), *New Essays on the Explanation of Action* (pp. 13-21). New York: Pelgrave Macmillan.

Elpidorou, A. (2012). Where is My Mind? *Avant* 3, 1: 145-160.

Enc, B. (2004). Causal Theories of Behavior and Wayward Causal Chains. *Behavior and Philosophy* 32: 149-166.

Farkas, K. (2012). Two Versions of the Extended Mind Thesis. *Philosophia* 40, 3: 435-447.

Ferrero, L. (2013). Intention. In E. Lepore and K. Ludwig (eds.), *A Companion to Donald Davidson* (pp. 75-90). Malden, MA: Wiley.

Fischer, J. (2012). *Deep Control: Essays on Free Will and Value*. New York: Oxford University Press.

Fodor, J. (1990). *A Theory of Content and Other Essays*. Cambridge, MA: MIT Press.

Gallese, V. (2000) The Inner Sense of Action: Agency and Motor Representations. *Journal of Consciousness Studies* 7, 10: 23-40.

Gertler, B. (2007). Overextending the Mind? In B. Gertler and L. Shapiro (eds.), *Arguing about the Mind* (192-206). New York: Routledge.

Giles, J. (1993). The No-Self Theory: Hume, Buddhism, and Personal Identity. *Philosophy East and West* 43, 2: 175-200.

Glenberg, A., Cowart, M., and Kaschak, M. (2001). An Affordance Field for Guiding Movement and Cognition. *Behavioral and Brain Sciences* 24, 1: 43-44.

Glenberg, A., and Kaschak, M. (2002). Grounding Language in Action. *Psychonomic Bulletin and Review* 9, 3: 558-565.

Glenberg, A., Kaschak, M., Malter, A., and Robertson, D. (2003). Embodied Meaning and Negative Priming. *Behavior and Brain Sciences* 26, 5: 644-647.

Goldenberg, G. (1995). Imitating Gestures and Manipulating a Mannikin – The Representation of the Human Body in Ideomotor Apraxia. *Neuropsychologia*, 33, 1: 63-72.

Goldman, A. (2012). A Moderate Approach to Embodied Cognitive Science. *Review of Philosophy and Psychology* 3, 1: 71-88.

Goldman, A. (2014). The Bodily Formats Approach to Embodied Cognition. In U. Kriegel (ed.), *Current Controversies in Philosophy of Mind* (pp. 91-108). New York: Routledge

Goldman, A. and de Vignemont, F. (2009). Is Social Cognition Embodied? *Trends in Cognitive Sciences* 13, 4: 154-159.

Greenwood, J. (2015). Is the Mind Extended or Scaffolded? Ruminations on Sterelny's (2010) Extended Stomach. *Phenomenology and the Cognitive Sciences* 14, 3: 629-650.

Gustafson, D. (2007). Neurosciences of Action and Noncausal Theories. *Philosophical Psychology* 20, 3: 367-374.

Hacker, P. (2007). Thought and Action. In A. Leist, ed. *Action in Context* (pp. 53-74). New York: de Gruyter.

Hieronymi, P. (2011). Reasons for Action. *Proceedings of the Aristotelian Society* 111: 407-427.

Hornsby, J. (1997). *Simple Mindedness: In Defense of Naïve Naturalism in the Philosophy of Mind*. Cambridge, MA: Harvard University Press.

Hurley, S. (1998). *Consciousness in Action*. Cambridge, MA: Harvard University Press.

Hurley, S (2001a). Overintellectualizing the Mind. *Philosophy and Phenomenological Research* 63, 2: 423-431.

Hurley, S. (2001b). Perception and Action: Alternate Views. *Synthese* 129: 3-40.

Hurley, S (2003). Action, the Unity of Consciousness, and Vehicle Externalism. In A. Cleeremans (ed.), *The Unity of Consciousness* (pp. 72-91). New York: Oxford University Press.

Hurley, S. (2008). The Shared Circuits Model (SCM): How Control, Mirroring, and Simulation Can Enable Imitation, Deliberation, and Mindreading. *Behavioral and Brain Sciences* 31: 1-58.

Hutto, D. (2013). Exorcizing Action Oriented Representations: Ridding Cognitive Science of its Nazgul. *Adaptive Behavior* 21, 3: 142-150.

Hutto, D., Kirchhoff, M., and Myin, E. (2014). Extensive Enactivism: Why Keep It At All?. *Frontiers in Human Neuroscience* 8: 1-11.

Hutto, D., and Myin, E. (2012). *Radicalizing Enactivism: Basic Minds Without Content*. Cambridge, MA: MIT Press.

Hutto, D., and Myin, E. (2014). Neural Representations Not Needed – No More Pleas, Please. *Phenomenology and the Cognitive Sciences* 13, 2: 241-256.

Jeannerod, M. (1997). *The Cognitive Neuroscience of Action*. Cambridge, MA: Blackwell.

Kaschak, M., Aveyard, M., Blanchard, A., Madden, C., Therriault, D., Yaxley, R., and Zwaan, R. (2005). Perception of Motion Affects Language Processing. *Cognition* 94: B79-B89.

Kinsborne, M. (1995). Awareness of One's Own Body: An Attentional Theory of Its Nature, Development, and Brain Basis. In J. Bermudez, A. Marcel, and N. Eilan (eds.), *The Body and the Self* (pp. 205-224). Cambridge, MA: MIT Press.

Leist, A. (2007). Cognition and Action. In A. Leist, ed. *Action in Context* (pp. 315-343). New York: de Gruyter.

Lowe, E. J. (2009). Free Agency, Causation and Action Explanation. In C. Sandis (ed.), *New Essays on the Explanation of Action* (pp. 338-355). New York: Pelgrave Macmillan.

Ludwig, D. (2015). Extended Cognition and the Explosion of Knowledge. *Philosophical Psychology* 28, 3: 355-368.

Mackenzie, M. (2011). Enacting the Self: Buddhist and Enactivist Approaches. In M. Siderits, E. Thompson, and D. Zahavi (eds.), *Self, No Self?: Perspectives from an Analytical, Phenomenological, and Indian Traditions* (pp. 239-273). New York: Oxford University Press.

Madary, M. (2012). Showtime at the Cartesian Theater? Vehicle Externalism and Dynamical Systems. In F. Paglieri, ed. *Consciousness in Interaction: The Role of the Natural and Social Context in Shaping Consciousness* (pp. 59-72). Philadelphia, PA: John Benjamins.

Mandik, P. (2005). Action-Oriented Representation. In A. Brook and K. Akins, eds. *Cognition and the Brain* (pp. 284-308). New York: Cambridge University Press.

Maloney, J. C. (1994). Content: Covariation, Control, and Contingency. *Synthese* 100, 2: 241-290.

Martin, A. (2007). The Representation of Object Concepts in the Brain. *The Annual Review of Psychology* 58: 25-45.

Martin, A., and Chao, L. (2001). Semantic Memory and the Brain: Structure and Processes. *Current Opinion in Neurobiology* 11: 194-201.

Mele, A. (2009). Action and Mind. In J. Symons and P. Calvo (eds.), *The Routledge Companion to Philosophy of Psychology* (pp. 609-620). New York: Routledge.

Menary, R. (2007). *Cognitive Integration: Mind and Cognition Unbounded*. New York: Pelgrave Macmillan.

Metzinger, T. (2011). "The No-Self Alternative". In S. Gallagher (ed.), *The Oxford Handbook of the Self* (pp. 279-296.). Oxford, UK: Oxford University Press.

O'Brien, G. (1998). The Mind: Embodied, Embedded, but not Extended. *Metascience* 7: 78-83.

Olson, E. (2004). Animalism and the Corpse Problem. *Australasian Journal of Philosophy* 82, 2: 265-274.

Olson, E. (2011). The Extended Self. *Minds and Machines* 21, 4: 481-495.

O'Shaughnessy, B. (1995). Proprioception and the Body Image. In J. Bermudez, A. Marcel, and N. Eilan (eds.), *The Body and the Self* (pp. 175-204). Cambridge, MA: MIT Press.

Prinz, J. (2009). Is Consciousness Embodied? In P. Robbins and M. Ayede (eds.), *The Cambridge Handbook of Situated Cognition* (pp. 419-436). New York: Cambridge University Press.

Rowlands, M. (2009a). Enactivism and the Extended Mind. *Topoi* 28: 53-62.

Rowlands, M. (2009b). Extended Cognition and the Mark of the Cognitive. *Philosophical Psychology* 22, 1: 1-19.

Ruben, D. (2009). Con-Reasons as Causes. In C. Sandis (ed.), *New Essays on the Explanation of Action* (pp. 62-74). New York: Pelgrave Macmillan.

Rundle, B. (1997). *Mind in Action*. New York: Oxford University Press.

Rupert, R. (2004). Challenges to the Hypothesis of Extended Cognition. *Journal of Philosophy* 101, 8: 389-428.

Rupert, R. (2009). *Cognitive Systems and the Extended Mind*. New York: Oxford University Press.

Rupert, R. (2011). Embodiment, Consciousness, and the Massively Representational Mind. *Philosophical Topics* 39, 1: 99-120.

Santos, F., and Sia, S. (2007). *Personal Identity, the Self, and Ethics*. New York: Pelgrave Macmillan.

Schlosser, M. (2007). Basic Deviance Reconsidered. *Analysis* 67, 3: 186-194.

Schlosser, M. (2008). Agent-Causation and Agential Control. *Philosophical Explorations* 11, 1: 3-21.

Schlosser, M. (2010). Bending it Like Beckham: Movement, Control, and Deviant Causal Chains. *Analysis* 70, 2: 299-303.

Schroder, T., Stewart, T., and Thagard, P. (2014). Intention, Emotion, and Action: A Neural Theory Based on Semantic Pointers. *Cognitive Science* 38: 851-880.

Searl, J. (1979). The Intentionality of Intention and Action. *Inquiry* 22: 253-280.

Shapiro, L. (2004b). *The Mind Incarnate*. Cambridge, MA: MIT Press.

Shapiro, L. (2011). *Embodied Cognition*. New York: Routledge.

Shapiro, L. (2013). Dynamics and Cognition. *Minds and Machines* 23, 3: 353-375.

Shapiro , L. (2014). When is Cognition Embodied? In U. Kriegel (ed.), *Current Controversies in Philosophy of Mind* (pp. 73-90). New York: Routledge.

Shapiro, L. Matters of the Flesh: The Role(s) of the Body in Cognition. Unpublished manuscript. March 22, 2016. Microsoft Word File.

Shenton, J., Coslett, H., and Schwoebel, J. (2004). Mental Motor Imagery and the Body Schema: Evidence for Proprioceptive Dominance. *Neuroscience Letters* 370: 19-24.

Spackman, J., and Yanchar, S. (2013). Embodied Cognition, Representationalism, and Mechanism: A Review and Analysis. *Journal for the Theory of Social Behavior* 44, 1: 46-79.

Sprevak, M. (2009). Extended Cognition and Functionalism. *Journal of Philosophy* 106, 9: 503-527.

Steiner, P. (2014). Enacting Anti-Representationalism: The Scope and Limits of Enactive Critiques of Representationalism. *Avant* 5, 2: 43- 86.

Steiner, P. (2015). A Problem for Representationalist Versions of Extended Cognition. *Philosophical Psychology* 28, 2: 184-202.

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

Stout, R. (2007). Two Ways to Understand Causality in Agency. In A. Leist (ed.), *Action in Context* (pp. 137-153). New York: de Gruyter.

Stout, R. (2010). Deviant Causal Chains. In T. O'Connor and C. Sandis (eds.), *A Companion to the Philosophy of Action*. New York: Blackwell.

Strawson, G. (1999). The Self. In S. Gallagher and J. Shear (eds.), *Models of the Self* (pp. 1-24). Exeter, UK: Imprint Academic.

Tani, J. (1999). An Interpretation of the 'Self' from the Dynamical Systems Perspective: A Constructivist Approach. In S. Gallagher and J. Shear (eds.), *Models of the Self* (pp. 149-176). Exeter: Imprint Academic (1999).

Thagard, P. (2008). How Cognition Meets Emotion: Beliefs, Desires, and Feelings as Neural Activity. In G. Brun, U. Doguoglu, and D. Kuenzle (eds.), *Epistemology and Emotions* (pp. 167-184). Burlington, VT: Ashgate.

Thompson, E. and Varella, F. (2001). Radical Embodiment: Neural Dynamics and Consciousness. *Trends in Cognitive Sciences* 5, 10: 418-425.

Valaris, M. (2015). The Instrumental Structure of Actions. *The Philosophical Quarterly* 65, 258: 64-83.

Van de Laar, T. (2006). Dynamical Systems Theory as an Approach to Mental Causation. *The Journal for General Philosophy of Science* 37: 307-332.

de Vignemont, F. (2010). Body Schema and Body Image – Pros and Cons. *Neuropsychologia* 48, 3: 669-680.

de Vignemont, F. (2011). Embodiment, Ownership and Disownership. *Consciousness and Cognition* 20, 1: 82-93.

de Vignemont, F. and Haggard, P. (2008). Action Observation and Execution: What is Shared? *Social Neuroscience* 3, 3: 421-433.

Wheeler, M. (2005). *Reconstructing the Cognitive World*. Cambridge, MA: MIT Press.

Wheeler, M. (forthcoming). The Revolution Will Not be Optimised: Radical Enactivism, Extended Functionalism and the Extensive Mind. *Topoi*: 1-16.

Wilson, M. (2001). The Case for Sensorimotor Coding in Working Memory. *Psychonomic Bulletin and Review* 8, 1: 44-57.

Wilson, M. (2002). Six Views of Embodied Cognition. *Psychonomic Bulletin and Review* 9, 4: 625-636.

Wilson, R (2000). The Mind Beyond Itself. In D. Sperber (ed.), *Metarepresentations: A Multidisciplinary Perspective* (pp. 31-52). New York: Oxford University Press.

Wilson, R. (2004). *Boundaries of the Mind: The Individual in the Fragile Sciences*. New York: Cambridge University Press.

Wilson, R. (2014). Ten Questions Concerning Extended Cognition. *Philosophical Psychology*, 27, 1: 1-15.

Zhang, J. and Patel, V. (2006). Distributed Cognition, Representation, and Affordance. *Pragmatics and Cognition* 14, 2: 333-341.