

THE EXTENDED CONSCIOUS MIND

by

Michael George Bruno

Copyright © Michael George Bruno 2013

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF PHILOSOPHY

**In Partial Fulfillment of the Requirements
For the Degree of**

DOCTOR OF PHILOSOPHY

In the Graduate College

UNIVERSITY OF ARIZONA

2013

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Michael George Bruno entitled *The Extended Conscious Mind* and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

_____ Date: 9/28/2013
David Chalmers

_____ Date: 9/28/2013
Terence Horgan

_____ Date: 9/28/2013
Uriah Kriegel

_____ Date: 9/28/2013
Shaun Nichols

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

_____ Date: 9/28/2013
Dissertation Director: David Chalmers

_____ Date: 9/28/2013
Dissertation Director: Terence Horgan

STATEMENT BY THE AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

SIGNED: Michael George Bruno

TABLE OF CONTENTS

| | |
|--|-----|
| ABSTRACT | 5 |
| 1. INTRODUCTION | 6 |
| 2. ACTIVE VEHICLE EXTERNALISM | 18 |
| 2.1 Externalism Generally | 23 |
| 2.2 Factive states and de re contents | 26 |
| 2.3 Twin Earth | 29 |
| 2.4 Direct Realism, Wide Intentionalism, and Inverted Earth..... | 33 |
| 2.5 Vehicles..... | 41 |
| 2.6 EM | 47 |
| 2.7 ECM..... | 55 |
| 2.8 Ecological theories of perception..... | 59 |
| 2.9 Enactivism..... | 65 |
| 2.10 Dynamic sensorimotor theories..... | 68 |
| 2.11 Two-level interdependence view..... | 72 |
| 3. OBJECTIONS AND GROUNDING | 75 |
| 3.1 Extended Minds, Extended Selves? | 75 |
| 3.2 Demarcating the Mental..... | 78 |
| 3.3 Mistaking Evidence of Coupling for Evidence of Constitution | 79 |
| 3.4 Locating Constitutively Grounded Events..... | 86 |
| 4. TEMPORALLY EXTENDED VISION..... | 93 |
| 4.1 The dynamic content problem for cinematic models..... | 96 |
| 4.2 The extensionalist and retentionalist alternatives | 101 |
| 4.3 Why the vehicles of visual experience are never instantaneous events | 103 |
| 4.4 Enactivism and retentionalism | 107 |
| 4.5 Why visual experiences fail to supervene on instantaneous events..... | 114 |
| 5. EXTENDED VISUAL CONSCIOUSNESS | 117 |
| 5.1 An Argument for EVC: Temporal Extension plus Dynamical Entanglement | 121 |
| 5.2 Why Bodily Duplicates Decoupled from Any Environment Are Insufficient to Constitutively Ground Most Pedestrian Visual Experiences | 130 |
| 5.3 Conclusion | 134 |
| REFERENCES..... | 136 |

ABSTRACT

Do minds ever extend spatially beyond the boundaries of the bodies of their subjects? I argue that they do. More precisely, I argue that some of our visual experiences are constitutively grounded by events that include parts of the world that are not parts of any subject's body. After surveying the development of externalist theories in the philosophy of mind, I present some of the motivations common to ecological, enactive, dynamic sensorimotor and two-level interdependence accounts of perception and explain how some of these accounts support the case for active vehicle externalism about consciousness. I then discuss and respond to three well-known objections. The first concerns whether the extended mind thesis implies that there extended selves, the second concerns what exactly demarcates mental events from non-mental events, and the last concerns what is required to demonstrate constitutive dependence. To address what distinguishes constitutive from nomological or causal forms of dependence, I develop an account of constitutive grounding. My account draws on recent work in analytic metaphysics on the notion of ontological dependence or grounding, where grounding is taken to be a non-causal relation of ontological priority. After showing how this notion is different than any kind of nomological dependence and how it can be constructively used to decipher the spatiotemporal extent of events, I argue positively that the grounds of visual experiences are always temporally extended and often include parts of the world external to the seeing subject's body. My argument for temporally extended vision begins by considering three different models of the temporal structure of consciousness: cinematic, retentional, and extensional. I then draw on the dynamic sensorimotor theory to object to the cinematic model and explore whether enactivists are really committed to retentionalism. I end up arguing that any account one gives of the intentional contents or phenomenal characters of individual conscious visual events will have to make reference to a briefly enduring process and not just an instantaneous event involving the subject. Lastly, I argue as follows: (P1) in the explanation of visual experience, the brain internal parts of the temporally extended events that constitutively ground visual experiences often cannot be decoupled from parts of the non-bodily world; (P2) if event A is a constitutive ground of event E and event B cannot be decoupled from A in the explanation of E, then B is also a constitutive ground of E; therefore, (C) some visual experiences are constitutively grounded by events that include parts of the non-bodily world. I call this conclusion the extended visual consciousness thesis. If my argument for it is sound, our conscious minds do, in some cases, extend beyond our bodies.

1. INTRODUCTION

Where is the interface between a mind and the rest of the world? Although today most would point to the boundaries of a subject's body or brain, I will argue that some conscious states and processes literally extend spatially beyond the subject's body. If this is right, then while a subject's mental life may always be centered on that subject's brain, his or her mind also sometimes includes parts of the non-bodily world. The vehicles of some of his or her conscious states and processes, in other words, include parts of that subject's surrounding environment. Following Clark (2009), I will call this view the extended conscious mind thesis, or ECM. My main task in what follows will be to defend its plausibility. Clark (2009, 967) characterizes ECM and distinguishes it from the extended mind thesis, EM, which he elsewhere defends, in the following way:

[I]n the case of ECM, just as in the case of EM, what is at issue is the location of the material vehicles of certain mental or cognitive states. But whereas EM was concerned only with the vehicles of non-conscious mental states such as states of dispositional believing, ECM makes the even more striking claim that the local material vehicles of some of our conscious experiences might include more than the whirrings and grindings of the brain/CNS.

Florid mechanistic language aside, I quibble with Clark's use of the epistemic modal 'might' in characterizing ECM. I interpret ECM as a claim about what the vehicles of some conscious states *do* include, not simply what they might include for all one knows.

Consider also Ward's (2012, 2) recent characterization of ECM in defending it against some of Clark's (2009) objections:

[T]he subpersonal basis of some conscious experiences can include events and properties outside the organismic boundaries of the experiencing subject.

Again, while this is close to what I have in mind, I think we should replace the modal ‘can’ with a ‘do’. ECM is not merely about what the vehicles of mental states and processes *can* include; one cannot establish it by showing that extended consciousness is possible (e.g., in intelligent aliens or future machines with strong AI). The issue is not directly about whether mental properties are multiply realizable; instead, it is about the minimally sufficient conditions for actually realizing particular conscious experiences on particular occasions. Sometimes a minimally sufficient basis includes events and properties outside the boundaries of the experiencing subject’s body. That, at least, is what I plan to argue.

ECM is a species of the more general extended mind thesis, or EM. Roughly, EM is the thesis that the vehicles that ground some mental states and processes spatiotemporally extend beyond the boundaries of the body of their subject. Since EM is compatible with the claim that only unconscious mental states and processes ever extend beyond the body, it does not entail ECM. Both EM and ECM are versions of externalism, which means they maintain that some mental properties are relationally determined. In particular, the relevant properties are determined by relations between a subject’s body and the surrounding world and thus are not intrinsic to what happens just within the subject’s body.

There is something strange about saying that states and processes are spatially extended at all, mental or otherwise. This is because states and processes are properties and properties, on most accounts, are abstract. Things that instantiate properties, e.g. subjects, objects, systems, and events, are concrete and spatially extended; but not, if properties are by nature abstract, their token states and processes. This way of framing

things raises some foundational metaphysical questions about the nature of properties. But questions about whether properties are immanent or transcendent universals, bundles of tropes, sets of cross-world objects, etc. need not concern us. Instead, our focus will be on whether claims about the location and extension of particular instantiations of properties make sense. I think they do, insofar as the thing instantiating the property is something physical and concrete.

To simplify matters, we can usually confine our discussion to events and relations among different events. Events clearly have mereological and spatiotemporal structure – for example, the battle of Gettysburg occurred prior to and in a different place than the battle of Normandy. Thus, events are well-suited to framing debates about the locations of property instantiations.

There are two general questions to consider when locating property instantiations or events. First, to what extent does the type of property instantiated constrain the location of the instantiation? Second to what extent does the type of entity instantiating the property constrain the location of the instantiation? The first question raises its own set of issues when dealing with mental properties, which I discuss momentarily. With respect to the second, while the location of a property instantiation is constrained by the location of the thing instantiating it, it is not the case that the instantiation always coincides precisely with the extent of the entity instantiating that property. An apple can be red or spherical even though its core is neither. A subject can have a fever and a bruised toe on some occasion and though the fever and the bruise may spatially and temporally overlap, only the fever spatially extends throughout the body. The bruise is presumably located in a region

occupied by just one of the subject's toes. Closer to the topic at hand, if some types of mental states and processes only occur when there is characteristic activity in a subject's pre-frontal cortex or thalamus and others occur even in the absence of such activity, it may be that states and processes of the first type always spatially extend into brain regions where states and processes of the second type do not. But we say the apple itself is red and spherical and the person herself is bruised. In these cases, we might say the location of the vehicle is less extensive than the location of the instantiator on the occasion of instantiation. But it can go the other way too. If I beat someone in a game of backgammon, then I instantiate the property of winning, but the minimal extent of my win includes regions occupied by the board, my opponent, etc. In other words, the location of the vehicle of an instantiation can be more or less extensive than the location of the instantiator.

Concerning the first question, it can seem perplexing to claim anything at all about the spatiotemporal extension of minds, mental properties, events, etc. Mental processes such as acts of thinking, desiring, perceiving, feeling, remembering, and so forth clearly last or endure for some time and hence are temporally extended. But mental states such as content-individuated thoughts, desires, perceptions, feelings, memories, etc. are often conceptualized synchronically, i.e., as states of a subject at an instant. Even for mental states that do endure, unlike processes, they persist *without change*. States of a subject at a time are often individuated in terms of their contents or what they are about, and (intentional or representational) contents are generally taken to be abstract, hence without location, regional boundaries, and spatial extent. Moreover, logically speaking, even temporally extended mental processes may not always be spatially extended. Perhaps instantaneous

vehicles carry dynamic contents; if so, change at the level of experience does not straightforwardly constitutively implicate an event with temporal extension.

Questions about the extension or location of property instantiations become especially difficult when dealing with phenomenally conscious properties or phenomenal characters. Conscious states and processes are those there is something it is like for the subject to be in or undergo. We can individuate or type-identify conscious states and processes in terms of their phenomenal characters, e.g. experiences of seeing red, feeling cold, hearing a B flat, etc. But phenomenal characters are often thought to be essentially independent of anything spatiotemporal. Articulating this perspective, McGinn (1995) writes that experience “is not located at any specific place; it takes up no particular volume of space; it has no shape; it is not made up of spatially distributed parts; it has no spatial dimensionality; it is not solid.” If he is correct, this raises a problem for understanding how conscious mental states and processes could ever be spatially extended at all. It would follow that with respect to any conscious experience that “to ask for its spatial properties is to commit some sort of category mistake, analogous to asking for the spatial properties of numbers.” Of course, this points toward a familiar argument for dualism. If phenomenal characters are essentially non-spatial, but material or physical things like bodies, brains, trees, and tables always have volume, shape, extension, and location, they must be fundamentally distinct and different in kind. Provided we countenance both as real, this rules out any comprehensive materialistic/physicalistic or idealistic theory. So, if McGinn (and Descartes) are right, it must be a mistake to assign locations or specifications of spatial extension to conscious mental properties, as ECM requires.

But to be extended in the relevant sense, a property instantiation need not itself have shape or volume. If phenomenal characters are literally type-identical to physical properties perhaps they would, but this is not what ECM requires. Rather, all that is necessary is that particular instantiations are grounded by or ontologically dependent upon something that is spatially extended, or by something which will eventually be grounded or dependent upon something that is. This could simply be the region the subject occupies on the occasion of the instantiation. Provided we are dealing with actual conscious states and processes, individual tokens are always located somewhere in space and time. The general point here is that it is only the vehicle or realizing state or process that needs to have shape, volume, and extension but not phenomenal characters themselves. Analogously, while it may be a category mistake to ask about the spatial properties of the number two, it is entirely legitimate to ask about the spatial properties of any collection of two objects.

Still, reconciling the underlying psychophysical issues that arise about the relation between phenomenal consciousness and spatiotemporal extension becomes more promising when dealing with versions of dualism, which, unlike Descartes' and McGinn's, do not rely on the putative non-spatiality of the phenomenal. This includes most contemporary versions of property dualism, e.g. Chalmers (1996), and perhaps even some non-Cartesian versions of substance dualism, e.g. Lowe (1996); nevertheless, complications remain. I will address some of the issues surrounding the mind-body problem as they pertain to locating instantiations of conscious mental properties over the next two chapters.¹

The issues that arise for advocates of ECM are basically the same as those for anyone

¹ Succinctly put, the underlying issue concerns whether phenomenal consciousness can be accounted for in terms of something else or, instead, must be regarded as irreducible, hence as emergent and itself fundamental.

willing to argue that conscious states and processes have extension within parts of the brain or body. Appeal to extended minds does not resolve these issues, but neither does it add unnecessary complications to them.²

In the next chapter, I will discuss different versions of externalism at some length, including the relationship between EM and ECM. While externalist theses have been popular in the philosophy of mind for several decades, explicit discussion of EM is a more recent development. Most of the current debate can be traced to Clark and Chalmers' (1998) influential and programmatic defense of EM, or what they call 'active externalism', which has provoked a great deal of criticism. Clark and Chalmers both, together and independently, explicitly reject ECM. However, a number of influential theories of perception seem to entail that consciousness sometimes extends beyond the body. For instance, the ecological theory of Gibson (1966, 1979/1986), the enactive theory of Varela and Thompson (2001), the dynamic sensorimotor theories of O'Regan (2011) and Noë (2004, 2009), the two-level interdependence theory of Hurley (1998, 2010), and disjunctivist direct realist theories of perceptual experience have all been interpreted as entailing ECM. In each case, however, whether or not the core commitments of the theory actually commit it to ECM is somewhat controversial. My positive argument for ECM draws on and refines arguments and evidence advanced in support of these theories, but without adopting any of them wholesale.

² If one accepts even the most non-reductive versions of functionalism about phenomenal consciousness, then one ought to concede the possibility of extended conscious minds. Questions of vehicles, of what realizes or grounds a token conscious state or process (of a conscious subject or creature) are always questions about what is included in the subpersonal basis of actual, personal-level, conscious thoughts, feelings, and experiences. Even emergent property instantiations can be, and in the case of conscious experiences are, constitutively grounded in regions of space and time.

In the third chapter, I present some of the more pressing objections to EM and ECM that have emerged. These objections pertain to the possibility of spatially extended selves, demarcating the mental, and whether it is always fallacious to infer constitution from coupling. The most important objection is that the arguments put forth in their favor only succeed in demonstrating causal or nomological dependencies between a subject's mental properties and the surrounding environment. The theses themselves, however, require constitutive or metaphysically necessary dependence. Without constitutive dependence on the non-bodily world, the location claims central to EM and ECM do not follow. Unfortunately, it is difficult to say just what is sufficient to demonstrate constitutive dependence in all cases. If an account cannot be provided, however, the whole issue may begin to look terminological, vacuous, and without purpose or point. In order to better appreciate the distinction between causal and constitutive relations, I will explicate a notion of *constitutive grounding* and show how it can be constructively applied to defending various theses about the spatiotemporal extension of human visual experience.

Shapiro (2011, 160) writes that the debate over EM “turns on whether the body and world are *important* or *central* constituents in cognitive processes, such that cognition would break down, or be incomplete, or be something other than what it is, without their constituency.” The necessity of the constituents (grounds) to the constituted (grounded) must thus be stronger than nomological. A ground may not need to be metaphysically necessary for all instantiations of a particular type, but it is metaphysically necessary for a particular realization or instantiation or tokening on an occasion; hence it is part of *what* that particular instantiation is. My account of the causal/constitutive distinction is inspired

by recent attempts to vindicate a neo-Aristotelian notion of ontological dependence within contemporary analytic metaphysics.³ This notion is often just called ‘grounding’, which is why I call my notion ‘constitutive grounding’. Grounding is a relation roughly captured by explanatory expressions like ‘in virtue of’ and ‘because’ in ordinary language. Grounds exist prior to and are more fundamental than what they ground. Since grounds are more fundamental than what they ground, we can say that what’s grounded obtains or occurs because of its grounds, i.e. in virtue of those grounds occurring or obtaining.

Drawing on Clark and Chalmers’ (1998) characterization of EM as ‘active externalism’ it is possible to distinguish EM from other forms of externalism in terms of constitutive grounding, without making explicit reference to vehicles. Externalism generally is the thesis that some mental events are constitutively grounded by events that include parts of the environment surrounding a subject that are not parts of that subject’s body. EM is the thesis that some mental events are constitutively grounded by events that include parts of the environment outside a subject’s body and do so in virtue of those parts being actively coupled in the right way to parts of that subject’s brain.

Externalism in terms of grounding generally implies that there are some essential explanatory determinates in the world, while EM places a further constraint on what those worldly determinates are. This constraint also helps clarify why all versions of EM are relevant to the location of the instantiation, occurrence, or tokening of specific mental properties, events, states, and processes in a way more traditional versions of externalism are not. Events within a subject’s brain and events that occur outside that subject’s body are

³ See, for example, Fine (1994), Schaffer (2009), Lowe (2010), and essays in Correia and Schneider (2012).

actively coupled in the right sort of way when they are interdependent and inseparable with respect to some token-explanatory end. The activeness of a coupling, more generally, can be understood in terms of Simon's (1969) notion that components of a system can be distinguished in terms of intensity of interaction. Roughly speaking, high-bandwidth interactions tend to occur within components and while low-bandwidth interactions occur between different components. If this is right, the coupling relations within components will always be more active than coupling relations between components.⁴ More generally, high-bandwidth information flow can be taken as an indicator of the extent to which some coupling relation is active.

The non-bodily determinates of content according to traditional content externalists include events that involve evolutionary history, lengthy learning processes, diffuse social practices, etc. While these kinds of non-bodily grounds may also be functionally tuned for action in many cases, they generally do not stand in high-bandwidth interactive relations to subjects' brains and are not actively coupled to them. So while the external determinates of content for traditional externalists are passive in that they "play no role in driving the cognitive process in the here and now" (Clark 2008, 79), the requirement of active coupling

⁴ Picking up on this idea, Haugeland (1998, 213) has defined 'interface' as "a point of interactive 'contact' between components such that the relevant interactions are well-defined, reliable, and relatively simple." By 'relatively simple', Haugeland means low-bandwidth. Haugeland applies this bandwidth criterion to challenge the idea that there are any important interfaces between an acting agent's body and the outer world. Instead, he argues, we find an "intimate intermingling of mind, body and world" (Haugeland 1998, 224). Clark (2008, 32) objects that there are still well-defined interfaces between computers engaged in high-bandwidth interactions with each other as a part of a computational grid. Moreover, Velliste et al. (2008) have shown that motor command signals from 20-60 neurons in the primary motor cortex are sufficient to produce smooth and successful reaching behavior in a prosthetic limb. An implication of their study might be that a high-bandwidth coupling between brain and world is unnecessary for many kinds of action. Haugeland's use of the bandwidth criterion of component individuation is also criticized by Grush (2003, 79) who instead endorses the "plug points criterion," according to which "components are entities that can be plugged into, or unplugged from, other components and/or the system at large." For further discussion of the issues raised in this footnote, see Adams and Maher (2012) and Kaplan (2012).

ensures that the relevant non-bodily determinates “are part of the local equipment that plays a causal role in the generation of action” when it comes to EM (Clark 2008, 79).

In chapter four, I will employ the account of constitutive grounding to help argue for the temporally extended vision thesis, TEV, i.e., that no conscious visual experiences are constitutively grounded by instantaneous events involving a subject’s body. Any internal vehicles of conscious visual experiences will always be extended in time. In chapter five, I will argue that some states and processes that are conscious and visual are constitutively grounded by events that include parts of the non-bodily world that are continuously and actively coupled in the right way to parts of that subject’s brain. I call this version of ECM extended visual consciousness, EVC. Even if my arguments for TEV and EVC fail to convince, in making them I hope to show that constitutive grounding is itself a coherent, useful, and independently philosophically interesting relation. In both chapters, I proceed by arguing that sufficient explanations of why visual experiences of certain types occur must make reference to what happens beyond some spatiotemporal boundary and that the implications regarding the falsity of the relevant instantaneous supervenience and diachronic bodily supervenience theses are in fact plausible.

William James once wrote: “[t]he pragmatic method is primarily a method of settling metaphysical disputes that otherwise might be interminable” (1997, 2). Since vacuity threatens to undermine debates about EM and its species, adopting a pragmatic methodology may be appropriate when approaching these debates. A pragmatist, according to James, “turns away from abstraction and insufficiency, from verbal solutions, from bad *a priori* reasons, from fixed principles, closed systems, and pretended absolutes and origins”

(5) and instead looks toward “last things, fruits, consequences, facts” (7). While my account of constitutive grounding itself is unapologetically abstract, the imperative to focus on what is concrete and consequential is often an indispensable tool for good metaphysical theorizing. In keeping with this much pragmatism, I grant that the viability of my account of constitutive grounding should be assessed by the uses to which it can be put in locating events in the actual world. Moreover, my arguments in chapters four and five, while theoretical, largely prioritize evidence that is concrete and empirical. Both James and his fellow American pragmatist, John Dewey, can plausibly be interpreted as early advocates of ECM. Moreover, Putnam’s (1999) recent account of perceptual experience in transactional terms has inspired some of the most well-known defenses of ECM in the literature, particularly Noë’s (2004, 2009).⁵ Dewey (1916, 336) lamented in his time that “the older dualism of soul and body has been replaced by that of brain and the rest of the body.” Yet despite the rise of externalism in twentieth century philosophy of mind, Dewey’s complaint still seems apt in the twenty-first.⁶

⁵ Ward (2012, 5) points out that the “enactive view of experience as essentially world-involving interactions” contrasts with conceptions of perceptual experience as “mere affectation of a person’s subjectivity” (Putnam 1999, 11). I agree with Clark (2012) and disagree with Ward (2012) that the enactivist case for ECM is significantly different than the naïve realist or disjunctivist case.

⁶ Perhaps there are fairly deep psychological reasons for this that would be explained if Bloom (2005) is right that our propensity to understand the world in dualistic terms is innate. Also apt to my eventual defense of EVC is Dewey’s (1896, 357) earlier lament that “[t]he older dualism between sensation and idea is repeated in the current dualism of peripheral and central structures and functions.”

2. ACTIVE VEHICLE EXTERNALISM

In order to sufficiently understand mind and behavior, we must look beyond what happens within our heads. That seems to be one important lesson from a number of influential developments in philosophy and psychology throughout the 20th century. Before setting out to survey some of these developments, it is worth briefly acknowledging the *prima facie* plausibility of the contrasting internalist perspective. A global form of internalism holds that everything about our minds depends entirely on and can be sufficiently explained in terms of the states and processes intrinsic to our brains. Internalists acknowledge that there are external causes of our mental states and processes, of course, but they deny that anything that happens outside the brain and body is ever strictly necessary. Internalism intuitively captures one way we often think about the interface between our minds and the rest of the world, i.e., our minds are locally determined islands of subjectivity, interactively navigating an abyss that is in no way part of us. After all, as various skeptical possibilities make salient, provided that what happens internally stays the same, the world beyond our heads could be different in all sorts of undetectable ways without altering the contents and characters of our minds.

Of course, it is not pre-theoretically intuitive that it is the brain itself rather than some other part of the body like the heart or the pineal gland that constitutes the local physical basis of our minds. Brain-based internalism is plausible, however, not merely because it fits well with common sense conceptions of privacy and subjectivity. In addition to the ways in which it has been demonstrated that mental events systematically depend on neural events, the development of cognitivist theories of mind in the mid-20th century

suggests how the brain could be the mechanism that implements the mind. Like a computer running a software program, a brain may act as a physical symbol system, i.e. a system that sorts physical patterns into symbols and uses them in conjunction with other symbolic, i.e. formally specifiable and semantically invariant, representations in cognition to guide behavior. Newell and Simon (1976, 116) hypothesize that a “physical symbol system has the necessary and sufficient means for general intelligent action.” If they are right, then since what is typed on the keyboard and what appears on the screen of a computer are only causally related to the states and processes of the software program running but do not determine those states and processes in the way properties internal to central processing unit of the computer do, we should expect the same may be true of brains. In other words, cognitivism suggests a way of understanding how properties intrinsic to the brain could be sufficient on their own. Sensory organs and enervated muscles are mere causal inputs and outputs respectively, with the meat of the mind sandwiched in between.⁷

Developments in logic, AI, and computer science in the mid-20th century have encouraged an internalist orientation toward investigating mental phenomena. Moreover, the linguistic theory that emerged from Chomsky’s (1959) repudiation of behaviorism and his claim that knowledge of grammar is innate and universal also supports internalism. If the basic architecture of the human brain encodes all the information necessary for our linguistic abilities, then the role of the environment in enabling us to learn a specific natural language may seem more a matter of calibration than something strictly necessary. Once

⁷ On this cognitivist conception of cognition, what Hurley (1998, 20) calls ‘the classical sandwich’: “[t]he subpersonal underpinnings of the mind are conceived as *vertically modular*, with cognition interfacing between perception and action.” Vertical modularity, as opposed to Hurley’s notion of ‘horizontal modularity’, is the sort Fodor (1983) famously articulated.

learned, it stands to reason that our linguistic abilities can be sufficiently explained by the internal processes that enable production and comprehension. Fodor's (1975) language of thought hypothesis and defense of a computational theory of mind provide further basis for accepting internalism. According to Fodor, cognition, like language production and comprehension, involves the manipulation of formal symbols or concepts that are innate. A natural upshot of such an understanding may be methodological solipsism (Fodor 1980), i.e. the view that mental properties can and should be individuated without reference to the wider environment. After all, if cognition just involves transformations of innate symbols, the vehicles of cognition are likely always intrinsic to a subject.⁸

I will eventually be arguing for a kind of externalism about visual experience.

Within the cognitivist tradition, Marr's theory of vision is the most relevant foil. There has been substantial debate about whether or not Marr's theory really implies content internalism. While Segal (1989) and Egan (1992), for instance, have argued affirmatively, Burge (1986), Davies (1991), and Shapiro (1993) have all argued that Marr's theory is in fact compatible with externalism. That debate will not figure prominently in the present exposition because the issues at stake between internalists and externalists regarding contents are somewhat orthogonal to the issues between internalists and externalists about vehicles. Whether or not Marr's theory of vision is compatible with content externalism, it is clearly a form of vehicle internalism.

Ironically, while cognitivism suggests that the local basis of our mental states and processes is entirely in our heads, it has also encouraged the idea that psychology and

⁸ Fodor's vehicle internalism is compatible with either content internalism (Fodor 1987) or content externalism (Fodor 1994).

cognitive science can proceed without much attention to the details of the brain. Once we have determined the nature of some cognitive task and the algorithm needed to execute that task, we need look at the brain only to discover how it manages to implement the relevant algorithm. However, beginning in the 1980s with the growth of connectionist models of cognition and an explosion of research in neuroscience, purely computational theories lost some of their luster. The details of the brain are relevant not just as matters of implementation but also in understanding the nature of the task at hand. An outgrowth of the rise of connectionism was the dynamical systems approaches to cognition, which were championed by Beer (1995) and Van Gelder (1995) who further emphasized the relevance of the details about the niche in which a brain and body is embedded. Moreover, some of the work on perception, cognition, and action that fell outside the cognitivist tradition even in the heyday of cognitivism, e.g. Gibson's (1966, 1979/1986) ecological approach and the early enactivist theories of Maturana and Varela (1973, 1980, 1987), were always wary of the idea that we should expect to find brain-internal sufficient explanations for all aspects of our minds.

As the 17th Century English poet John Donne famously observed: "no man is an island, entire of itself." Our minds are embedded not just in a surrounding physical world but a world inhabited by other minded beings, and internalist theories may only be able to explain these social dimensions in artificially derivative ways. If social and normative aspects of our minds cannot be sufficiently understood by appeal to states and processes intrinsic to the brains of individuals, however, then global internalism obviously fails. Similarly, it may be that we need to look beyond the individual and focus on the histories

of its ancestors and the greater evolutionary context in order to fully appreciate how and why its mind is the way it is. Millikan (1984, 1989, 1993, 2004), for instance, has influentially argued for externalism about the contents of our mental states and processes on the basis of evolutionary considerations.⁹ Even if thoroughgoing teleosemantic accounts are guilty of externalist overreach, appreciating the social and evolutionary context in which our minds are embedded should lead us to be dubious of the methodological solipsism presumed by most in the cognitivist tradition.

Debates between internalists and externalists over different dimensions of mind can be difficult to track. In part, this is due to just how many different combinations of positions there are with respect to the dimensions that have been distinguished. An internalist about moral reasons may be an externalist about knowledge, for instance. To better understand what is at stake when it comes to active vehicular versions of externalism, I will survey some of the debate between internalists and externalists and clarify how EM and ECM are unique versions of externalism. I will then review and preliminarily describe some recent unorthodox accounts of perception, which likely imply ECM. Where we end up should leave the reader with a sense of why externalism about the vehicles of conscious states and processes is a plausible and interesting view worthy of sustained defense.

⁹ Papineau (1987) also provides an influential defense of an externalist teleosemantic account of mental content.

2.1 Externalism Generally

Ludwig Wittgenstein's (1953) worries about the privacy of thought and Gilbert Ryle's (1949) analytic behaviorism helped introduce externalist themes into contemporary philosophy, but it was Saul Kripke's (1980) and Hilary Putnam's (1973, 1975) defenses of externalism about the meanings of proper names and natural kind terms in language that really marked a shift. Since then, philosophers of mind have been concerned with understanding whether externalism is also true of the thoughts and ideas that are communicated through the use of such expressions. If one retains something like H.P. Grice's (1957, 1968, 1969) account of the distinction between utterer's meaning and semantic meaning, one can accept externalism about the semantics of natural language while retaining internalism about mental contents. But many have become convinced that externalism creeps into the mind as well.

Externalists hold that some of the mental properties subjects instantiate are individuated by relations that they bear to things other than themselves. One influential proponent, Tyler Burge (2010, 61) has recently characterized externalism, or what he calls 'anti-individualism', as the following thesis:

[The natures of many mental states constitutively depend on relations between a subject matter beyond the individual and the individual that has the mental states, where relevant relations help determine specific natures of those states.

Externalism implies that some mental properties are neither intrinsic properties of the subject instantiating them, nor exclusively dependent upon the subject's intrinsic properties. If any version of externalism is true it follows that it is possible for there to be a full duplicate of a subject on some occasion that has all the same intrinsic properties, but who

instantiates different mental properties. Mental properties that can vary among intrinsic duplicates (or so-called 'twins' or 'doppelgängers' or 'counterparts') must depend in part on what the world external to the subject is like.

To get a better idea of what externalists claim, consider some uncontroversial examples of non-mental properties for which comparable externalist theses hold. If someone has a mosquito bite at some point in time, it is possible for there to be an intrinsic duplicate of that person, one with exactly the same skin damage, inflammation, etc., who does not have a mosquito bite. If the exact same symptoms were caused by anything other than a mosquito, it would not be a mosquito bite. The same is true for sunburn. An intrinsic duplicate of a sunburned subject, burnt by something other than the sun, would not be sunburnt. The relations the subject bears to a mosquito or to the sun, in these cases, partly determines whether or not the subject instantiates a property of the specified type. Another straightforward case involves the conventional value assigned to human artifacts. An intrinsic duplicate of a coin or an artwork can vary in value depending on its origin. An intrinsic duplicate of a real coin or original artwork could be a forgery, hence worth less than the original, while an intrinsic duplicate of forgery could be a real coin or original artwork, hence worth more. It follows that the value of the coin or artwork depends on more than its intrinsic properties; it is also determined by the artifact's origin. Traditional externalists about content maintain that some types of mental properties are like this. Whether they are instantiated depends on relations between a subject and things that are not ordinarily taken to be part of the subject.

Even if externalism is true for some types of mental properties, many philosophers believe that numerous types of mental properties are by their nature intrinsic properties of an individual. If they are right, it follows that these types of mental properties will be instantiated by all duplicates, thus any kind of global externalism would be ruled out. Internalism would have to be true about at least some types of mental properties. Internalism itself involves an acceptance of the weaker claim that subjects' mental properties depend or supervene on their intrinsic properties. Identity is not required. Roughly, one class of properties supervenes on another if and only if it is impossible for there to be a difference with respect to the first unless there is a difference with respect to the second. If certain types of mental properties are intrinsic properties of the subject, then the supervenience claim is obviously satisfied.

If all mental properties are intrinsic properties, a global version of internalism would follow trivially; anything that duplicates all of a subject's intrinsic properties will thereby also instantiate all the same mental properties. Externalists typically reject internal supervenience for at least some mental properties. In other words, they accept that there can be differences with respect to whether a subject instantiates those properties even if everything intrinsic to that subject or to that subject's body or brain stays the same. In principle, however, while internalism entails internal supervenience, externalism does not strictly entail its falsity. For instance, if there are necessary connections between what's internal and what's external to the relevant boundary, internal supervenience might hold even in cases where the supervenient property itself could not be instantiated nor explained independent of something external. Nevertheless, the most straightforward way to argue for

externalism is to argue for a failure of internal supervenience for some important class of cases.

2.2 Factive states and *de re* contents

A plausible place to begin examining externalism about some aspects of mind is knowledge. Consider whether a subject knows something. Since truth is a condition on knowledge and whether a belief is true or false varies depending on factors that are entirely extrinsic to the subject, it follows that a subject knowing something depends on more than the subject's intrinsic properties. We can believe things and be mistaken. When we are mistaken, we do not know what it is we believe. Knowledge and other so-called 'factive' mental states and processes require accuracy. Since, as Austin (1962) and many others have noted, there are clearly factive interpretations of perceptual verbs like 'to see', 'to hear', 'to taste', 'to smell', 'to feel', 'to sense', 'to perceive', 'to experience', 'to observe' etc., an inference to a version of externalism that goes beyond just states of knowledge and processes of knowing is warranted. But this is still a somewhat uninteresting form of externalism. After all, there are also non-factive interpretations of the same perceptual verbs. Subjects who have lost limbs can sometimes feel pain as though it were coming from the missing limb. A subject on LSD having extreme visual hallucinations is still, in a sense, seeing. There are also locutions such as 'seeming to' or 'appearing to' that, when appended to a perceptual verb, cancels any requirement of accuracy and so rules out factive interpretations. Finally, many types of mental states central to ordinary folk psychology, e.g. beliefs and desires, simply carry no presumption of accuracy or fit at all.

Any simple case for externalism about knowledge and other factive mental states and processes does not apply across the board. On occasion externalists impressed by the kinds of considerations just discussed will respond to internalists by adopting an explicitly teleological conception of mental kinds. Accordingly, states like ‘false belief’ and ‘seeming to see’ are considered deviations from successfully functioning veridical states. Articulating this point, Williamson (2006, 303) explains that teleological externalists may go on to say, “the failures differ internally amongst themselves, and that what unifies them into a theoretically useful category with the successes is only their relation to those successes.” For instance, a visual hallucination of a pink elephant and an image of one in a dream are only of a common type by virtue of their relation to actually seeing a pink elephant, according to this kind of teleological externalist.¹⁰ Internalists are likely to counter that a common characterization of successful and non-veridical cases is required. After all, pink elephants were likely hallucinated, dreamed of, or imagined long before anyone ever painted some poor elephant pink and so actually *saw* a pink elephant. The external world or environment must be kept separate, in fact, so that it can provide the extra conditions that distinguish common cases into successes and failures, e.g. knowledge from justified false belief.

When it comes to ECM, my strategy will be to focus on a certain range of phenomenal character types that plausibly occur in the course of everyday human life. I claim they are more like knowledge than belief in this respect. While truth distinguishes

¹⁰ Cf. Burge (2010, 68): “For an individual to have any representational state (such as a belief or perception) as of a subject matter, that state must be associated with some veridical representational states that bear referential, indicational, and attributional representational relations to a suitably related subject matter.” By analogy, Ryle (1954, 94): “There can be false coins only where there are coins made of the proper materials by the proper authorities.”

knowledge (and any other factive state or process), within epistemology justification has traditionally been thought to depend only on evidence intrinsic to a subject's mind. Early modern rationalist and empiricist theories were internalist theories of justification in this sense. Causal theories of knowledge such as reliabilism, on the other hand, are externalist. Reliabilists hold that beliefs are justified only if they are produced by objectively truth-tracking processes.¹¹ Since whether or not a belief forming process tracks the truth depends on features of the world external to the subject's body, it follows that a subject's body could be duplicated without duplicating the justificatory status of the belief under examination. The issues at stake between internalists and externalists in epistemology, however, are not central to the issues that surround EM.

Most of the debate over externalism in the philosophy of mind has focused on the meanings or contents of representational states like belief. When thinking about the nature of the contents of mental states and processes, it helps to distinguish between *de dicto* and *de re* types. Consider the following belief ascriptions.

Oscar believes of water that it fills the oceans.

Oscar believes that water fills the oceans.

In order for the first statement to be true, i.e. in order for Oscar to correctly have the belief ascribed to him, there must be water. The expression 'of' indicates a *de re* content, one that is directly referential so requires the existence of the thing denoted. Like knowledge and other factive states, states with *de re* contents straightforwardly depend on more than the subject's intrinsic properties. The second statement, on the other hand, can be true

¹¹ For classic defenses of reliabilism, see Goldman (1976) and Dretske (1981).

whether or not there is any water; Oscar could still have that belief even if it was false. This kind of mental attitude ascription, involving a ‘that’-clause, expresses a *de dicto* content. Putnam’s famous argument for externalism targets *de dicto* contents. Unlike *de re* contents, *de dicto* contents seem more likely to be determined by what is intrinsic to a subject.

2.3 Twin Earth

In Putnam’s (1973, 1975, 1982) Twin Earth case, one imagines a version of reality that is superficially indistinguishable from how things actually are, in terms of macroscopic observation and high-level function, but different in some more basic respect. One then considers whether the contents of a duplicate subject’s mental states and processes would be different if things were that way. Suppose that Oscar was a person living on Earth in the year 1750 (before the discovery of the chemical composition of water) and suppose he believed that water fills the oceans. Now imagine another planet, Twin Earth (either in an entirely different possible world or in a spatiotemporally distinct but qualitatively indistinguishable part of the actual world), in which there is a duplicate of Oscar, Twin Oscar, who instantiates all of the same intrinsic properties as Oscar. The only difference between Earth and Twin Earth is the chemical composition of the clear drinkable liquid pervasive in each of their environments. While on Earth, the pervasive clear drinkable liquid is H₂O, on Twin Earth it is XYZ. Does Twin Oscar have the same belief as Oscar; does Twin Oscar believe that water fills the oceans? Externalists have argued he does not. While Twin Oscar will utter the same sounds to express his belief, both his word ‘water’ and his corresponding concept, refer to something quite different. His beliefs are about

XYZ, whereas Oscar's are about H₂O. If this is correct, it shows that preserving a subject's intrinsic properties is not always sufficient for preserving the *de dicto* contents of that subject's mental states.

An obvious problem with the case as told is that it is impossible for there to be an intrinsic duplicate of an actual human subject who is not composed of tissues that are largely composed of H₂O. So even if XYZ on Twin Earth functions just as H₂O on Earth, some of Twin Oscar's intrinsic properties will nevertheless be different. If we suppose that somehow Twin Oscar's tissues just happen to be composed of H₂O on the occasion in question, while everything else watery on Twin Earth is XYZ, it is no longer clear what his thoughts are about. Nevertheless, duplication cases can be constructed that avoid this problem. For instance, whether a duplicate of an actual subject has thoughts about tigers or elms depends in a similar way on whether the local tiger-like and elm-like things have the same underlying structure and composition as actual tigers and elms (and perhaps the same evolutionary lineage). Still, it is generally important to be clear about which things are being held fixed and which are allowed to vary in twin cases and other supervenience thought experiments involving, e.g., swamp people, brains in vats, or zombies.

Following convention, let us say that mental contents for which externalism is true are 'wide' contents, and contents for which internalism is true are 'narrow'. The wide contents of a subject's mental states and processes depend on relations the subject bears to things other than himself/herself/itself, whereas narrow contents depend only on the subject's intrinsic properties. Many philosophers, impressed by arguments for externalism about content, have attempted to blunt the force of the view by claiming that mental states

and processes always have narrow contents as well as wide contents. Narrow contents are commonly understood as functions that map contexts onto wide contents. If something like this picture is adequate, then there is a sense in which narrow contents must be prior to and partially determinative of any wide contents a mental state or process might have. So-called ‘dual factor’ accounts attempt to accommodate the existence of wide contents while retaining internalism about a special class of contents.¹² While wide contents depend on context; narrow contents are autonomous and independent of context. Hence, narrow contents are determined by, or identical to, the intrinsic properties of the subject in question.

Even if wide contents are required to explain some of the social and normative dimensions of our mental lives, narrow contents, many think, are needed for providing psychological explanations of behavior. An important objection to externalism about *de dicto* contents is that if mental states and processes sometimes have wide contents, such contents play no role in explaining subjects’ actions and movements. Since thoughts and desires cause behavior and causation is primarily a local affair – causes and effects are spatiotemporally contiguous – it is claimed that only narrow contents can be relevant to psychological explanations of behavior. This objection suggests a thoroughgoing externalism about mental content cannot be right. This is the causal powers argument for internalism.¹³ One way for an externalist to respond is to point out that relational properties are explanatory in other scientific domains. For instance, Burge (1986, 14-20 and 1989,

¹² E.g. Block (1986), Field (1978), Loar (1981), and McGinn (1982).

¹³ Fodor (1987, 12): “It is not just that, in a psychology of propositional attitudes, content and causal powers are attributed to the same things. It’s also that causal relations among propositional attitudes somehow typically contrive to respect their relations of content, and belief/desire explanations often turn on this.” For these kinds of explanations, Fodor argues, wide contents play no role.

309-310) discusses the following examples of kinds individuated by appeal to more than intrinsic causal powers: battles, continents, plates, organs, and species. These properties figure into accurate explanations and predictions within the special sciences despite requiring relational individuation. Wilson (1995) adds the properties of being a volcano, a mountain, having syphilis, lead poisoning, or a birth trauma to the stock of examples. Wilson (2004, 2005) elaborates and defends the explanatory possibilities that extrinsic properties hold in the social sciences and biology as well as a psychology.

Externalists have argued that many types of actions can only be individuated relationally. To shake someone else's hand, that person must exist. To swing a golf club, the club must exist. In such cases, it is plausible for the externalist to insist, "the content of the agent's intention typically corresponds to the broad individuation of the action in terms of the external objects that it involves" (Williamson 2006, 295). Since such actions cannot be sufficiently explained by the narrow contents of subjects' mental states and processes alone, wide contents seem to be required. Hence, the contents that explain behavior are typically wide. An internalist may insist that some basic set of actions or movements require only narrow contents to explain, but this still leaves many types of actions requiring explanations in terms of wide contents. While wishing to remain as neutral as possible on this, I think the accounts that individuate content partially in terms of downstream effects are more plausible than those that individuate content entirely in terms of causal origin.¹⁴

¹⁴ In other words, I am more inclined to endorse an account like Millikan's (1984) than Dretske's (1981) or Fodor's (1987). Here I agree with McGinn (1989, 65-66) that "when it comes to a competition between action and environment, in the fixation of perceptual content, action wins". As Hurley (1998, 424) usefully observes while writing about these issues: "Evolutionary teleology inherently involves relations between input from environment to organism and output from organism to environment, a dynamic system of causes and effects."

2.4 Direct Realism, Wide Intentionalism, and Inverted Earth

So far I have been examining externalism in general terms and focusing on how it applies to cases of *de dicto* and *de re* contents, but ECM is an externalist thesis about the vehicles of mental states and processes with phenomenal characters. I suspect that externalism about phenomenal consciousness strikes many as especially implausible in part because there is no settled opinion on the relationship between phenomenal character and representational or intentional content and because many have strong intuitions that how the world introspectively appears to a subject is always potentially independent of how the world actually is. Externalists about consciousness believe that fixing all of the intrinsic properties of a subject's body at a time fails to fix whether or not the subject has an experience with some specific phenomenal character. Vehicle externalists think this because they think the vehicles of some experiences are not determined by properties intrinsic to the experiencer's body; externalists about phenomenal character think this because they think properties intrinsic to the experiencer's body do not completely determine what the relevant experiences are like for the experiencer.

Here I wish to discuss two versions of externalism about phenomenal character: wide intentionalism and disjunctive direct realism. These views reject the idea that phenomenal characters are either identical to or completely determined by the intrinsic properties of a subject. Nevertheless, they generally accept the idea that subjects or agents or selves are demarcated by the organismic boundaries, the entire body, or by just a part of the body, such as the brain. On these accounts, it is just that the phenomenal characters are

partly individuated by subjects' relations to the environment. As we will see, ECM, while compatible with the idea that subjects are body bound, is better able to do justice to certain internalist intuitions than other accounts if it is combined with the idea that subjects spatially extend into regions more expansive than their bodies.

Direct realists about perceptual experience insist that objects and properties in the surrounding environment are literally parts of the perceiving subject's experiences. Recall that perceptual verbs like 'to see' and 'to hear' admit of factive interpretations. In order to see a blue car, there must really be a blue car there. Of course, it is not merely the possibility but the actual existence of certain illusions, afterimages, hallucinations, and dreams which suggests that subjects often have experiences that are phenomenally indistinguishable, despite the fact that no or different external objects are present in such non-veridical cases. There simply does not seem to be anything phenomenally different in kind about the experience of actually seeing a blue car as compared to vividly hallucinating one.

A duplication case makes this point with precision. Suppose that René is a normal human subject experiencing redness as he looks at a tomato. Twin René's body, however, is actually embedded in a vat hooked up to a computer in a world with no tomatoes. Nevertheless his body on a certain occasion is an intrinsic duplicate of René's body at the time René was seeing red. Most people have the intuition that in such a situation Twin René would still have an experience with the same phenomenal character as René's. But if that is right, then, unlike the original Twin Earth case, there seems to be no obvious dependence of phenomenal character on anything external to René's body. Certainly the

phenomenal character does not depend on the presence of red objects with specific surface reflectance properties and so forth.

In contrast to internalist views that insist that veridical and hallucinatory experiences share a core phenomenal character, disjunctivist direct realists deny this.¹⁵ Instead they claim that there are always further fundamental phenomenal differences between any pair of such cases, even if indistinguishable to the subject. Two experiences may be phenomenally indistinguishable from a subjective point of view without necessarily having the same phenomenal character. Since successful veridical perception is the normal case, direct realists distinguish veridical perception from cases of hallucination, illusions, etc. and claim that these defective cases are always phenomenally different in kind from real cases of seeing, hearing, etc. There remains a burden for the disjunctive direct realist to explain something about why the unsuccessful states and processes of subjects are indistinguishable to their subjects. Part of their idea, I think, is that all of the non-veridical cases involve some kind of epistemic impairment or cognitive deficiency.

The other main way philosophers argue for externalism about phenomenal character is by arguing that the phenomenal characters of conscious states and processes are determined by their representational contents and that those contents are always wide. The argument for the second conjunct relies on the plausibility of externalism about content generally. The representational point is motivated by the claim that perceptual experiences are transparent in the sense of making their subjects aware of the world itself

¹⁵ Byrne and Logue (2008) and Haddock and Macpherson (2008) are two recent collections of papers on disjunctivism.

rather than their own intrinsic properties.¹⁶ On the basis of these considerations, wide intentionalists maintain that the phenomenal characters of different experiences are determined by properties of the environment external to the subject having those experiences, such as those things that subject has learned to interact with or those things that subject's ancestors have regularly and reliably encountered in the past and so have helped shape the wide contents of the subject's current experiences.¹⁷

Internalists about phenomenal character can respond to wide intentionalists in a number of different ways. One option is to simply deny that there is a close connection between phenomenal character and intentional content. Making this move is compatible with direct realism, but in conjunction with the idea that phenomenal characters are intrinsic features of a subject's experiences and that subjects are bound by their bodies, it leads to internalism. Traditionally, this kind of internalism about phenomenal characters has been associated with the view that experiences have sense data or qualia, which may help to build up contents, but do not themselves represent. While many are still drawn to this view, developments in 20th Century philosophy of mind have led to it falling out of favor.

Allowing that there is an intimate relationship between the intentional or representational and the phenomenal, there are two routes an internalist might take in responding to the wide intentionalist. One is to accept the tight connection between the phenomenal and the intentional, perhaps even allowing that conscious states and processes

¹⁶ On transparency, see Harman (1990) and Moore (1992, 23): "When we try to introspect the sensation of blue, all we can see is the blue: the other element is as it were diaphanous." Tye (2002), for instance, explicitly argues for what I am calling 'wide intentionalism' on the basis of transparency.

¹⁷ Byrne and Tye (2006), Tye (1995, 2000), Dretske (1995), Lycan (1987, 1996), Thau (2002) all defend wide intentionalism.

always have wide contents by adopting a dual-factor view, but claim that phenomenal characters only depend on the narrow contents of the relevant experiences.¹⁸ Another option is to argue that phenomenal character is both intrinsic to experience and furnishes representational contents by determining a set of accuracy or application conditions which carve up the space of accessible possible worlds simply in virtue of the character itself.¹⁹

Block (1990) objects to wide intentionalism by appeal to another influential duplication scenario in which phenomenal character is unaltered despite alterations in wide content. Suppose that Ned is a normal subject. A duplicate of Ned, Twin Ned, inhabits a world, Inverted Earth, in which everything is the same except that the colors of objects and surfaces are systematically inverted. For instance, while everyone on Inverted Earth calls the sky 'blue', as we do, since the sky there is actually yellow and the inhabitants of Inverted Earth are normal perceivers, they experience it as yellow. Suppose further that Twin Ned has been unknowingly fitted with color inversion lenses, so that the sky actually appears to him to be blue, not yellow. More generally, the phenomenal characters of color experiences throughout his life would always be the same as Ned's, so it seems. Nevertheless, given the environment on Inverted Earth, any standard externalist theory of content determination will entail that the contents of his bluish visual experiences actually represent yellow, not blue. If Block is right, this shows that there can be differences in the phenomenal character of an experience without differences in its representational content, contrary to what wide intentionalism requires.

¹⁸ Carruthers (2000), Rey (1998), Chalmers (2005) all argue for a kind of 'narrow intentionalism' about consciousness in this way.

¹⁹ This is the strategy taken by proponents of 'phenomenal intentionality', see essays in Kriegel (2013).

Debate over the implications of Block's Inverted Earth and related cases is voluminous.²⁰ More recently, Pautz (2006) has developed an empirically motivated version of this challenge to wide intentionalism. In Pautz's (2006) case, Twin Maxwell inhabits a world in which human evolution has resulted in hominid organisms with slightly different opponent processing systems. Specifically, orange light is processed along a single opponent channel, like red, yellow, blue, and green are on Earth, instead of two (as orange is on Earth). Pautz argues that Twin Maxwell has different experiences when looking at an orange square than his Earthly counterpart Maxwell. Opponent color processing theory predicts differences in phenomenal character given the same optical stimuli in this case, specifically that Twin Maxwell, unlike Maxwell, will see orange as a unique hue rather than as a mixed one. Pautz goes on to argue that the standard causal co-variation accounts of content fixation available to the wide intentionalist cannot account for this.²¹

None of the preceding discussion counts decisively against wide intentionalism or disjunctive direct realism. My concern at the moment is only to highlight some important differences between these kinds of externalism about phenomenal consciousness and ECM. Most significantly, perhaps, despite Putnam's slogan 'meaning ain't in the head', neither traditional content externalism nor wide intentionalism actually implies that token mental states and processes ever partly reside in the non-bodily world.

To see why, recall the sense in which externalism is clearly true about the value of currency. A duplicate of a dime that was not produced at an official US mint would not

²⁰ It is worth noting that Hurley's (1998, 245-337) early arguments for EM are driven in part by an attempt to make sense of the assumptions about duplication that were being presumed in cases like Block's.

²¹ Byrne and Tye (2006) respond by questioning both whether Twin Maxwell's experience would be different and whether externalist semantics is committed to the verdict that they will be.

have the property of being worth 0.1 USD. Nevertheless, if I have a dime in my pocket, that dime is nowhere other than in my pocket. Coming from a mint may be a determining factor in its having the monetary value it has, but it is not as though the mint from which it originated is currently part of the dime. Similarly, it may be that a subject consciously seeing the setting sun in some particular way is in part determined by the ubiquitous presence of the sun in the environment that has shaped the life of that subject and the lives of that subject's ancestors, but that gives us no reason to think that those ancestors, or the sun itself, are literally parts of subject's current visual experience.

The lesson here is that relational individuation of content or character type is not germane to the location of specific tokens of those types. Perhaps token states and processes with contents and characters always stay in the head, a kind of vehicle internalism, even if contents are often or always wide and even if phenomenal characters are also often or always wide.²² In other words, just as vehicle externalism is compatible with some versions of internalism about phenomenal characters, externalism about phenomenal characters is compatible with internalism about the vehicles of all conscious processes and states.²³ While I will be moving next to a discussion of vehicles, it may be worth it here to simply stipulate a conceptual or ontological link between vehicles and location. Hence,

²² Relatedly, consider that though McGinn (1989) and Burge (1986, 1993) have argued that externalism refutes token-identity theory, Davidson (1987) responds by drawing an analogy to sunburns. A sunburn may be identical to a token body state, even though it would not have been a sunburn had the burn been caused by anything other than the sun. If Davidson's line of reasoning works, it is important because it shows that even essential, type-individuating, dependence is not sufficient for constitution (in the location constraining sense). Externalism about type-individuation does not entail externalism about token-location.

²³ For instance, Dretske (1995), who advocates a fairly global form of externalism, including wide intentionalism about consciousness, accepts internalism about the vehicles of conscious and cognitive states and processes.

Putnam's slogan would follow if vehicle externalism about content-bearing mental states and processes is correct even though it does not follow from content externalism alone.

Though externalism about contents or characters does not entail externalism about vehicles, disjunctive direct realists endorse both. They maintain that perceived objects instantiating properties are literal constituents of the relevant states and processes. To reiterate, disjunctive direct realism about perceptual experience is different from wide intentionalism in that it implies that the locations of veridical experiences include parts of the world. Since it implies that some subjects' mental states and processes extend beyond their bodies, it is a version of ECM.

ECM, however, does not entail disjunctive direct realism, which requires that the actual perceived object or property, something that can be thought about, desired, etc. after the fact, be present in the environment. To see a barn, there must be a barn, not merely a barn façade. ECM, however, does not require this. It is compatible with only some relatively abstract property that the barn and barn façade share being instantiated. An upshot of this is that ECM is compatible with the claim that external vehicles (as well as internal vehicles) are multiply realizable. But ECM is also compatible with the claim that there is an important core similarity between veridical and hallucinatory cases of perceptual experience that is phenomenologically manifest. It is just that the states and processes that preserve all fine-grained phenomenal similarities will only be tokened with some kind of environment there for the perceiver's brain and body to couple to and interact with in the right sorts of ways. This does not require veridical perception, since a computer properly hooked up to a brain in a vat could suffice. To recapitulate: disjunctive direct realism

entails ECM, wide intentionalism does not, and ECM entails neither version of externalism about phenomenal characters.

2.5 Vehicles

One way to express the takeaway point about wide intentionalism is to say that externalism about what individuates content or character types is compatible with internalism about the vehicles that carry or realize or ground content or character. The type of mental state or process might require relational determination even though the individual tokens themselves can be identified in ways that depend only on the intrinsic properties of the subject's brain or body. This raises, of course, the tricky question of what exactly vehicles are. I think a plausible place to start is in terms of their ontological category. Since we want there to be some kind of necessary or constitutive relation between a vehicle and the location of a token state or process, I think it helps to understand vehicles as events. Events are either identical to regions or at least partially individuated in terms of their relations to regions. Instead of identifying events as regions, I take events to be individuated by an instantiation relation between some property, a state or process of some type, and a region. Events can and do figure into both causal and non-causal relations of determination or dependence.²⁴ This still leaves many of the tricky questions about what

²⁴ Individuating events in terms of properties does have some counterintuitive consequences. For instance, it entails that a person's flight to Paris and that same person's flight to France (on the same day and time) are two distinct events. That person's flight to Europe might be yet a third. This leads to an obvious proliferation of events within any given region, when many are wont to suppose there is and could only be one event per region. If this is a cost of adopting intensional individuation conditions for events, it is one I am willing to pay. I suspect individuating events this way is necessary for them to play the explanatory and determinative roles they clearly do. Moreover, taking events to be regions under a description, as Davidson (1969) does, only yields cases of constitutive dependence between spatiotemporally overlapping events that are conceptual or

vehicles are and about what their relationship is to contents and characters they carry unanswered, but it is a start.²⁵

While the notion ‘vehicle’ commonly appealed to in the literature it is not entirely clear, Clark (2008, 76), explicitly claims that EM “is really a hypothesis about extended vehicles.” Most other proponents of the view have made similar claims. One promising place to gain traction is by looking at Hurley’s (1998) token-explanatory conception of vehicles. According to this conception, vehicles are “what is left when you subtract the explanation of why content of that type exists at all from a fuller explanation of why one particular mental state with that content obtains” (330). In other words, vehicles “do the *differentially* token-explanatory work” in explaining why and how various mental events of various content or character types occur when they do (330).

Examining the relation between the assumptions about modal duplication that are made in various twin cases and the token-explanatory role vehicles are supposed to play, Hurley emphasizes a certain contingency in the relationship between vehicles and contents and characters. Fixing vehicle type does not fix content or character type; and fixing content or character type does not fix vehicle type. Some content/character internalists accept that fixing vehicle type fixes content type, but Hurley argues at length that this assumption has problems. The token-explanatory conception encourages us to instead adopt the view that the spatiotemporal boundaries of vehicles and the boundaries of duplication are “nonaccidentally related” (331). This still leaves open questions about how to type vehicles, but Hurley claims it has the advantage of vindicating the idea that an event is a vehicle if

analytic. But part of my motivation for developing an account of constitutive *grounding* is that I think doing so makes viable non-analytic priority relations.

²⁵ For rhetorical convenience, I will continue use ‘vehicle’ to refer to states and process on occasion.

and only if it can be duplicated. In this way “[w]e can accept that differentially-token explanatory factors, or vehicles, must in principle be duplicable, without accepting that they must be internal” to a brain or body (331).

Another way to think of vehicles is by analogy to Marshall McLuhan’s (1964) famous claim ‘the medium is the message’. While McLuhan’s statement expresses a controversial message, the distinction it draws upon seems straightforward enough. The message of a state or process is its content or meaning, the medium is what carries that content. When dealing with most types of media, the difference between, e.g., a written and a spoken word, and what they have in common in terms of their meaning, is clear enough. If vehicles are media, this helps to highlight the contingency between vehicles and their contents or characters. There is little in common among the material manifestations of a certain piece of news conveyed through a radio, television, newspaper, the internet, ‘private’ thought, etc., so vehicles of very different types can realize or ground the same content or character. Moreover, the same vehicle type, e.g. an inscription or sound, can express different contents in different languages and contexts. It is likely that many mistakes throughout the history of philosophy lie in conflating issues about contents and characters with issues about vehicles, and vice versa.²⁶

Given the contingent nature of the relation between vehicles and contents/characters, if vehicles really play a differential token-explanatory role, then there must be some way of grouping them into types that is independent of their relations to contents or characters. Otherwise, it is hard to see how they could provide non-circular

²⁶ Hurley (1998), Dennett (1991), and Millikan (1984, 1993) all detail evidence to substantiate the claim the vehicle/content confluations are common in philosophy and psychology.

explanations of how and why mental events occur when they do. When dealing with public media, appeal to social convention clearly plays a large role, but this only highlights how mental vehicles may be different. Many insist that it cannot be a matter of convention that some of the enduring material structures that comprise my brain are sometimes experiences, memories, desires, and beliefs. Especially if the contents of mental states and processes are ‘original’ or ‘underived’ in the way Searle (1983), Fodor (2009), and Adams and Aizawa (2008) maintain, it would seem absurd to claim that the messages carried by mental events are conventionally determined.

Let us consider one internalist account of vehicles by looking at the cognitivist understanding of the mind popularized by Fodor, Chomsky, Marr, etc. If one adopts a language-of-thought type account of cognitive processes, whereby there must be criteria for formally or syntactically individuating concept-types independently of meaning, then the notion of a vehicle is straightforward.²⁷ Vehicles are symbolically typed events tokened in some internal mental language, a so-called language of thought. Once one abandons such a view, however, which many think there are very good reasons to do, it becomes much less clear how various enduring material structures are supposed to be identified and classified. It may also be unlikely that different tokens of individual concepts, say, can always be identified by some kind of signature neural or molecular process. But unless there is some non-mental way of classifying vehicles it is unclear exactly what explanatory work they can be claimed to do. A vehicle event need not be identical to a mental event, but it must be able to provide a non-causal explanation of how and why a mental event of the relevant

²⁷ See Fodor (1975) and Field (1978).

content/character type has occurred. It must tell us in virtue of what that event occurred on that occasion. If vehicles can be typed symbolically, then we can explain how and why mental events with various contents occur when they do simply by looking at what symbols are tokened. And we can do this without requiring a full story about how and why particular events have the contents they do. This is why even if content or character fails to supervene on what goes on in the head, many think there are independent reasons for preserving the idea that the vehicles that provide token-explanations of mental events do.

Without resolving all the pertinent questions about vehicles, the distinction between what the individuation of a mental property type depends upon and what a vehicle of an instantiation of that type depends upon, might provide a sufficiently sound basis for proceeding. The latter is what matters when it comes to EM and ECM; whereas traditional content externalism and wide intentionalism concern only the former. These later views are compatible with an internalist account of the vehicles that provide token-explanations of mental events with wide contents or characters. Before looking at the case for active vehicular forms of externalism in some depth, it may help to say a bit more about the relationship between ECM and disjunctive direct realism.

While disjunctive direct realism about perceptual experience entails ECM, ECM does not entail it. There are epistemological considerations driving disjunctivists to position that the objects or properties that can be intentional objects of thought must be literal constituents of the states and processes in question. To see that EM does not require this, it is worth briefly considering Wilson's (2004) version of ECM. While rejecting the 'global externalism' of Dretske, Lycan, Tye, et al., because he thinks pains and bodily sensations

never extend beyond the body, Wilson's TESEE (temporally extended, scaffolded, embodied, and embedded) account of consciousness is a version of ECM. It is similar to wide intentionalism, however, in that Wilson is "arguing for an externalist view of consciousness on the basis of assuming externalism about intentionality" (240).²⁸

Nevertheless, unlike those wide intentionalists who allow that the vehicles of contents and characters are always internal, Wilson (1994) has argued that parts of the environment sometimes play the right computational role to count as vehicles, i.e. they sometimes play the role of symbols in the computational processes that realize mental properties. Insofar as the relation between experience and judgment based on experience is a concern, this shows that proponents of ECM can appeal to developments in externalist theories of content to support their case, so need not rely on the epistemic resources of standard disjunctive direct realist theories. I take this all to show one way that, despite what is often claimed about it, ECM is in some ways less counterintuitive and less of an affront to what many regard as common-sense than wide intentionalist and disjunctivist theories of consciousness. Moreover, as we will later see, just as content or character externalism is compatible with vehicle internalism, vehicle externalism is compatible with internalism about contents and characters.

²⁸ For a sketch of the argument for TESEE, see Wilson (2004, 221). In making a more direct case for extended visual consciousness, Wilson (2010, 277) characterizes his earlier arguments as showing "that at least some of the various phenomena collected under the rubric of consciousness (higher order thought, introspection, and some aspects of attention) fall under the umbrella of the extended mind thesis."

2.6 EM

Before elaborating some of the motivations advanced specifically in favor of ECM, it will be instructive to step back and consider the most influential statement of the argument for active vehicle externalism about non-conscious mental states and processes. Clark and Chalmers (1998) begin by considering a number of cases in which a subject's behavioral competence relies on close interactions with some sort of external artifact or tool. For example, a subject might constantly re-arrange letters while playing scrabble to determine the available moves, press a button to flip Tetris pieces to figure out where to drop them (Kirsh and Maglio 1994), or use pencil and paper to do arithmetic. Clark and Chalmers (1998, 8) take such cases involving 'epistemic actions' to motivate the following principle:

If, as we confront some task, a part of the world functions as a process which, were it to go on in the head, we would have no hesitation in accepting it as part of the cognitive process, then that part of the world is (for that time) part of the cognitive process.

Since it is what something does rather than what it is made of that matters with respect to its status as a vehicle, we are not entitled to rule out non-biological artifacts, which may also play a role as conventional media, as candidate vehicles of mental states and processes in virtue of their being non-biological. That, roughly, is the upshot of this principle. However, I think that if you accept the nomological possibility of non-carbon based cognition or consciousness, then this principle of parity should look plausible independent of considerations about epistemic actions.

The role the parity principle plays in Clark and Chalmers' argument, however, is somewhat controversial. Commenting on how it has been misunderstood and attempting to

clarify its role, Clark (2008, 77) writes “the parity principle thus provided a ‘veil of ignorance’ style test meant to help avoid biochauvinistic prejudice” and goes on to provide the following refined formulation: “for the purposes of identifying the material vehicles of cognitive states and processes we should (normally speaking) ignore the old metabolic boundaries of skin and skull and attend to the computational and functional organization of the problem-solving whole.”²⁹ This is somewhat different than the initial characterization since it no longer requires that a part of the world must function in the same way as a part of the brain does in order to be part of the cognitive system. Chalmers’ (2008, xv) reformulation is closer to the original: “if a state plays the same causal role in the cognitive network as a mental state, then there is a presumption of mentality,” one which cannot be defeated by mere appeal to the location of the thing in that state.³⁰

Insofar as one of the important things about mental states and processes is that they enable us to explain an agent’s behavior, Clark and Chalmers (1998, 8-9) claim that cases of epistemic actions ought to drive us to adopt the following perspective on the problem-solving whole, i.e. the system or subject that cognizes, perceives, and acts:

In these cases, the human organism is linked with an external entity in a two-way interaction, creating a *coupled system* that can be seen as a cognitive system in its own right. All the components in the system play an active causal role, and they jointly govern behavior in the same sort of way that cognition usually does. If we remove the external component the system's behavioral competence will drop, just as it would if we removed part of its brain. Our thesis is that this sort of coupled

²⁹ Functionalists should already be on board with this, but one need not be a functionalist about content or character to accept the parity principle. A functionalist theory of the middle level that connects events with enduring material structures to events with contents or characters will suffice and it may be that any science of mind requires this much functionalism.

³⁰ Note that in these reformulations, unlike the original statement of the parity principle, there is an explicit appeal to the instantiator, i.e. the ‘cognitive network’ and ‘problem-solving whole’. Since this requires that we already have some understanding of what the instantiating system is and isn’t like, it may be a point in favor of the original version that it does not make such an appeal.

process counts equally well as a cognitive process, whether or not it is wholly in the head.

Unlike other forms of externalism, in which monadic relations between a subject and some passive distal features of the world are required, their ‘active externalism’ requires that parts of the subject’s body and parts of the surrounding environment are in constant interactive contact. In other words, they hold that vehicles of some mental states and processes are events that include coupled causal interactions between a brain and the surrounding world. To see the difference, consider an Earthling who was actually transported to Twin Earth would, at least initially, still have thoughts about H₂O, not XYZ. But switching a subject’s environment would immediately lead to changes in the subject’s mental properties in the case of EM.

While the putative external type-individuating determinants in the cases relevant to content or character determination involve time scales that require individual learning, cultural and generational factors, and evolutionary inheritance, the putative environmental determinants when considering EM and ECM exist at a time scale involving the dynamic sensorimotor structure of the perceiving and acting subject. That is why the extended mind thesis, EM, is a form of active vehicle externalism.

Clark and Chalmers argue for EM by appeal to the following thought experiment involving dispositional mental states. Suppose that Inga believes that the Museum of Modern Art (MoMA) is on 53rd Street and that her belief is stored in memory somewhere in the standing structure of her brain. Whenever she hears about an exhibit at the MoMA she wishes to see, she recalls the fact that the MoMA is on 53rd Street. Otto, however, suffers from a mild form of Alzheimer’s disease but has learned to adapt to his declining

ability to rely on his brain to remember things by carrying around a notebook in which he writes down important information. He consults, updates, and revises his notebook as needed. One piece of information recorded in his notebook is that the MoMA is on 53rd Street. Whenever he hears about an exhibit at the MoMA he finds interesting, he looks at his notebook and his doing so is necessary for explaining his arriving at the MoMA successfully on those occasions. Inga clearly has the dispositional belief that the MoMA is on 53rd Street even when she is not consciously entertaining that belief as part of some actual ongoing mental process. What should we say about Otto? If Otto has the dispositional belief that the MoMA is on 53rd Street at all, his belief cannot be located completely in his head; rather, it must be partially located in his notebook.

Since it would be absurd to hold that just any kind of interaction with a notebook would constitute an extended mind, as Hurley (1998, 330) puts it “the vehicle externalist needs to make discriminating appeal to causal spread,” Clark and Chalmers (1998, 17) provide the following extra conditions necessary for determining when a non-bodily item is a candidate for being part of a mental vehicle. The item must be easily accessible and used regularly and reliably by the subject. It must also contain information the subject has previously endorsed and will automatically endorse upon further consultation. With these conditions met, which is much more likely to be the case today with cell phones and tablets than with any paper notebook, Clark and Chalmers’ point is that Otto’s notebook reliably functions for him in a way relevantly similar to the way parts of Inga’s brain function for her with respect to the belief in question.

Clark and Chalmers (1998, 14) also construct an important twin case based on the thought experiment involving Otto and Inga. Suppose that Twin Otto's body is an intrinsic duplicate of Otto's body on some occasion and that the only difference in the rest of the world is that Twin Otto's notebook contains the inscription 'The MoMA is on 51st Street', rather than 'The MoMA is on 53rd Street'. Granting that Otto and his notebook together form the vehicle for his memory of the MoMA's location, it follows that Otto and Twin Otto have different beliefs despite the fact that their bodies are exact duplicates on the occasion in question. Hence, it is possible to fully duplicate the region occupied by a subject's body on some occasion without thereby duplicating all of that subject's mental states and processes. Just as with Twin Earth, duplication of all properties internal to the subject's brain and body does not suffice for duplication of all the subject's mental properties.

Clark and Chalmers' case for EM can be buttressed by considering Rowlands' (1999) defense of what he calls 'environmentalism'. Rowlands' environmentalism is the conjunction of an ontological claim that "[c]ognitive processes are not located exclusively inside the skin of cognizing organisms" and an epistemological claim that "[i]t is not possible to understand the nature of cognitive processes by focusing exclusively on what is occurring inside the skin of cognizing organisms" (22). Obviously, these are distinct claims and the difference between them is at the heart of much of the debate that surrounds EM.³¹ But, let us take it for granted for now that the epistemological claim at least provides evidence for the ontological claim as it pertains to cognitive processes (like remembering

³¹ Consider that Rupert (2004, 2009) endorses the explanatory hypothesis HEMC even though he thinks the ontological claim about the location of vehicles does not follow.

what street the MoMA is on). Still, we are apt to be left with the impression that the Otto/Inga case is somewhat artificial. External devices do not yet tend to function with sufficient parity; not even in cases of a savvy smartphone coupled with a contemporary teenager, say. So it is important to see why we should expect to find incorporation of non-bodily items with some predicable regularity in the situations that we do or likely will encounter.

The subtitle of the first part of Rowlands' (1999, 21-30) early defense of EM is 'Don't work hard, work smart'. There he endorses a principle that is at play in Occam's razor, "Don't multiply *effort* beyond necessity" (21). For Rowlands, the central point here is that we have principled reasons to suspect that physical manipulation of non-bodily structures are often incorporated as parts of mental processes because we have evolved to be as we are in a particular context with various constraints. What he calls "the barking dog principle" states that it is selectively disadvantageous to evolve entirely internal mechanisms and capacities when an adaptive task can be carried out by some combination of manipulation and internal mechanisms (80). If something like this is right, we have a presumptive case for thinking that human and other organisms are often the sorts of things that will extend themselves into the environment absent countervailing influences that select against doing so.³²

³² Cf. Clark's (1989, 63) '007 Principle': "In general, evolved creatures will neither store nor process information in costly ways when they can use the structure of their environment and their operations upon it as a convenient stand-in for the information-processing operations concerned." See also Clark (2008, 118-122, 197) on the hypothesis of cognitive impartiality. Clark (2008, 143-146) further suggests that evidence from studies on change blindness and inattentional awareness provide some support for a stronger hypothesis of motor deference, i.e. that "[o]nline problem solving will tend to defer to perceptuomotor modes of information access. That is, we will often rely on information retrieved from the world even when relevant information is also neurally represented" (197).

Menary's (2007, 2010) and Sutton's (2010) so-called "second-wave" arguments in terms of integration (Menary) and complementarity (Sutton) help further strengthen the case for EM. Building on Clark's (2003) suggestion that we already are 'natural born cyborgs', Menary and Sutton articulate another reason to expect incorporation rather than mere enabling, scaffolding, exploitation, or use. As technology advances, it is likely that we will continue to have more and more opportunities for off-loading cognitive tasks onto reliable devices and artifacts in seamless ways. Thus, even if it often has been selectively advantageous to develop internal mechanisms for adaptive tasks in the past, we should expect these selection pressures will dissipate going forward as we continue to drastically transform the environment we inhabit, and thereby ourselves.

By pointing out differences between epistemic actions involving the manipulation of artifacts and states and processes internal to a brain (screened off by sense organs and muscles), critics of EM have attempted to undermine the claim that there is functional parity (at the relevant level of description). For instance, Sterelny (2004, 246) worries that in contrast to processes that are entirely internal to a subject's body, "perception operates in an environment of active sabotage by other agents [and] often delivers signals that are noisy, somewhat unreliable and functionally ambiguous." In response, Clark (2010a, 60) suggests, plausibly I think, that our brains have evolved in such a way that they have enabled us to implicitly learn that "on a day-to-day basis the chances of these kinds of espionage are sufficiently low that they may be traded against the efficiency gains of treating the perception-involving loop as if it were an inner, relatively noise-free channel, thus allowing them to use the world as 'external memory' (O'Regan 1992; O'Regan and Noë 2001)." The

idea that the world itself can often be relied upon and exploited as a repository of information in ways that are less metabolically costly than internal representation and storage is central to the best overall case for ECM as well.

Before considering ECM explicitly, it may help to step back and summarize just what EM is committed to. Rowlands (2009, 53) aptly characterizes EM as an ontic thesis about the composition of some mental processes that is partial and contingent. Qua ontic thesis, EM has epistemic implications about how best to understand mental processes, but those implications are shared by the weaker thesis that minds are simply causally embedded in the environment in significant and surprising ways. EM is a compositional thesis about the parts of mental processes. It is partial because it allows that what is internal to the brain always makes a necessary contribution as well. In other words, it is compatible with the claim that all (human) mental processes have brain internal parts. Departing a bit from Rowlands' characterization, I also take EM to apply to mental states, like Otto's dispositional belief about the location of the MoMA. The inscription in Otto's notebook is merely part of the vehicle of his dispositional belief, but it is not the state itself. Finally, Rowlands claims that to preserve the functionalist presumptions shared by most versions of EM, and by many of the theories of perception I will be surveying next, both *de dicto*, "necessarily, some mental processes are partly constituted by processes of environmental manipulation", and *de re*, "some mental processes are necessarily constituted by processes of environmental manipulation", versions of EM should be eschewed (54).³³ Hence, EM and its species make only contingent claims about the locations of token mental processes

³³ I'm not sure the *de re* version is so bad, particularly if (i) one is focusing on fine-grained types in order to extrapolate existential claims, and (ii) one can furnish a reasonable account of *de re* modality applicable to processes (and states).

(and states) of various types. Inga's dispositional belief has exactly the same content as Otto's, after all; but it stays in her head.

Rowlands' (2010, 83) preferred version of EM is what he calls the amalgamated mind thesis, i.e. that "[c]ognitive processes are an *amalgam* of neural structures and processes, bodily structures and processes, and environmental structures and processes." It is clear that Rowlands takes that to be true of some cognitive processes, not all, and so the amalgamated mind thesis fits the characterization as an ontic thesis about the composition of mental processes that is partial and contingent.

2.7 ECM

Clark and Chalmers (1998) and Clark (2008, 2009, 2012) resist the notion that EM, active vehicle externalism, ever holds as a claim about conscious mental states or processes.³⁴ Anticipating likely objections to their version of EM, they write:

Some find this sort of externalism unpalatable. One reason may be that many identify the cognitive with the conscious, and it seems far from plausible that consciousness extends outside the head. But not every cognitive process, at least on standard usage, is a conscious process. It is widely accepted that all sorts of processes beyond the borders of consciousness play a crucial role in cognitive processing: in the retrieval of memories, linguistic processes, and skill acquisition, for example. So the mere fact that external processes are external where consciousness is internal is no reason to deny that those processes are cognitive (1998, 10).

Clark and Chalmers' point about not conflating the cognitive and the conscious is fair enough. Given that EM is a partial claim about the vehicles of some mental states and

³⁴ Clark (1999, 2009), however, provides sympathetic expositions of the arguments for ECM via EVC. I offer a refined version of the strongest of these arguments in chapter 5.

processes, it does not by itself entail anything about the vehicles of conscious mental states and processes.³⁵ It may be that cognition sometimes extends into the world, but consciousness never does.

Nevertheless, I suspect that at least some of the resistance to ECM is based on misplaced concerns about consciousness that at best support property dualism or emergentism rather than internalism. After all, if a subject is simply located wherever that subject's mind is, the vehicles of conscious states and processes may sometimes extend beyond the subject's body while still being internal to the subject. In this way vehicle externalism is compatible with content or character internalism; I will examine this further in the section on extended selves at the beginning of the next chapter. Throughout this dissertation, I will also revisit whether there are compelling reasons for restricting EM to the cognitive but unconscious, but for now let us turn to the theories of perception that are often taken to imply ECM, survey what motivates them, and ask whether they really entail ECM.

While much of the initial discussion over EM focused on cognition, to begin to motivate ECM it is best to shift our focus to perception. While many cognitivists view perception as a species of cognition, if not simply an encapsulated input process that functions for cognition, the phenomenal characters of perceptual states and processes are

³⁵ Noë and Thompson (2004, 94) claim that “[e]xternalism about the vehicles of content... entail[s] that neural systems are not sufficient for consciousness”, but Clark (2009, 967) is right to point out that no such implication holds directly. One way to make the connection might be by appeal to Chemero’s (2009) ‘radical embodied cognitive science’, according to which “intentionality and consciousness are inseparable” (202). Or perhaps it can be shown that EM entails ECM because they are motivated by similar but more general externalist considerations. In any case, many proponents of EM who argue from similar starting points as Clark and Chalmers are less reticent to accept ECM, e.g. Wilson (2004, 214-261), Rowlands (1999, 2003, 2010), and Hurley (1998, 2010).

proprietary to and closely associated with different sensory organs. Moreover, while cognition and perception can occur consciously or unconsciously, talk of qualia, sense-data, what is given or present in experience, etc. suggest that perceptual properties are more paradigmatically conscious. Even though we are most familiar with our own conscious thoughts, we are happy with the idea, especially post-Freud, that much of our thinking is unconscious. On the other hand, empirical demonstrations of unconscious perception remain surprising, e.g. in masked priming experiments or in patients with blindsight. Since there are quite likely many forms of unconscious sensation and perception, we need to be clear that we are talking about perceptual states and processes that are conscious.

Two dimensions for distinguishing the accounts I will survey that lend support for ECM from the more orthodox cognitivist accounts concern the function of perception and the nature of systems sufficient to implement or realize perceptual properties. Orthodox accounts of perception take it that the primary function of perceptual states and processes is to furnish the subject with detailed representations of the world. These representations can then be used, along with others, to guide thought and action. Orthodox accounts conceive of perception as an input process with each of the different sense modalities, e.g. vision, touch, smell, etc., providing different kinds information about the external world. It is only once the different sources of perceptual information have been synthesized and cognized that they are apt to guide (non-reflexive) action. In contrast, the accounts of perception that support ECM all take the primary function of perception to be guiding real-time action in the world. By rejecting the idea that perception is primarily for the construction of detailed models and emphasizing a more direct link between perception

and action, these alternative accounts also reject taking cognition to be central and autonomous while delegating perception and action to the periphery. Instead, crisscrossing interdependencies suggest that cognition is much more intricately dependent upon unfolding sensorimotor activity than cognitivist orthodoxy would have it.

Most scientific and philosophical work on perception has focused on visual perception. This is unsurprising given the centrality of vision to humans. Orthodox accounts take the starting point of vision to be the retinal image at a moment, which is incredibly sparse and upside down, and so take the task of perceptual processing to be solving ‘inverse optics problems’ by constructing models of the world on the basis of that sparse incoming optical information. Perceptual processing must then ‘fill-in’ missing information in order for the subject to have an updated representation of the state of the world. This involves using earlier visual information and background assumptions to constantly draw inferences about what the world is like.

With a sense of the contrasting orthodoxy, let us now examine in some depth four unorthodox accounts of perception – the ecological approach, the enactive approach, the dynamic sensorimotor theory, and the two-level interdependence theory. Hopefully, this will help readers appreciate some of the reasons for taking ECM scientifically and philosophically seriously. However, as mentioned earlier, I do not intend to endorse any of these theories wholesale.

2.8 Ecological theories of perception

To begin to appreciate this alternative to orthodox cognitivist accounts of visual perception, let us look at the work of Gibson (1966, 1979/1986). Gibson developed his views in tandem with but largely outside of the mainstream cognitivist tradition. Indeed, elements of Gibson's ecological theory have been harshly criticized by well-known cognitivists such as Marr (1982), and Fodor and Pylyshyn (1981). I have already mentioned that the starting point for orthodox cognitivist accounts of visual perception like Marr's is the retinal image. The retinal image, determined by the mosaic of firing rods and cones, is sparse and inverted. So, according to orthodox approaches, it must be supplemented with inferential activity internal to the brain in order to 'fill-in' what is missing and realize visual experience. In contrast, the starting point for vision according to Gibson's ecological approach is the so-called 'ambient optic array' that surrounds the seeing subject.

What is the ambient optic array? In any visible environment, light from the sun and other sources fills the air creating a steady state of reverberating energy. At any given point in space and time in a visible environment, there is also a convergence of light rays of various wavelengths and intensities reflecting off different surfaces. The optic array can be defined as this spatial pattern of light and a subject's ambient optic array as the pattern of surrounding light that is accessible through movement. The light that reaches a subject's eyes is reflected off surfaces of various shapes, textures, and colors, and as the subject moves around in the world some structural properties of the light remains invariant, while others vary more or less systematically. Since the topological structure of the ambient optic array specifies the sources of the objects and properties in it, it is simply there to be

perceived directly and non-inferentially. The takeaway point is that a subject's ambient optic array is an external information-bearing structure – one that can be quantified, tracked, and studied – and one that is not informationally impoverished in the way retinal images are.³⁶

Already we may begin to appreciate why this view is likely to be amenable to externalism about visual phenomenal characters as well as the vehicles of visual experiences. According to Gibson (1972, 293), "[i]nvariance in the ambient optic array is not constructed or deduced; it is there to be discovered". While subjects must ordinarily move around to 'pick-up' or 'resonate' invariant information, the information that determines the contents and characters of visual experience is already present in the environment. Moreover, by starting with the ambient optic array, instead of the retinal image, we are more inclined to understand the brains and bodies of perceivers as tightly coupled to the environment in some kind of high-bandwidth interaction. Since perception occurs through movement, sense organs, like the eyes, are not conceived as low-bandwidth filters that naturally demarcate the outer bounds of perceptual systems. While there will be more to say about this point going forward, it suggests that we should understand the visual system itself and the vehicles of visual experiences as spatially extending beyond the body. So, for instance, Gibson (1979/1986, 1) suggests, "natural vision depends on the eyes in the head on a body supported by the ground, the brain being only the central organ of a complete visual system." Even if one maintains that the vehicles of visual experience must

³⁶ So, "[i]nstead of adopting the idea that visual representations result from the hierarchical processing of impoverished sensory stimuli, Gibson argued that information reaching the senses is already structured in ways that specify the layout of the environment to an animal" (Gangopadhyay and Kiverstein 2009, 64).

be intrinsic to an instantiator, this suggests they may still fail to be intrinsic to any brain or body.

We have tentatively seen how Gibson's theory of perception might suggest that the subpersonal system sufficient to realize visual states and processes may not be entirely contained in subjects' brains. In contrast to cognitivist orthodoxy, Gibson also emphasized the extent to which vision is action-dependent and so its main function is not the representational one of producing internal maps of the world that can be stored and used in later cognition. Rather, he thinks perception should be understood as an achievement of an animal successfully navigating its environment. Perception functions to keep subjects in contact with the relevant parts of their local niche, and thereby directly guides action within that niche. If visual perception is primarily for guiding action, then it is not primarily a kind of information processing over discrete explicit representations (or so Gibson reasoned). Since the information is already present in the structure of the ambient optic array, to be picked-up and exploited when and as needed, detailed representations are likely unnecessary. Nevertheless, it is natural to wonder whether 'resonating' invariant optical structures require some kind of representing. Most commenters, even sympathetic commenters, regard some of Gibson's rhetorical hostility to representational and information-processing accounts of visual perception as somewhat overblown and unwarranted.

We can now appreciate how in Gibson's ecological theory of perception, as with the remaining theories I will discuss, we find an "attenuation of representation combined with augmentation of action" (Rowlands 2010, 37). The action-oriented dimension of this

account can be fleshed out by considering what Gibson had to say about the contents and characters of visual experience. What we perceive, according to Gibson, are affordances, i.e. information about possible ways to move, act, or behave in the context. This information is available in the invariant structures of the optic array to be perceived by organisms in a direct and non-inferential way. Affordances are supposed to be relational properties of objects and surfaces in the environment that specify the types of action that may be taken. For example, we may see a piece of fruit as edible or a tree as climbable. Importantly, what affords action varies across animals and even individuals. What looks climbable to me may not look so to a small child or to a mouse. In this way, affordances are both in the environment and specific to different types of perceivers. Ecological laws relate affordances to perceivers and it is “in virtue of ecological laws, [that] particular patterns in the ambient optic array specify the presence of affordances” (Chemero 2009, 111).

In a famous passage, Gibson (1979/1986, 129) further explains what affordances are as follows:

An affordance cuts across the dichotomy of subjective-objective and helps us to understand its inadequacy. It is equally a fact of the environment and a fact of behavior. It is both physical and psychical, yet neither. An affordance points both ways, to the environment and to the observer.

Gibson seems to be claiming that affordances are both the phenomenal characters of visual experiences and relational properties of structures, surfaces, objects, etc. in the external environment. Affordances are properties of the environment that can be perceived directly and thereby constitute or determine the character of perceptual experience.

So how exactly are invariants in the optic array picked up by subjects and perceived as affordances? Ordinarily, Gibson would no doubt insist, by moving around in it. But for all this emphasis on action, Gibson seems to have regarded the dependence of perception on action on this account as merely causal or instrumental rather than constitutive. Gibson would allow, for instance, that a paralyzed subject who was moved in just the right ways could have exactly the same perceptual experiences as a subject whose movements were self-generated.³⁷ This is because he thought that perception is direct and not usefully understood in terms of information processing. Thus, efferent motor signals, or copies thereof, play no constitutive role in realizing experience according to the approach.³⁸ This feature of Gibson's account in some ways makes it more like wide intentionalism than ECM. It is also at the heart of Thompson's (1995, 229) critique: "I think Gibson and his followers have not fully embraced the animal-dependence of the environment because of their desire to uphold direct realism against idealism in perceptual theory."

After investigating this issue at some length, Mossio and Taraborelli (2008, 19) conclude "ecological theories conceive invariants as properties of the sensory stimulation that remain constant through transformation produced by the perceiver's self-motion." In other words, perceptual invariants, for Gibson and his followers, are entirely 'transformation-specific'. In contrast, the three other unorthodox accounts of perception I will discuss all allow that we also perceive 'motor-specific' invariants "by extracting regularities from co-occurring efferent motor and reafferent sensory patterns" (19). Mossio and Taraborelli discuss a number of "cases in which different motion patterns coupled with

³⁷ For further discussion of this point, see Hurley (1998, 430-435, and 2001).

³⁸ See Hafed and Krauzlis (2006) on how efferent motor commands help perceivers to resolve spatial ambiguity in the retinal signal though.

the same sensory stimulation result in difference in performance” (18). For instance, experiments with sensory substitution devices and inversion lens show that self-generated movement is necessary for perceptual adaptation, see, e.g., Bach-y-Rita (1972) and Held (1965). Moreover, Wexler and Van Boxtel (2005) review evidence that subjects’ ability to see the depths and shapes of objects and surfaces by motion parallax sometimes depends on self-generated movement. The optical flow changes that would be brought about by such movements are not sufficient in all cases, contrary to what proponents of Gibson’s ecological approach usually maintain.

In responding to commentary by Pascal and O’Regan (2008) on their earlier paper, Taraborelli and Mossio (2008) contrast enactivism and dynamic sensorimotor theories with Gibson’s ecological approach by claiming that the former both “take an explicit *interactivist* position, which sees the perceiver’s motor activity as a determinant of perceptual information itself.” This emphasis on agents’ motor activities and movements in helping to constitute its visual experiences is distinctive of the two-level interdependence view as well. While Gibson’s account is clearly a version of externalism, since it holds that perception only depends instrumentally on movement and motor activity, it may be best seen as a version of wide intentionalism. Despite his hostility to representations, it may be that the best interpretation of Gibson would be to combine a form of vehicle internalism with a teleological theory of content. Teleosemantic theorists, with their focus on real-world evolutionary constraints, can accommodate the idea that perception functions primarily to guide action rather than to construct storable models of the world. Allowing that motor-specific invariances, which co-occurrent afferent and efferent signals help extract,

constitutively determine some types of experiences makes the case for vehicle externalism far more compelling.

2.9 Enactivism

In contrast to Gibson, the three remaining unorthodox accounts of perception allow that “both movement through real environments by whole organisms and brain activity play essential roles in extracting information from the environment and enabling a creature to have a perceptual perspective” (Hurley 2001, 20). Gangopadhyay and Kiverstein (2009, 64) characterize enactive theories in terms of the following two commitments: “(1) Perception and action are interdependent processes. (2) The vehicles of perception are distributed across brain, body and world.” If this is right and perception is sometimes conscious, then enactivism entails ECM; in fact, (2) alone does. While I sometimes use ‘enactivism’ in roughly this sense – as a convenient label for theories of visual perception that entail ECM – I usually use it to mean a more specific account associated with Thompson and Varela.

Thompson and Varela’s enactivism is not simply a theory of perception, but an account of cognition and emotion as well. Moreover, the underlying motivations are deeply tied to Varela’s earlier work with Maturana on the concept of autopoiesis in mathematical biology.³⁹ Central to the enactivist framework overall then is the promise of providing a unifying, though non-reductive, account of mind, life, and consciousness. But it is not

³⁹ Thompson (2007, 51): “For the enactive approach it is the autonomy perspective on natural cognitive agents that remains the reference point for understanding mind and life, not a predefined input-output task structure.”

entirely clear to me that their view entails ECM, despite occasional explicit claims to the contrary.⁴⁰ For one thing, Thompson (2007) and Varela (1999) commit themselves to a Husserlian retentionalist account of time consciousness that may be at odds with vehicle externalism.⁴¹ For another, there is an idealist and constructivist tenor to their account that may undermine the idea that conscious experiences have locatable vehicles at all. Let us look a bit more closely at what they say.

What exactly does ‘enactive’ mean? Obviously, it contrasts with the meaning of ‘inactive’, so we can infer that enactive perception requires action. It does not simply involve the passive reception of images and contents. Moreover, in the case of visual experience, it connotes that the relevant states and processes are acted out by agents. Varela, Thompson, and Rosch (1991, 173) emphasize two central tenets of their approach: “1) perception consists of perceptually guided action and 2) cognitive structures emerge from the recurrent sensorimotor patterns that enable action to be perceptually guided.” So, for enactivists, perception is a kind of action and the vehicles of cognition depend constitutively on structural coupling between sensory and motor activity just as the vehicles of perception and action do.

Summarizing the main points of Varela et al. (1991), Thompson (2007, 13) further characterizes their enactive approach in terms of five tenets: (1) “Living beings are autonomous agents”, (2) “the nervous system... actively generates and maintains its own

⁴⁰ E.g. Thompson and Varela (2001, 425): “we conjecture that consciousness depends crucially on the manner in which brain dynamics are embedded in the somatic and environmental context of the animal’s life, and therefore that there may be no such thing as a minimal internal neural correlate whose intrinsic properties are sufficient to produce conscious experience... the processes crucial for consciousness cut across the brain-body-world divisions rather than being located simply in the head.”

⁴¹ Husserl’s influential tri-partite account of temporal consciousness was first explicated in his 1905 lecture and developed in subsequent work, see Husserl (1991).

coherent and meaningful patterns of activity”, (3) “cognition is the exercise of skillful know-how”, (4) “a cognitive being’s world is... a relational domain enacted”, and (5) “experience is not an epiphenomenal side issue, but... needs to be investigated in a careful phenomenological manner.” Thompson’s characterization is somewhat broader than the passage I quoted last paragraph. Tenets (1)-(3) are tied to how the enactive approach treats human agents as distinctive types of autopoietic systems.⁴² Let us focus a bit further on tenets (4) and (5).

Tenet (4) is what seems to entail EM. For the world or environment to be relational and enacted, for enactivists, is for it to emerge or come into being along with the states and processes of the agent. The notion of emergence relevant here is what Thompson (2007, 60-65, 417-441) calls “dynamic co-emergence.” In cases of dynamic co-emergence, a whole is not reducible to its parts nor can those parts be understood independently of the whole they compose. The membrane and reaction network of a cell, in even the simplest of autopoietic systems, and the molecules that compose them, dynamically co-emerge, according to Thompson.⁴³ In just this way, so do an agent and its niche. In the case of vision, the niche will include parts of the ambient optic array. Tenet (5) is connected to Thompson and Varela’s attempt to understand the brain’s role in consciousness in naturalistic yet anti-functional terms, what they call ‘neurophenomenology’.⁴⁴ Pascal and

⁴² Thompson (2007, 49): “Whereas autopoietic closure brings forth a minimal “bodily self” at the level of cellular metabolism, sensorimotor closure produces a “sensorimotor self” at the level of perception and action.”

⁴³ Thompson (2007, 65): “The membrane and reaction network (as well as the molecules that compose them) do not pre-exist as independent entities. Rather, they co-emerge through the integrative, metabolic relation to each other.”

⁴⁴ The enactive approach should be distinguished from other ‘biological’ theories of consciousness, e.g. Searle’s and Block’s, which are resolutely internalist. Still, this might point at a way in which enactivism may be less clearly a version of ECM than the dynamic sensorimotor and two-level interdependence theories,

O'Regan (2008) object that this element of enactivism is based on a problematic commitment to a kind of constructivist idealism. The other three accounts I have in my purview are more resolutely realist. But can we make sense of the idea that visual experiences are constituted by movement and motor activity, without having to say agents make their own worlds? O'Regan, Noë, and Hurley all develop accounts that attempt to do just this.

2.10 Dynamic sensorimotor theories

Despite some commentators conflating dynamic sensorimotor theories with Gibson's ecological approach, e.g. Pylyshyn (2001) and Scholl and Simons (2001), as we have seen, the former embrace a modally deeper understanding of the ways in which perceptual invariants are action-dependent. O'Regan and Noë (2001, 1019) write: "whereas Gibson stresses the use of sensorimotor invariants as sources of information, we are stressing the idea that sensorimotor invariants are part of what constitute sensations and perceptual content."

Dynamic sensorimotor theories are most notably defended by O'Regan (2011) and Noë (2004, 2009). As an alternative to orthodox cognitivist account of visual perception, dynamic sensorimotor theories are like the others in being wary of the idea that minds must be understood by reference to internal representations. O'Regan and Noë (2001a,

which are compatible with functionalism about consciousness. According to Rowlands (2009, 57), the sort of functionalism presupposed by most versions of EM and enactivism maintain: "(1) cognitive processes are functional roles, (2) the vehicles of cognition are whatever structures realize those roles, and (3) as a matter of contingent fact, some of those structures are external to the skin of the subject of the cognitive processes." This characterization allows that mental properties of the same type can be instantiated on different occasions and only one have a vehicle that extends beyond the subject's body.

946) express this when they write: “visual experience does not arise because an internal representation of the world is activated in some brain area... there is no “re”-presentation of the world inside the brain: the only pictorial or 3D version is the real outside version.”

Dynamic sensorimotor theories hold that perceptual states and processes non-causally depend on the subject’s expectations of how afferent sensory stimulation is likely to change in response to a subject’s movements as well as the independent movements of things in the environment. Some of the relevant sensorimotor relations hold only if particular states of the non-bodily environment either remain constant or co-vary systematically with the changing states of a subject’s brain and body. Without the relevant parts of the non-bodily world present to preserve the necessary dynamic sensorimotor relations, many of the most pedestrian conscious perceptual experiences that ever have occurred could not have. This is why this view seems, more so than ecological and enactive accounts, to clearly entail ECM. Consider what Noë and O’Regan (2002, 567) say about the neural basis of visual consciousness:

[V]ision is not a process in the brain. Though the brain is necessary for vision, neural processes are not, in themselves, sufficient to produce seeing. Instead, we claim that seeing is an exploratory activity mediated by the animal’s mastery of sensorimotor contingencies. That is, seeing is a skill-based activity of environmental exploration. Visual experience is not something that happens *in* individuals. It is something that they *do*.

Here the commitment to ECM, EVC in fact, is quite explicit.

In the case of vision, to see is to explore the optical environment by means of one’s understanding or knowledge of how retinal stimulation changes with movement, particularly eye, head, and neck movements. Other sense modalities are individuated by different sets of sensorimotor contingencies, as opposed to via proprietary “specific nerve

energies” (Müller 1843/2003). This feature of the account purportedly enables them to explain how sensory substitution devices are able to replicate the phenomenal characters associated with one modality without routing through the ordinary channels. Moreover, it promises to provide the resources for a unified explanation of both inter and intra modal differences in phenomenal characters.

Dynamic sensorimotor theorists also appeal to sensorimotor contingencies, more specifically, subjects grasp and use thereof, to explain distinctions in phenomenal character within a sense modality. For instance, while, according to O’Regan and Noë (2001a, 942), the “visual quality of shape is precisely the set of all potential distortions the shape undergoes when it is moved relative to us, or when we move relative to it”, the visual quality associated with seeing green or red would each be characterized by a different set of potential distortions. Here, the aspirations for an all-encompassing account of perception are similar to the enactivists. And while Pascal and O’Regan (2008) object to the idealist underpinnings of Thompson and Varela’s enactivism, Taraborelli and Mossio (2008) argue persuasively that both views accept that perception is “built upon structural coupling between sensory patterns and motor activity.” The structural congruence that emerges between what’s within an agent’s body and its non-bodily niche in virtue of such sensorimotor coupling is key to how both views support vehicle externalism.

When it comes to the question of what perceivers perceive, Myin and O’Regan (2002, 30) list the following properties of perceptual experiences that need explaining: ongoingness, forcible presence, ineffability, and subjectivity. O’Regan and Noë’s (2001a, 1012) dynamic sensorimotor account explains ongoingness and forcible presence in terms

of “bodiliness” and “grabbiness”. Bodiliness is measured in terms of the amount of sensory change is a result of bodily movement and grabbiness is measured in terms of how exogenous sensory change attracts the subject’s attention. Ineffability is explained by our inability to ever fully describe the operative but implicit practical knowledge claimed to be constitutive of any experience. Subjectivity is explained by the fact that interacting with something and thereby becoming conscious of it guarantees that consciousness is always for the subject.

Thompson (2007) criticizes this account’s explanation of subjectivity and notes that both ongoingness and forcible presence depend on subjectivity as well. Moreover, the account does not explain intransitive consciousness, since it simply identifies consciousness with attention and transitive (object-directed) consciousness. In contrast, the enactive view accounts for subjectivity in terms of agency and autonomy. Even if the enactive account falters elsewhere, I think Thompson is right that “a complete account of perceptual experience requires an account of nonintentional (intransitive, non-object-directed), prereflective bodily self-consciousness” (265). And it is far from clear that the sensorimotor resources Noë, O’Regan, and Myin appeal to are sufficient for doing so.

Lastly, while Noë and O’Regan have often claimed that knowledge of sensorimotor contingencies must be acted out or enacted, they sometimes deny this. In other words, they claim that merely being in possession of such knowledge is sufficient. Critics, e.g. Aizawa (2010) and Prinz (2006), noting this difference in the way the dynamic sensorimotor theory is articulated, have distinguished strong and weak versions of the view. Whereas strong versions require that the vehicles of experience are sometimes constitutively determined by

bodily exercises of sensorimotor knowledge, weak versions merely require that the vehicles of experience sometimes constitutively depend on possession of sensorimotor knowledge. While the strong version requires active coupling with the environment, and so more clearly supports EVC, the weak version is likely compatible with an internalist and representationalist account. Though combined with the idea that knowledge is factive, even the weak version can likely vindicate some version of externalism.

2.11 Two-level interdependence view

Hurley (1998) defends a two-level interdependence view of the relations between the contents and characters of perceptual, cognitive, and intentional states and processes and the vehicles thereof. According to Hurley (1998, 6), “the contents of experiences and intentions are both functions of a complex dynamic system of relations between input and output, including both internal and external feedback loops.” This contrasts with the input-output picture, according to which perception is regarded as input from world to mind and action is regarded as output from world to mind, with cognition sandwiched in between. Instead, perception, cognition, and action are interdependent and this interdependence is crucial to understanding how we have a subjectively unified conscious perspective when we perceive. On the two-level interdependence view, the mind is horizontally rather than vertically modular, where “[e]ach horizontal module or layer is a content-specific system that loops dynamically through internal sensory and motor processes well as through the environment” (22).

For Hurley (1998), consciousness entails having a perspective and some kind of agency. Tracking interdependencies between what is perceived and what is done requires both consciousness of one's own agency and the ability to instigate and control one's own actions. Consciousness/awareness of one's own agency, whether conceptual or non-conceptual, is not sufficient for explaining the unity of consciousness, i.e. how, when, and why different states and processes are (synchronically or diachronically) co-conscious. Since agency presupposes unity, it is not taken to explain, determine, or account for perspectival self-consciousness on its own. This marks a difference between her account and Thompson and Varela's enactivism. With respect to the ability to instigate and control one's own actions, Hurley thinks there is a type of intentional motor access that can be deployed which is sufficient for a kind of minimal unified perspectival self-consciousness; access that does not require the deployment of a concept of self.

All of the non-orthodox theories discussed require some account of intentional, cognitive, or personal-level integration. Just displaying some sensitivity to sensorimotor contingencies is insufficient for consciousness; after all, roombas, drones, missile defense systems, etc. manage to do so without any conscious awareness. To address this issue, Noë and O'Regan (2002, 569) require that in the case of conscious vision, as opposed to mere visual sensitivity, a subject "integrates its coupling behavior with its broader capacities for thought and rationally guided actions." On Hurley's view, rational integration at the level of conceptual thought is not required. Instead, directly action-guiding access is equally important as any kind of cognitive access we may have to such contents when it comes to determining various conscious features of our perceptual states and processes. Moreover,

since non-human animals engage in normatively accessible behavior on the basis of perception and can perceive the world consciously, it is likely that practical integration of sensorimotor contingencies for real-time action guidance may be sufficient on its own for some forms of consciousness. Intentional or motor access may require a kind of self-reflexivity, but not the kind of conceptual sophistication cognitive integration requires.

That marks the biggest difference between the two-level interdependence view and the dynamic sensorimotor and enactivist accounts. Unlike the ecological accounts, all three allow that we can see motor-specific invariances and that motor commands and efference copies, not simply sensory reafference caused by movement, play a constitutive role in realizing visual experience. Hurley is also more consistent in holding that action and bodily movement itself, and not merely knowledge of sensorimotor contingencies, is necessary for many types of perception. The best case for EVC combines the ideas that both motor-specific transformations and bodily movements are sometimes necessary for visual perception.

3. OBJECTIONS AND GROUNDING

First, I consider whether EM and ECM have problematic implications about the boundaries of the subject or self and whether they rely on problematic criteria for determining what is truly cognitive or mental or conscious. I argue that they do not. Next, I evaluate whether the hodge-podge of empirical evidence presented by proponents of EM and ECM, which clearly demonstrates some kind of explanatorily interesting dependence on events that involve coupling among brain, body, and environment, ever really demonstrates that mental events are constitutively grounded by events partially located in the world external to the subject's body. This objection becomes especially significant since there seem to be alternative and theoretically more conservative explanations of environmental dependency or scaffolding available. At the end of this chapter, I explicate an account of constitutive grounding, one which is distinct from any notion of causal determination and is suitable for spatiotemporally locating mental and non-mental events alike.

3.1 Extended Minds, Extended Selves?

The first objection I will consider has to do with personal identity and the boundaries of the self. Both Hurley (1998, 2010) and Clark and Chalmers (1998), for example, accept that active vehicle externalism about any dimension of mind implies that persons or selves are not always confined to their bodies. Olson (2011) and Baker (2009), noting this and elsewhere having argued that persons are either identical to animals or

constituted by bodies, have mounted this into an objection to EM. They reason that since selves cannot extend beyond their bodies, if EM implies that they ever do, it must be mistaken.

While I think that extended selves should have some plausibility for those who accept EM, I do not think that this implication is demanded. Perhaps the most straightforward reason for denying that it does is articulated by Rowlands (2010, 8-10). As a claim about the vehicles of some token mental events, EM is in fact compatible with the non-existence of minds and selves. To wit, even if Hume is right that minds are no more than bundles of mental states and process, and hence are not really independently existing things as selves are ordinarily taken to be, this would not threaten EM.

More significantly, Wilson (2004, 141-143) explicitly rejects extended selves and maintains that the bodily or organismic boundaries ought to be retained. Wilson explains his rationale for doing so thusly: “there is a basis for marking out individuals as the subjects of properties, even those properties with wide realizations” (142). Whether this is a viable strategy depends on how we understand vehicles or realizers. Wilson understands vehicles as ontologically akin to the way traditional content externalists and wide intentionalists understand intentional content and phenomenal character, i.e. as monadic relational properties. But just as being sunburnt or winning a game of backgammon does not alter the location of the instantiator of such properties, if vehicles are monadic relational properties, then even if they depend constitutively on the environment, they need not alter the subject’s location. This is connected to the reasons why Putnam’s slogan about meaning

not being in the head does not really follow from accepting that mental content is sometimes wide.

On the other hand, if one accepts Hurley's (1998) token-explanatory conception of vehicles, according to which the boundaries of vehicles and duplication are "nonaccidentally related" (331), then we cannot say that vehicles are monadic relational properties. Rather, they will have to be intrinsic to their instantiator, whatever that is. While there still may be some room for debate depending upon how we understand personal identity, this conception of vehicles in conjunction with EM and a psychological account of personal identity will entail that subjects sometimes extend beyond their bodies.⁴⁵ One benefit of this account is that it has the resources for retaining a kind of internalism about content and character. Even if the contents or characters of mental events fail to supervene on what is intrinsic to a subject's brain or body, they may still supervene on what is intrinsic to the subject on this view. And this might be sufficient to capture the sense in which conscious states and processes are always for a subject. By rejecting the close relation between vehicles and duplicability, as Wilson seems to do, there is no longer the option of retaining this kind of internalist intuition.

⁴⁵ It is notable that James considers whether something like EVC is implausible because it entails that two people's minds can sometimes overlap, e.g. if they are both appropriately coupled to some object they are both looking at. See Chemero (2009, 140-141) and Heft (2001) for further discussion. One may still wish to hold an externalist account of the persistence of persons or selves, however, and perhaps the best way to do so would be to accept an account like Nozick's (1981) closest continuer view.

3.2 Demarcating the Mental

The next common objection concerns whether the extended states and processes in question are ever really mental. Adams and Aizawa (2001, 2008) and Fodor (2009), for instance, have argued that all and only mental events have underived contents and this is what marks them as mental. Manipulations of artifacts like Otto's notebook at best involve derived contents and so fail to meet a necessary condition for something's being mental. In response to this objection, Clark (2010a, 48) sketches an example of imagining Venn Diagrams to solve a logic problem, for which "the set-theoretic meaning of the overlaps between... two intersecting Euler circles is a matter of convention." Clark's point is that the contents of these images are clearly as derived as the contents of Otto's notebook, yet this does not disqualify them from serving as parts of cognitive tasks. Moreover, a mental demarcation condition in terms of underived contents presupposes a controversial psychosemantic theory. As Adams and Aizawa (2008, 55) acknowledge, "[i]t remains unclear just exactly what naturalistic conditions give rise to non-derived content; hence it remains correspondingly unclear just exactly what objects bear non-derived content." For this reason, we should be suspicious of the notion that there is some mark of the mental which only biological matter can have.

In any case, whether or not conscious states and processes have underived content, they clearly have phenomenal character and having phenomenal character is sufficient for being mental. While I think Clark's response to the mental demarcation objection is plausible, Clark and Chalmers' version of EM, which is restricted to unconscious mental phenomena, may face more difficulty with it than any version of ECM. Unlike the

inscription in Otto's notebook, and many other examples Clark discusses, the relevant structures in the ambient optic array that help ground visual experiences don't generally have derived contents, which it might be natural to assume must be re-processed and re-presented internally in order to serve cognition. It may be hard to comprehend how an extended event can have phenomenal character at all, but this is equally a worry for internal events. Still, if the target of one's thesis is consciousness, there can be no worry about whether what one is talking about is really mental. The question of why anything beyond the brain is necessary for grounding such events remains pressing, but not whether the events themselves are mental.

3.3 Mistaking Evidence of Coupling for Evidence of Constitution

By far the most frequently discussed objection to EM is Adam and Aizawa's (2001, 2008) allegation that proponents of it are making an invalid or fallacious inference when they take evidence of coupled causal interactions between a subject's body and the surrounding world to demonstrate constitutive dependence on events in the world. In addition to Clark and Chalmers (1998), Adams and Aizawa (2001, 408) single out van Gelder (1995), Haugeland (1998), Rowlands (1999), Menary (2006), Noë (2004), Rockwell (2005), Gibbs (2001), and Wilson (2004) as all committing the "constitution coupling fallacy" at various places. Here is how they characterize the fallacy:

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

To show why coupling is not generally sufficient for constitution, Adams and Aizawa discuss the example of the cooling of an air conditioning system. While all of the components of the system interact and are tightly coupled, it would be a mistake to say that all of the components play a constitutive role in determining the cooling properties of the system. Rather the cooling of air can be localized to the evaporation coil, whereas the ducts, vents, and thermostat play no essential role. However, the coupling in question is not, as the quoted passage above suggests, supposed to be between some cognitive agent and things that are not part of that cognitive agent, but between what probably ends up counting as different parts of a single cognitive agent.⁴⁶ Still, not everything that is part of an agent's body is constitutively relevant to that agent's cognitive states and processes, even though everything within an agent's body is, to some significant degree, causally coupled. Especially, since I've distinguished EM and ECM from other versions of externalism in terms of parts of the environment being continuously and actively coupled in the right way to parts of that subject's brain, the demand to say more is fair.

Nevertheless, while Adams and Aizawa are probably right that some arguments for EM have been hasty or at least too impressionistic in their appeals to coupling, I do not think that there is anything inherently problematic or fallacious about the sort of inference being made. Nobody ever really claimed that a body or brain being coupled to the environment is alone sufficient to demonstrate constitutive dependence. After all, my body, right now, is gravitationally coupled to all the stars in the sky. It would be absurd on that

⁴⁶ Cf. Clark's (2010b, 83) apt retort: "The appeal to coupling is not intended to make an external object "cognitive"... Rather, it is intended to make some object, which in and of itself is not usefully (perhaps not even intelligibly) thought of as *either cognitive or noncognitive*, into a *proper part of some cognitive system*, such as a human agent. It is intended... to ensure that the putative part is poised to play the kind of role that *itself* ensures its status as part of the agent's cognitive routines."

basis to infer that the vehicles of my current mental states and processes extend well beyond the Milky Way. They do not. They do, however, extend at least through regions of my brain, and there are now many ways to investigate which regions of the brain are necessary for occurrences of different types of mental events. One way to do this is by considering the manner in which the activities of different neurons in different regions are coupled. When coupling is sufficiently complex and high bandwidth, this is often taken as good, though defeasible, evidence that whatever mental events are the target of investigation, their occurrences constitutively depend on what happens in the entire region in which complex coupled activity takes place. If this inferential move is ever warranted in locating mental events within the brain, which it is, then it is not fallacious when made in defense of the claim that the vehicles of some mental events are located outside the brain.⁴⁷ Of course, this leaves it open what kind of coupling counts as sufficiently complex, and what, in addition to complexity, is important.

Next consider Rupert's (2004, 393) objection that while empirical evidence and theoretical considerations might support HEMC, advocates of EM have, as a matter of fact, taken them to support HEC.

Hypothesis of Embedded Cognition (HEMC): Cognitive processes depend very heavily, in hitherto unexpected ways, on organismically external props and devices and on the structure of the external environment in which cognition takes place.

Hypothesis of Extended Cognition (HEC): Human cognitive processing literally extends into the environment surrounding the organism, and human cognitive states literally comprise - as wholes do their proper parts - elements in that environment.

⁴⁷ See Sporns (2011) and Von der Malsburg, Philips, and Singer (2010) on the role of nonlinear coupling in neuroscience in determining the constitutive bases of mental states and processes.

The idea here is that the proponent of EM is claiming something more than dependence, whether explanatory or causal. Instead we have a compositional relation of some sort. Obviously, true dependency of a whole on its parts is different than the dependency of an effect on its causes, but Rupert, like Adams and Aizawa, does not see any compelling reason to adopt HEC when HEMC is a more conservative alternative that can do justice to the appearances.

While things get more complicated when thinking about coupling and consciousness, it is worth considering how mechanistic/interventionist strategies for localization can be put to use in an argument for EM. These strategies provide one way of specifying what sort of coupling might suffice for constitution. Kaplan (2012) has recently sketched how such a case might go by appealing to Craver's (2007a, 2007b) mutual manipulability account of constitutive relevance.⁴⁸ According to this account, something counts as part of the relevant system or mechanism, as opposed to a mere correlate, background condition, or sterile effect, when both top-down and bottom-up manipulations or interventions are possible. Bottom-up interventions involve selectively altering a putative part and measuring what, if any, changes occur to the system as a whole. Top-down interventions involve selectively altering the system as a whole and measuring whether there are corresponding changes in the putative part. If a proper pairing of these manipulations can be found, we have, in principle, met a sufficient condition for saying that the putative part is constitutively relevant to the system or mechanism. For instance, this may be achieved in neuroscience by pairing of a bottom-up study of some behavior using

⁴⁸ It should be mentioned that the general tone of Kaplan (2012) toward EM is critical. Accounts like Craver's are quite popular these days. For largely complementary accounts, see Bechtel and Richardson (1993/2010), Machamer, Darden, and Craver (2000), Pearl (2000), and Woodward (2003).

transcranial magnetic stimulation or direct electrical stimulation to the cortex and a top-down imaging study of that same behavior using EEG, fMRI, or direct recordings of cell firings to measure brain activity.

While this account is controversial, it has the advantage of theoretic neutrality, since it does not presume we know what could be constitutively relevant ahead of time. For example, while the properties of internal combustion engines, toasters, DNA transcription, and digestion are all mechanistically demarcated along pre-theoretical spatiotemporal boundaries, according to Kaplan (2012), “the neural mechanism underlying action potential generation fails to respect the relevant spatial boundaries of the neuron.” So whether a neuron is firing at any moment depends on more than its intrinsic properties at that moment. Relatedly, while one might be drawn to the idea that sense organs mark the outer bounds of perceptual systems, the mutual manipulability criterion itself doesn’t require settling whether this is so ahead of time.

Applying this to EM, consider the Otto example once again. As the case is described, we can expect to observe certain changes to his notebook as he forms, recalls, and corrects his memories. Moreover, we see from the Twin Otto case that changing the inscription in the notebook can by itself change Otto’s memory as well as his memory-related behavior. Thus, using the mutual manipulability criteria (in outline form, at least), a plausible case can be made that the inscription in Otto’s notebook is a part of the relevant memory realizing cognitive system. That is to say, it rules the inscription to be a constituent rather than merely a cause. This may be one promising way to distinguish what are really parts of something from what are causally coupled background conditions. For notice that

other cases where the coupled interactions between Otto's brain and his notebook are different, e.g. Otto carries it around but never looks at it, mutual manipulability would not be possible.

In some cases, however, for principled reasons that do not preclude any attempt at localization at all, mechanistic strategies for decomposition and localization may not be readily available. Specifically, the mutual manipulability criterion may only work when the putative system we are dealing with is component dominant. The behavior of component dominant systems can usually be understood in terms of linear functions that map the intrinsic properties of individual components onto the whole. In contrast, when the system in question is interaction-dominant, the behavior of the system as a whole depends much more on the relations between parts than on the intrinsic nature of the parts themselves. This is because the causal coupling among component parts is highly non-linear. When the properties of the whole system are not directly proportional to the properties of the components, a small change in the intrinsic properties of one part can drastically change the state of the system. In practice this can place barriers to selectively isolating and manipulating parts and hence to standard decompositional strategies for localization. One barrier might be that in the cases of conscious vision the relevant non-causally determining systems may sometimes only persist for a very short amount of time.

Interaction-dominant systems display what Clark (1998) calls 'continuous reciprocal causation', a kind of dynamic complexity in which alterations to individual parts are inseparable from alterations to the system as a whole. Still coupling, even complex, non-linear coupling, shouldn't always be taken as sufficient for establishing that a putative

component is a constitutive ground of a property instantiated by the whole. In arguing that bodily gestures are sometimes constitutive of cognitive processes, Clark (2008, 129-131) provides a plausible extra condition for when complex nonlinear coupling implicates what otherwise could be taken as output as part of the grounds. In some cases, the effects of one part are generated specifically in order to serve as later inputs to the system.

Clark has us consider, as an illustrative example, a triad of cases. The first is a case where the sound of rain against the window is coupled to one's brain activity and thereby helps pace of one's thoughts. The second involves a robot designed to make use of the sound of rain to pace its internal processes. And the third involves a robot that rhythmically spits water onto a plate and uses the auditory feedback produced by its spits to pace its internal processes. While the drops of water may merely form part of the causal background in the first two cases, it is quite plausible that they are constituents in the third. Another of Clark's examples involves the exhaust produced by a turbo-driven engine. The exhaust fumes, an effect of the engine's combustion, are fed back into the engine itself enabling it to produce more power and hence to move the vehicle more quickly and efficiently. In cases of continuous reciprocal causation where output is produced in order to be recycled as input to the system (or system component) that produced it, we have reason to suspect the minimal system responsible for the behavior we're interested in spans that interface.

It is plausible that feedback from bodily gestures is sometimes necessary for realizing various cognitive processes on various occasions, even if the gestures are also outputs of other cognitive processes. The same hand movement, say, equally coupled to

the band of air that surrounds it may be part of a cognitive process in some cases but not others. It will be when visual, proprioceptive, and tactile feedback of its effects are continuous and proximate causes of the unfolding cognitive activity internal to the subject's brain that produced it. It won't be when any feedback from the movement does not continuously influence the processes that produced in this way. Similarly, the exhaust from a non-turbo engine will merely be an effect rather than part of the system itself.

This may not be the only way to go in order to address the question of what more than non-linear coupling is necessary to call something a part, but it is worth noting that Clark's (1997) example of tuna propulsion and Wilson's (2010a) examples of the sensory systems of bats and eels fit this mold. Moreover, this 'recycled outputs' condition can be combined with the mutual manipulability criterion. For example, in the case of gesturing, we see that people gesture spontaneously while thinking and that you can impair someone's train of thought by inhibiting their ability to gesture. But while compatible with the relevant system being component dominant, it also provides guidance for distinguishing components from mere causes, effects, or correlates, when dealing with highly nonlinear, interaction-dominant systems where selective manipulation may not always be possible.

3.4 Locating Constitutively Grounded Events

Supervenience itself is not an explanatory relation. It is not a "deep" metaphysical relation; rather, it is a "surface" relation that reports a pattern of property covariation, suggesting the presence of an interesting dependency relation that might explain it. (Kim 1993, 167)

We take *ground* to be an explanatory relation; if the truth that P is grounded in other truths, then they *account* for its truth; P's being the case holds *in virtue of* other truths' being the case. (Fine 2001, 15)

In order to have a more general criterion of constitutive grounding, it helps to begin by noting that we are taking vehicles to be events that provide sufficient token-explanations of other events. To know whether something is part of a vehicle then is to know that it is a necessary part of a sufficient explanation for that event. As Kim (1993, 167) and others have noted, supervenience itself is not an explanatory relation, there can be modally-deep covariation even in cases where the base events or properties fail to explain the supervenient events or properties. Enter recent discussion of 'grounding' in analytic metaphysics. Something is grounded only if it is accounted for and holds in virtue of what grounds it. Moreover, the way in which grounds are supposed to explain what they ground is not a causal form of explanation.

It is the business of the empirical and mathematical sciences to discover and systematically explain correlations between natural events. Some correlations between events of different types are accidental, while others indicate the presence of some kind of lawful connection between events of the respective types. Lawful relations may be present when one event causes another or when a set of correlated events have a common cause. They are also present whenever there is an entailment relation between two events, such as the relation between the event of someone's saying 'hello' loudly and the event of that person saying 'hello'. Saying 'hello' loudly non-causally depends on saying 'hello'. One needs to explain how and why the correlations that are actually observed obtain.

Explanations, as opposed to non-explanatory descriptions or models, inform us of how and

why events of different types are related to each other in the ways they are. In so doing, explanations are able to provide a basis for predictions, hypothesis formation, and other less tangible types of illumination.

Causal or nomological relations are compatible with their relata being ontologically independent. Following Hume, it is plausible to regard causes and effects as ontologically distinct. Hume wrote, “[t]here is no object, which implies the existence of any other if we consider these objects in themselves” (THN 1.3.6). In other words, there are no metaphysically necessary connections between any distinct, intrinsically classified, entities. Hume’s dictum accommodates the following modal recombination principle: “anything can exist with anything else, at least provided they occupy distinct spatiotemporal regions,” (Lewis 1986, 88). While one can make sense of synchronic nomological determination, just as one can make sense of various synchronic and diachronic metaphysical determination relations, it may be infelicitous to call such relations forms of causation. Moreover, it is natural to take the temporal priority of causes to effects to entail their ontological distinctness. That is to say, C is temporally prior to E if and only if neither C nor E is ontologically prior to the other.

The explanatory dimension of grounding is tied to the idea that grounds are ontologically prior to the events they ground. Certain uses of “in virtue of” and “because” are claimed to capture this non-causal form of explanation appeals to grounding are invoked to provide. Given this, it might be worthwhile briefly examining what some of the proponents of EM say about “in virtue of.” Rowlands (2010) takes “in virtue of” to connote logical sufficiency. Wilson (2010b, 174-175) develops an argument for EM, one which

crucially builds on the ideas of parity and augmentation, and the comments on ‘in virtue of’ in the closing section of the paper entitled “Rethinking the Problem of Intentionality” (182-184). He thinks part of the problem is that critics of EM insist on an answer to the question “In virtue of what does internal state **R** represent condition (state, object, property) **C**?” (183). Instead, Wilson thinks that EM shows that the question should be reformulated to be “In virtue of what is activity **A** the representation of **C**?” (183). I agree. And I also agree that the relevant activity in the case of visual consciousness (2004, 2010a) often includes things that happen external to anyone’s body, but I do not quite understand why Wilson remarks that that this question “seems hardly pressing at all” (183). He goes on to briefly advocate that “the methodology appropriate here... is not traditional conceptual analysis of representational practices but an interdisciplinary, pluralistic motley” (183). Fair enough, I suppose, but all versions of EM stake out modally robust ontological claims so one of the disciplines’ disputants over EM ought to be in dialogue with is contemporary analytic metaphysics. It is a mischaracterization of analytic metaphysics to presume it works by appeals to conceptual analysis alone; the literature on the concept of grounding is a testament to that. There are many varieties of grounding; probably as many or more than the varieties of externalism.⁴⁹ But the endeavor articulate accounts of this notion should engender some optimism about making progress on a general understanding of ‘constitutivity’ that is highly relevant to the causal coupling v. constitution issue when discussed about EM. The notion can be used without providing a conceptual analysis of intentionality or consciousness and it is not simply a relation of logical sufficiency.

⁴⁹ For a useful recent review of the literature on grounding, its motivations, and the objections that have been raised about its viability, see Clark and Liggins (2012).

To get a better sense of ‘grounding’, let us consider Schaffer’s (2009) discussion of its uses and the motivation for regarding it as primitive notion, which I will follow him in doing. To differentiate grounding from any kind of supervenience relations, Schaffer (2009, 364) notes that grounding is anti-symmetric (not merely asymmetric) and hyperintensional. Hyperintensional relations cut at a fine-grain, i.e. between possible worlds that are co-intensive or necessarily co-extensive. To illustrate how this works, Fine’s (1994) example of Socrates grounding the set of Socrates is often appealed to. Other examples Schaffer (2009, 375) mentions of grounding between co-intensive relata involve holes in material substances, moral and natural properties, abundant and sparse properties, and truths and truthmakers.⁵⁰ Schaffer also rejects an attempt to analyze grounding in terms of “existential dependence counterfactuals” and registers agreement with Lowe’s (2010) criticisms of these attempts in a comprehensive and influential Stanford Encyclopedia entry on “Ontological Dependence”. Another example involving Socrates concerns his famous question for Euthyphro. Even if ‘what the gods love’ and ‘what is pious’ are co-intensive, it is legitimate to insist one must be explanatorily and ontologically prior to the other.

Schaffer (2009, 355) also distinguishes between flat, sorted, and ordered target structures for metaphysical inquiry. For flat approaches it is “an unstructured list of existents”; for sorted approaches the target of metaphysical inquiry is the number of categories and a list of the members of each; and for ordered approaches, it is “an ordered hierarchy generated from (i) a list of substances... [and] (ii) a list of grounding relations” (355). Schaffer goes on to defend an ‘ordered structure’ approach, which he labels “neo-

⁵⁰ More recently, Nolan (2013) has argued that hyperintensional relations are likely required for giving an adequate metaphysical account for many non-representational features of the world.

Aristotelian” and contrasts this account with “Carnapian/Quinean” approaches, which are flat. To be ‘derived’ or ‘grounded’ on this account requires bearing a relation to something more fundamental that is irreflexive, asymmetric, and transitive.

If this is what’s required for grounding, it is natural to ask the following questions. Is consciousness derived or grounded in this way? Or is it itself the more fundamental category? These are big questions I don’t propose to answer or even really address, but I think it is obvious that conscious events are dependent upon non-mental events, in that they are caused and sustained by them. Those non-mental events have spatiotemporal extension, which is why we can readily discuss issues about the locations of vehicles without begging too many questions about the essence of consciousness. Even if we cannot provide an explanation of consciousness that is structured, say because there’s a transitivity failure somewhere along the chain of material (to mental) constitution, and must take consciousness to be emergent, ECM could well be true. To not be dependent would mean being duplicable without the events that cause and sustain token conscious events, to be capable of what Lewis and Langton (1998) call “a lonely existence”. But unlike numbers and abstracta, a “Platonic” categorization of consciousness is not at all plausible when it comes to understanding its real-world manifestations. Perhaps a structural ordering can be provided; we would end up with a reductive explanation of consciousness in terms of its grounds in that case. But I am not supposing that we do have that. In fact, one might suspect that a form of emergence about consciousness to be supported by the appeal to non-linear coupling, interaction dominance, non-decomposibility, and the like, but most recognize that an explanatory understanding of emergence, in terms of reductions of

degrees of freedom say, is not sufficient for the kind of ‘ontological emergence’ that the British Emergentists embraced, for example, nor what is generally at stake in debates over the mind-body problem.

Taking a step back, we can say that ECM implies that there are non-bodily events, actively and continuously coupled to brain events in the right sort of way, that are parts of the constitutive grounds of some conscious events. I’ve indicated above what that ‘right sort of way’ might be. This distinguishes ECM from other forms externalism about consciousness in requiring that the external determinates play the token-explanatory role played by vehicles. In the next two chapters, I will begin by stating the extension thesis I plan to defend in terms of a constitutive grounding. I will then clarify the types of phenomenal characters the visual experiences under investigation have and argue that sufficient explanations of how and why visual experiences with those specific phenomenal characters occur on the occasions they do must make reference to something that happens beyond what is internal to some particular spatiotemporal boundary. Nothing intrinsic to just those regions fully grounds visual experiences of the relevant phenomenal types. Since even partial constitutive grounding entails that the grounding event is part of the grounded event’s metaphysical supervenience base, I will also consider what the duplication implications of the extension theses defended are and argue for their plausibility. To secure that these factors are partial constitutive grounds it needs to be shown that there can be changes with respect to whether experiences of the relevant type occur just due to changes in the aforementioned factors, even while everything else about the world remains the same.

4. TEMPORALLY EXTENDED VISION

In this chapter I focus on time and temporal extension. I will argue that all visual experiences have temporally extended or non-instantaneous constitutive grounds. In other words, the vehicles of all visual experiences are trajectories through time rather than punctate events at a time. This implies that no visual experiences are located at any instant anywhere in space. Even visual experiences conceptualized as momentary mental states are only realized or instantiated by processes and, as a consequence, by events that have temporal parts and duration. If the temporally extended vision thesis, TEV, is correct, then subjects' visual experiences always fail to supervene on those subjects' intrinsic properties at any single instant in time. Even if *some* visual experiences have full instantaneous constitutive grounds, however, what is most important for my argument for ECM by way of EVC is that many do not.

Throughout the philosophy of mind and the cognitive sciences, the expressions 'extended', 'embodied', 'embedded', 'situated', and 'scaffolded' are used in a variety of specialized and technical ways. Some commentary on these terms is necessary before proceeding. Let us start by reconsidering Rupert's (2004, 2009) distinction between HEC and HEMC.

Hypothesis of Extended Cognition (HEC): Human cognitive processing literally extends into the environment surrounding the organism, and human cognitive states literally comprise - as wholes do their proper parts - elements in that environment.

Hypothesis of Embedded Cognition (HEMC): Cognitive processes depend very heavily, in hitherto unexpected ways, on organismically external props and devices and on the structure of the external environment in which cognition takes place.

If we take this as our starting point, we can immediately extrapolate that ‘extended’ connotes a mereological claim, about what is part of cognition, while ‘embedded’ connotes a merely causal and explanatory kind of dependence on the environment. Of course, these may not be mutually exclusive, depending upon how we understand causation and modality more generally, but clearly HEMC is not sufficient for HEC. Unlike ‘extended’, I take the expressions ‘situated’ and ‘scaffolded’, like ‘embedded’, to imply no more than causal-explanatory dependence upon the environment.

Recall also Rowlands’ (2009, 53) characterization of EM (or HEC) as an ontic thesis about the composition of some mental property tokens that is partial and contingent. Pulling these together, we can apply it to the case at hand, i.e. conscious vision.

Extended Visual Consciousness (EVC): Some visual experiences are constitutively grounded by events that include parts of the non-bodily world

Recall that continuous and active coupling is required to distinguish EVC from other versions of externalism about conscious vision. My main goal in this dissertation, and the focus of chapter five, is to defend EVC. To defend this thesis, more is required than showing that visual experiences are sometimes environmentally embedded. It must be shown instead that parts of the environment are essentially parts of the vehicular events that constitutively ground some visual experiences. I take the expression ‘embodied’, like ‘extended’, to connote essential mereological dependence. To say that visual consciousness is embodied is to say that some visual experiences are constitutively grounded by events that include parts of a subject’s body that are not part of that subject’s brain. Some take the claim that something is embodied to rule out non-bodily extension, but I do not

understand things this way.⁵¹ While I will not explicitly argue for embodiment here, I think visual experiences are often both embodied and extended.

TEV is different than EVC in that it includes a condition that violates Rowland's characterization of extension theses as holding for only some conscious visual events. TEV makes a claim about the vehicles of all visual experiences. In other cases, it only matters that some types of visual experiences are sometimes extended in the relevant way.

Temporally Extended Vision (TEV): Some visual experiences are constitutively grounded by non-instantaneous events and no visual experiences are constitutively grounded by instantaneous events involving a subject's body.

TEV is restricted to denying that there are instantaneous bodily grounds because it is plausible, at least if determinism is true, that an instantaneous world event along with the laws grounds all subsequent events, including all conscious visual ones. One reason for advancing the universal negative claim is that appreciating the case for it will help the reader understand my defense of non-bodily spatial extension. More importantly, it seems to me both obvious and underappreciated. TEV is underappreciated because many simply assume that synchronic mental *states*, even if not diachronic mental *processes*, have instantaneous supervenience bases. In other words, it is often assumed that an instantaneous event involving a subject will fix or uniquely determine that subject's state of mind at that instant.⁵² So, for example, Metzinger (2003, 547) reasons that "there must be a

⁵¹ Understandings of 'embodiment' that may be in tension with EM are articulated and defended by Shapiro (2004), Gallagher (2005), and Lakoff and Johnson (1999).

⁵² Eliasmith (2008, 134) offers a partial explanation of why this is: "Many of our physical, computational, and metaphysical theories turn a blind eye to the role of time, often for practical reasons. So, perhaps it is not surprising that in the philosophy of mind - where physical, computational, and metaphysical theories meet - there has been a consistent tendency to articulate theories that consider function and time independently." If we conceive of computation in terms of Turing computability, it should be clear why the tendency is pervasive. Even if we move toward a more connectionist or dynamical systems approach, the relevant

minimally sufficient neural correlate for the content of consciousness at any given point in time. If all properties of this local neural correlate are fixed, the properties of subjective experience are fixed as well.” In the case of consciousness, if we think that we are only ever aware of the present moment and that the present moment has no duration, then we will be led to think that vehicles of our experiences must also lack duration. But if EVC is correct, then there is not a minimally sufficient *neural* correlate in all cases. If TEV is correct, then there is not a minimally sufficient correlate, at least not one involving the subject, at *any given point in time*. Let’s begin by considering some models of the temporal dimensions of consciousness.

4.1 The dynamic content problem for cinematic models

On one hand, consciousness seems to obviously have a temporal structure. We are conscious of some things before or after we are conscious of others and we seem to be conscious of things changing and moving. On the other hand, it seems that we are only conscious of things as they are *right now*, that is to say it seems that we are only conscious of the present moment. In this way, it may paradoxically seem that consciousness must be both temporally extended and temporally punctate. In the Stanford Encyclopedia entry on ‘temporal consciousness’, Dainton (2010) distinguishes three historically venerable models of how this seeming paradox may be resolved and how consciousness relates to time more generally. Using the terminology of vehicles and contents, we can characterize Dainton’s

equations specify the overall state of the system at any given time; so it is tempting to suppose that the vehicle of any state of mind always corresponds to the instantaneous total state of the system at that very time.

models in the following way: Cinematic models hold that all conscious experiences have instantaneous vehicles with static contents; retentional models hold that all conscious experiences have instantaneous vehicles with dynamic contents; and extensional models hold that all conscious experiences have dynamic vehicles with dynamic contents.⁵³

The cinematic model faces a problem the other two do not; one which is connected to why proponents of the unorthodox accounts of perception discussed back in the second chapter often reject it quite disparagingly. According to cinematic models all we really ever see is a succession of images and so we lack immediate awareness of anything other than static and motion-free ‘snapshots’ or ‘stills’.⁵⁴ Dainton (2010) goes on to explain that according to the cinematic model “[o]ur streams of consciousness are composed of continuous successions of these momentary states of consciousness. In this respect they are analogous to movies, which (as displayed) consist of rapid sequences of still images.” The clearest argument against this approach is simply that it implies that we never directly experience change, succession, persistence, movement, etc., when it seems obvious that we do. That is to say, it implies that our visual experiences never have dynamic contents and that our sense that they do is an illusion. Since the cinematic model holds that consciousness comes in discrete static chunks, like a child’s flipbook or a film. The movement and change we think we see in such cases is illusory. We never see change in any kind of direct, non-inferential, and continuous way, according to it.

⁵³ This leaves open a fourth model, i.e. one in which dynamic vehicles only ever carry instantaneous contents. While compatibility with realism about dynamic contents is a point in favor of extensionalist and retentionalist accounts, anti-realism about dynamic contents is compatible with the idea that experiences sometimes or always have temporally extended vehicles. It is likely that the cinematic model collapses into this fourth option.

⁵⁴ Cf. Koch (2004, 264) “Perception might well take place in discrete processing epochs, *perceptual moments*, *frames*, or *snapshots*. Your subjective life could be a ceaseless sequence of such frames.”

Visual experiences with dynamic contents include things like seeing a baseball flying through the air, a train pulling away from a station, one's own hand moving toward a coffee mug, etc. In each of these cases, the phenomenal character of the visual experience itself unfolds over time and involves some kind of persistence through change. Attempting to account for dynamic contents, “[c]inematic realists hold that all our experiences of change are the product of gap-free successions of momentary stream-phases, each possessing momentary – and static, motion-free – phenomenal contents” (Dainton). But it is hard to see how this could adequately explain the sense in which consciousness seems to flow with continuity over time. After all, a succession of momentary experiences is not sufficient for an experience of succession. Some, such as Augustine and Reid, have embraced anti-realism about the phenomenal presence of dynamic features of experience, but it is both anatomically and introspectively plausible to insist that we often see movement and change as directly as we see color and shape. Motion is registered and processed quite early in visual processing, after all. Moreover, for anyone who thinks visual phenomenology is somewhat thick – that we ever see causation, agency, affordances, etc. directly – the cinematic model should look particularly unappealing.

This liability for the cinematic model, i.e. that it renders all dynamic experience illusory, can be usefully compared to the ‘grand illusion’ reaction to now well-known evidence about change blindness and inattention blindness. We often fail to notice large and dramatic changes in the visual scene, especially if our attention is otherwise diverted or engaged, and some have taken this to show that we actually see much less of the world than we take ourselves to. In one example, the leaves on a tree that figure prominently in an

image may change slowly from green to orange without the subject ever noticing. The ‘grand illusion’ reaction is that this shows that our sense that we continuously see a richly colored array of patterns is mistaken. That we do not notice the gaps that result from our blind spots, that we do not register optical information between saccades, and that we do not think we see the world as colored and detailed only where we foveate further suggests that visual experience is phenomenally more impoverished and less rich than ordinarily assumed. Some might take these facts about human vision to support a cinematic model, since they suggest visual input is somewhat jumpy, gap-filled, and discrete. But the dynamic sensorimotor theorists, O’Regan and Noë (2001b, 92), offer a different interpretation:

[V]ision does not consist in the reception and sticking-together of successive images captured by the eye. On the contrary, vision is constituted by knowledge of the changes that occur through eye movements and body movements. Vision occurs *through* movement, not *despite* movement.⁵⁵

The idea that the characters of experiences are merely summations or composites of instantaneous snapshots is what leads to anti-realism. But this may suggest we have misconceived our explanandum. Our explanandum, the visual experiences of conscious subjects, clearly do include these dynamics elements. The ‘now’ of which we are conscious is not static.

This is why James (1890) rejected the cinematic model as phenomenologically implausible. For James (1890, 631), “the original paragon and prototype of all conceived times is the specious present, the short duration of which we are immediately and

⁵⁵ Consider also, O’Regan (2010, 46): “The impression of seeing ‘everything’ and of doing so in a continual fashion derives from the immediate accessibility of information at the slightest flight of attention of the eye, not from the existence of information in an internal brain representation.” It is only if we assume that internal brain representations provide the full constitutive grounds of visual experiences that we are likely to be drawn to anti-realist conclusions.

incessantly sensible.” If James is right that it is the specious present rather than a truly instantaneous absolute present that we are directly conscious of, TEV will likely be true of any visual experiences we should be tempted to take any theoretical interest in. Since a specious present involves a short duration but is also what is minimally sensible, it follows that its contents can be experienced both simultaneously and successively. It is a non-trivial task to explain exactly how this works, but if it is right then clearly appealing to a succession of discrete frames, even if there are no temporal gaps between individual frames would be inadequate.

Even though they face a problem in accounting for dynamic contents, it is not clear that contemporary proponents of the cinematic model really commit themselves to there being instantaneous vehicles. For instance, while offering a tentative defense of the cinematic model and expressing a willingness to accept its anti-realist implications, VanRullen and Koch (2003, 207) write:

Do we experience the world as a continuous signal or as a discrete sequence of events, like the snapshots of a Multimedia Component camera? Although the subjectively seamless nature of our experience would suggest that the relevant underlying neuronal representations evolve continuously, this is not the only possibility. Conscious perception might well be constant within a snapshot of variable duration.

To say that the snapshot has variable duration, however, implies that it is not always instantaneous. In fact, they hypothesize that “oscillations in different frequency bands could serve as a neural substrate” of discrete perception (207). But nothing oscillates at an instant so the relevant proposed vehicles are not in fact instantaneous. Contemporary theorists who think visual perception is discrete rather than continuous, might find the following alternative position appealing: Temporally extended vehicles carry either entirely static

‘snapshot’ contents, or perhaps even some dynamic contents, e.g. enough to make sense of individual specious presents, but that each snapshot is still taken to be packaged discretely and so there are gaps or discontinuities between different specious presents. Our sense of longer term diachronic unities, on such an account, would turn out to be an illusion.

But even if one is willing to go anti-realist about dynamic contents, including the contents of a specious present, one of the most remarkable features of conscious experience is the way in which phenomenal characters and contents within and between different sense modalities are unified for the subject in the experience even at a moment. In seeing a baseball flying through the air, the ball and the sky must be present together in the experience, these features must be synchronically co-conscious, they must happen together within a window of perspectival simultaneity. For this to happen, however, there must be some kind of process that allows these two elements to come together into awareness, and that must be something that happens over time.

4.2 The extensionalist and retentionalist alternatives

While the strict cinematic model implausibly implies that we never actually see change succession, movement, persistence, etc., but instead only ever infer these things from the static images we do see, both retentional and extensional models can accommodate the idea that visual experiences always, or often, have dynamic contents. “[T]he defining trait of the extensionalist approach is the rejection of the confinement of consciousness to momentary stream-phases” and insistence that “our episodes of experiencing are themselves temporally extended” (Dainton 2010). Non-instantaneous

vehicles present no in principle difficulty to explaining how we can consciously see movement, change, succession, etc. as directly as we do colors, objects, and shapes. After all, changes take place within the relevant temporally extended intervals and those changes may well be sufficient to ground such experiences.

For retentionalists, even though consciousness is confined to momentary stream-phases, each phase is postulated to have a complex structure that preserves or retains parts of the recent past. According to Husserl's influential account, experiences have a tripartite structure in which they represent something as currently present and anticipate something about what is to come while retaining representations of the immediate past. To use Husserl's terminology, in addition to a 'primal impression' of what is present right now, every experience is also structured by 'retentions' of prior experiences and 'protentions' or expectations about what near future primal impressions are likely to be.

One of Husserl's favorite examples involves auditory perception. The primal impression of hearing a single note does not capture the unified synchronic experience of hearing that note on an occasion when it is played as a part of a melody, or held continuously, or played staccato. The exact same note at an instant will be experienced differently. Moreover, Husserl takes this example to illustrate that "any given now-phase of consciousness retains the whole just-past phase of consciousness, specifically this phase's primal impression, retention, and protention" (Thompson 2007, 321).⁵⁶ So, a momentary primal impression is merely one essential component of the whole, at the level of the explanandum, i.e. occurrent visual experiences of the present moment.

⁵⁶ Cf. Grush (2006, 421) explains that on Husserl's retentionalist account "the content of my retention of what happened a brief moment ago is not just what was primal impression a moment ago, but the full retention-impression-protention structure from that moment."

If retentionalism is correct, instantaneous vehicles may carry contents representing temporally extended intervals rather than discrete snapshots of optical information.

Articulating some of the underlying motivations for retentional models, Dainton (2010) writes:

Reid's contention... that 'the senses give us information of things only as they exist in the present moment' has a good deal of intuitive plausibility, as does taking the present moment to lack duration, in the manner of Augustine. Anyone who subscribes to each of these doctrines and who also believes we are immediately aware of change has little option but to opt for the Retentional approach.

Nevertheless, the retentional model, insofar as it is committed to experiences having instantaneous or punctate vehicles, is implausible. As I will argue next, it is independently plausible that vehicles are temporally extended and to infer otherwise based on considerations about the nature of the phenomenal characters or contents of experience is fallacious. It does no good to point out that in principle vehicles that can be fully present synchronically can sometimes carry dynamic contents, e.g. the inscription 'the ball bounced' or a drawing of someone mid-stride,⁵⁷ since any account we give of the vehicles of actual experiences will require events that have processes as parts. This leaves it open whether experiences with dynamic contents have experiential parts or not.

4.3 Why the vehicles of visual experience are never instantaneous events

There are quite general physiological reasons for thinking that the vehicles of consciousness are never instantaneous that do not rely on any special considerations about

⁵⁷ Crick and Koch (2003, 122) appeal to this second example as a way of making sense of how cinematic anti-realists might understand visual perception as consisting in "a series of static snapshots, with motion "painted" on them."

vision or consciousness. The neural events potentially sufficient to constitutively ground occurrent mental events always involve processes or activities among groups of neurons, never simply momentary or instantaneous brain events. Even at the level of single neurons within the brain, information is encoded and processed in terms of properties like firing rate and the rate at which a neuron fires does not supervene on its state at any single instant. In this way, a neuron firing is not like an on-off switch, nor the velocity of a simple moving object, which can be understood as present or absent at an instant. Action potentiation is a process that takes time to occur. Coordinated firing of many neurons throughout the brain takes even longer. In fact, candidate correlates of various types of mental events within the brain very often involve groups of neurons oscillating together at different frequency bands. Again, rate of oscillation is not fixed at an instant.

More generally, it is often the case that states we conceive of synchronically fail to supervene on any instantaneous event. For instance, according to Ladyman and Ross, even the liquidity of some quantity of water is not supervenient on its microstructure at a time. Suppose some water is in a liquid state, rather than being gaseous or solid. It may seem that the water's liquidity must supervene on the state of the H_2O that composes it at that time. But, as Ladyman and Ross (2007, 21) explain, water "is composed of oxygen and hydrogen in various polymeric forms, such as $(\text{H}_2\text{O})_2$, $(\text{H}_2\text{O})_3$, and so on, that are constantly forming, dissipating, and reforming over short time periods in such a way as to give rise to the familiar properties of the macroscopic kind water." For this reason, they contend that with respect to water, it "makes no sense to imagine it having its familiar properties synchronically" (Ross and Ladyman 2010, 160). I do not think this is quite right. We can

still imagine macro-level properties of complex dynamical systems synchronically – it is not as though the water ceases to be liquid at any instant – it is just that we cannot simply assume that synchronic properties always have synchronic grounds.

I mention this non-mental example as a way of illustrating why I think it is somewhat obvious that the vehicles of visual experiences are likely temporally extended, even if we are inclined to think of their contents purely synchronically. In fact, it is likely that only purely dispositional mental states like unrecalled episodic memories, stored declarative knowledge, implicit understanding of phonology and syntax, unexercised affective capacities, etc., are plausible candidates for having constitutive grounds involving just the standing structure of the brain at a time. Processes are always temporally extended, so if events involving them constitutively ground all occurrent mental states and process, then they are all temporally extended in the relevant sense. Since all visual experiences are occurrent, TEV follows straightforwardly.⁵⁸

What about the suggestion that conscious visual experiences are constitutively grounded by instantaneous brain events along with their dispositional profiles? If we trace back the trajectory of brain events that led to a given momentary event, perhaps one that is spatially distributed throughout much of the brain, wouldn't that be enough to determine existence and character of the subject's visual experience at that time? I think not. First of all, since grounding entails supervenience, it would have to be that the event's dispositional

⁵⁸ If we focus on not merely occurrent states and processes, but conscious ones, the point becomes even more apparent. As Lutz and Thompson (2003) put it: "It is now widely accepted that the neural processes crucial for consciousness rely on the transient and ongoing orchestration of scattered mosaics of functionally specialized brains regions, rather than any single neural process or structure... Both animal and human studies demonstrate that specific changes in neural synchrony occur during arousal, sensorimotor integration, attentional selection, perception and working memory, which are all crucial for consciousness."

profile is fixed as a matter of metaphysical necessity. But it's hard to see why that would be. Consider a twin case. Suppose $W1$ and $W2$ are distinct but duplicate possible worlds with the same laws of nature up until time T . $S1$ and $S2$, subjects inhabiting $W1$ and $W2$ respectively, are intrinsic duplicates at T . Since we are dealing with the realm of metaphysical possibility, however, we may suppose there is a sudden and very short-lived event in $W2$ that involves a violation of the laws of nature immediately after T .

Furthermore, we may even suppose this event is localized entirely within $S2$'s brain. In such a case, while $S1$'s and $S2$'s brains are in the same intrinsic state at T , the firing patterns of the neurons in $S2$'s brain will be different. Since the local dispositions fail to fix what happens what happens in the brain immediately after T , the experiences we can ascribe to $S1$ and $S2$ at T could be different. My reasoning here depends on Hume's dictum about there not being metaphysically necessary connections between spatiotemporally distinct existences. If there are not, dynamical processes that span T won't be fixed at T . So, similarly, if two samples of H_2O are identical up until T , depending on what happens immediately after T , it may that one sample is liquid but the other is not. The past and local dispositions at T don't fix the H_2O 's macroscopic properties at that time.

In a more mundane way, postdictive effects on perception such as perceptual masking, the flash-lag effect, and the phi illusion can help to illustrate the point. For instance, in the flash-lag effect, people will experience a flash of light that in fact occurs at the same location as a moving object to lag behind it. In one telling study, Eagleman and Sejnowski (2000) performed studies in which the movement of the object after the flash was altered in various ways. For instance, in some cases it reversed directions, in others it

stopped shortly after the flash. They found that changes in the trajectory of the moving object initiated up to 80 ms after the flash could alter where the flash was perceived to be. They argue that this tells against models of the flash-lag effect that appeal to predictive coding of where the moving object is likely to be and against models that propose a difference in how long the visual system takes to process motion rather than flashes. Such models may be compatible with the idea that subjects' visual perception of the flash is fixed by their momentary states plus local dispositions, but this is not an option if environmental events after the flash is initially processed affect its perceived location.

So we can construct another twin case that builds on these results. Suppose we have two duplicate subjects participating in one of these studies and light from the flash has just stimulated their retinæ and is now being processed in visual cortex. Nevertheless, if they are in different conditions and what happens to the moving object after the flash is different, they will experience the flash at different locations. The momentary intrinsic state of their bodies, including the recent trajectory of states and whatever dispositions those trajectories determine, will not be enough to ground the character of their visual experience of the flash at that time.

4.4 Enactivism and retentionalism

The proponents of EVC I discussed in chapter two and will draw on in chapter five are united in their opposition to the cinematic model. Nevertheless, the enactivists' Thompson and Varela endorse a Husserlian conception of time consciousness, which is a version of retentionalism. By looking at the details of their data and overall theoretical

framework, however, it is not clear whether the considerations they marshal really favor a retentional model over an extensional one. Elsewhere, they explicitly endorse ECM and TEV. For instance, clarifying how the notion of ‘representation’ is understood by enactivists, Thompson (2007, 59) writes:

Representational “vehicles” (the structures or processes that embody meaning) are temporally extended patterns of activity that can crisscross the brain-body-world boundaries, and the meanings or contents they embody are brought forth or enacted in the context of the system’s structural coupling with its environment.

Nevertheless, the retentional models – as explicated by Dainton and endorsed by (at least early) Husserl – hold that the vehicles of conscious experience are instantaneous. Active vehicle externalism about visual experience, however, does not fit naturally with a retentional model. If the vehicles are instantaneous, it is more likely that we can find minimally sufficient explanations within subjects’ heads. So proponents of EVC should on the face of it instead prefer extensionalism. If one allows that the vehicles themselves lack duration while embracing externalism about consciousness, this supports a view that is probably closer to wide intentionalism than ECM, and hence is compatible with internal supervenience with respect to the vehicles of conscious events.

Let us back up and look a bit more closely at the differences between retentionalism and extensionalism and consider whether the evidence the enactivists Varela and Thompson cite in support of their “neurophenomenological” theory of time consciousness really supports retentionalism. According to Dainton’s retelling, we see the development of explicitly retentionalist models starting with Kant, Brentano, and Husserl. Retentional models, once again, hold that “our experiencing of change and succession occurs within episodes of consciousness which themselves lack temporal extension, but

whose contents present (or represent) temporally extended intervals and phenomena.” Varela, Thompson, and Rosch (1991) distanced their view from Husserl, who they took to accept a problematic form of methodological solipsism. They argued that mindfulness based meditation techniques found in the Buddhist tradition are instead “a more promising phenomenological partner for cognitive science” (Thompson 2007, 413). But Thompson (2007, 413-416) now suggests that they had misinterpreted Husserl based in part on being influenced by Dreyfus’ (1982) interpretation. He thinks Husserl’s considered views actually move him away from solipsism toward a more Heideggerean being-there-in-the-world type of account. However, Thompson’s (2007, 312-359) and Varela’s (1999) accounts of temporal consciousness are explicitly Husserlian, even as accounts of the specious present.

“Neurophenomenology” refers to a methodological approach developed by enactivists to investigate the brain’s contribution to consciousness in explicitly dynamical cum phenomenological terms. According to this approach, the structure and dynamics of brain activity and the subjective reports different people give in psychophysically controlled situations place mutual constraints on adequate accounts of conscious explananda. With respect to visual experience of the present moment, of what is being seen as happening right now for the subject, this approach adopts Husserl’s views on the nature and structure of time consciousness. Consciousness of the present moment has a tripartite structure involving retention, primal impression, and protention; all three of which are synchronically present in any conscious experience whatsoever.

Given their emphasis on the nonlinearity and their explicitly expressed sympathy for EVC, and ECM more broadly, one might think enactivists would want to deny that the

vehicles of experience are ever instantaneous. Let us look more closely at the details of how they understand brain dynamics. Cosmelli, Lachaux, and Thompson (2007, 729) claim that the basic intuition behind their neurophenomenological approach is that “[c]onsciousness is an intrinsically dynamic phenomenon that must therefore be studied within a framework that is capable of rendering its dynamics intelligible.” This is the framework of dynamic systems theory. Part of their motivation for employing this framework is that it is well-suited to measuring what more orthodox, representationalist, approaches in cognitive neuroscience have discounted as irrelevant noise. They argue that shifts in transitive consciousness depend on dynamic patterns of large-scale neural activity rather than the momentary state of any single brain region or structure. This activity is not only distributed spatially, it also requires some kind of minimal temporal duration to emerge. The time scale at which a pattern of coherent neural activity must hold together to have any transitory efficacy is at the order of 250-500 ms. Phase-locking, in which the oscillatory discharges of different neurons synchronize for a short period of time, indicates large-scale integration of processing. But integration requires not only synchronized coupled neural activity (or phase-locking), but also periodic phase scattering or de-synchronization. De-synchronization time itself “defines a temporal window of simultaneity, such that whatever falls within this window counts as happening ‘now’ for the system” (Thompson 2007, 334). Following Varela (1999), Thompson suggests de-synchronization rate is what constitutively grounds the specious present.

Adopting a neurophenomenological approach, one can understand phenomenological retentions as constitutively grounded by the dynamical trajectory of

recent patterns of synchronized and scattered activity required for large-scale integration of processing in disparate regions of the brain. According to Varela (1999, 296), phenomenological protention is an affectively laden global order parameter that shapes the dynamics of large-scale integration in the brain. This notion is similar to Kelso's (1995, 137-158) account of intention as an order parameter for the coordination dynamics guiding behavior. The "primal impression involves no reference to either past or future" (Thompson 2007, 319). Perhaps a neurophenomenological approach can accommodate the idea that current visual primal impressions are constitutively grounded by the momentary intrinsic properties of the network of neurons and glia in the subject's brain. But primal impressions do not constitutively ground, nor in any other way comprise, make up, build, realize, or determine, by themselves any experience whatsoever. Instead, as we have seen, all occurrent experiences have both retentional and protentional features in addition to any primal impression(s). For this to happen, any candidate vehicular event must last for a short but non-trivial amount of time.

According to the results of the enactivists neurophenomenological investigations, we should understand the vehicles of consciousness in the following way. What we experience as present is constitutively grounded by events in which a dynamic pattern of phase-locking and phase-scattering between the firing rates of neurons in sometimes disparate brain regions occurs. The intervals between phase-scatterings demarcate vehicles of different specious presents. As Gallagher and Varela (2003, 123) put it, "the fact that an assembly of coupled oscillators attains a transient synchrony and that it happens within a certain temporal window is the explicit substrate of the living present." Moreover, Cosmelli,

Lachaux, and Thompson (2007, 748-752) detail how metastable patterns of resonate gamma assemblies, i.e. populations of firing neurons temporarily synchronized in the gamma frequency band, indicates that neural activities in disparate areas of the brain are often part of a conscious event's constitutive ground. Such a methodology has been used successfully, for instance, in studies on object perception, binocular rivalry, and attentional blink.

So the enactivist account of temporal consciousness is not really a version of retentionalism, as characterized, but instead entails TEV. Unlike other extensionalist and retentionalist models discussed, however, it is not immediately clear that enactivism can make sense of longer-term dynamic contents. In this way, it may end up being more like the cinematic model than one would expect. To get a better appreciation of the issue I am raising, let us consider a problem that Kelly (2005) raises in a paper on temporal experience in a section entitled "What is Wrong with the Specious Present" (230-231).

Kelly writes: "[t]he doctrine of the Specious Present... proposes that we are at every moment in direct perceptual contact not only with what is now occurring but also with what has recently occurred and indeed is about to occur as well" (230). He then raises three objections to James' doctrine. The strongest of these begins with the observation that "[i]n order to account for perceived motion that lasts longer than the duration of a single Specious Present, the view would have to allow for the possibility of stringing Specious Presents together" (231). Kelly's point is that the resources of a Husserlian retentionalist account (or something else entirely) will be necessary to explain the dynamic contents of conscious events longer than a specious present, e.g., to explain the diachronic unities

necessary to experience complex melodic variations in music.³⁹ It is not clear how an enactivist account in terms of metastable coordination and gamma synchronization will do this. Especially if one ties specious present individuation to de-synchronization rate, it would seem that the conscious “now” is packaged discretely. In other words, given the story of vehicles in terms of transient synchronization of neurons, it seems that the contents of individual specious presents would not overlap with the contents of any others. So there would have to be discontinuities. Dainton (2010) objects to Varela (1999) on this point, “the idea that experience is packaged into discrete chunks is not in the least Husserlian.”

Insofar as enactivists posit an account of the vehicles of visual specious presents in terms of intervals of de-synchronization, this may suggest that they, like contemporary proponents of cinematic model, are not taking up the burden of explaining conscious events with longer-term dynamic contents. Then again, perhaps what Thompson and Varela are actually illustrating, given their qualified endorsement of Husserl, is precisely that “the doctrine of the Specious Present” and retentionalism should not be opposed. Perhaps supposing short-term dynamic contents are packaged discretely into specious presents is not in tension with availing oneself of a retentionalist account of longer-term dynamic continuity. After all, as we have seen, all viable accounts of the vehicles of visual experience entail TEV. No doubt the contents of these discrete perceptual episodes will have dynamic contents, for enactivists, rather than simply being static snapshots, as traditional proponents of the cinematic model hold. Moreover, unlike Kelly (2005, 230), and probably many others, I do not find James’ notion of the specious present in anyway

³⁹ Kelly (2005, 230) writes “to say that we are not aware *of* past events, but only aware of them *as past*... is to turn the doctrine of the Specious Present into a Husserlian kind of intentionalist theory” (230).

“baffling.” It, rather than some snap-shot absolute present, should be the starting point for understanding active conscious minds like ours.

One option for accounting for experiential continuity beyond a specious present is by appeal to what Dainton calls ‘the overlap model’. This is a version of extensionalism that holds that streams of consciousness “are composed of successions of stream-phases which overlap by sharing common parts.” As Russell (1913/1984, 77-78) and others have recognized, diachronic co-consciousness is not transitive. For example, if someone is playing a scale I may experience the C and the D together and the D and the E together without experiencing the C and E together. Using just noticeable differences (a standard and robust psychophysical measure) to distinguish specious presents, streams of consciousness will overlap significantly, but “there is no phenomenologically unrealistic repetitions” (Dainton). Rather we get “a temporally extended stretch of consciousness that is phenomenally continuous, and which includes all the experienced transitions which *are* experienced” (Dainton). So, according to the overlap model, streams of consciousness are gap-free, experientially connected and display moment-to-moment qualitative similarity, without having to say we experience everything all at once. This account answers Kelly’s (2005) objection. So even if the enactivist account of temporal consciousness falters, there’s no reason for the proponent of TEV to worry.

4.5 Why visual experiences fail to supervene on instantaneous events

At the end of the day, the extent of the compatibility between the results of the enactivists’ neurophenomenological investigation of time consciousness and Husserl’s views

is unimportant. What is important is that any view that is remotely plausible will allow that the vehicles of visual experiences are events that have some non-trivial duration, whatever their experiential structure. The cinematic and retentional models either entail that the vehicles of visual experiences are punctate or instantaneous, or they do not. If they do not, then they entail TEV. If they do, then they are empirically implausible and likely involve some kind of content/vehicle conflation. After all, even synchronically conceived states like being liquid have temporally extended constitutive grounds. Regardless of how one conceives of the content/character, occurrent mental events like visual experiences always as a matter of fact have temporally extended vehicles containing processes involving patterns of firing neurons.

The events that constitutively ground visual experiences are always enduring events, ones involving large-scale integration of spatially extended neural activity. If TEV is correct, it follows that it is metaphysically possible for there to be a duplicate of the brain or body of a seeing subject at a moment without duplication of that subject's visual experiences. In contraposition, it is metaphysically necessary that if there is no enduring event, there will be no visual experience. An instantaneous event involving the subject will not suffice to determine any visual experiences on its own. I have already argued that there is a failure of instantaneous supervenience even if we build in the past and local dispositions, but the case can be made more easily by considering a mere instantaneous duplicate of a seeing subject. Suppose that Fred is an actual subject currently having a visual experience of a mountain. Within the realm of metaphysical possibility, we can imagine a scenario in which the entirety of reality contains only a momentary molecule for molecule duplicate of Fred that

pops into and immediately vanishes from existence. Is it plausible to claim that Twin Fred would nevertheless be visually experiencing a mountain in that moment of existence?

Surely not, since none neurons in Twin Fred's brain would actually be firing. While perhaps we can ascribe some of the same non-occurrent mental states to Twin-Fred as Fred, such as his dispositional beliefs, none of Fred's occurrent states would be preserved. So there can experiential differences among instantaneous duplicates.

If all visual experiences have partial constitutive grounds that are processes extended in time, then it follows that there can be changes with respect to the phenomenal characters of a subject's visual experiences without there being any change to that subject's momentary intrinsic properties. The explanatory considerations put forth earlier were intended to show that even a statically conceived experience can only occur in virtue of some kind of temporally extended neural process which itself cannot be divided up or decomposed into a punctate series of instantaneous events. Even if the case for this universal claim has not been sufficiently demonstrated, the specific claim that some visual experiences have temporally extended grounds has been. Experiences with dynamic contents, including specious presents, always have temporally extended vehicles.

5. EXTENDED VISUAL CONSCIOUSNESS

While most discussion of the extended mind thesis, EM, has focused on cognitive but unconscious aspects of our mental lives, my aim is to demonstrate that a comparable claim holds about the vehicles of consciousness. Specifically, I defend the extended conscious mind thesis, ECM, by arguing for active vehicle externalism about visual experience. I call this the extended visual consciousness thesis or EVC. One of the key premises in my argument for EVC is the temporally extended vision thesis, TEV, which was the focus of the last chapter. TEV is the thesis that the events that constitutively ground familiar visual experiences are never instantaneous events involving the subject. Of course, even if the constitutive grounds or vehicles of visual experiences are temporally extended, this is not enough to show that they ever spatially extend beyond a subject's body. It suggests, however, that it is important to pay attention both to the length of putative vehicle events and to details about what happens within the relevant window of time, especially among those things which most clearly form parts of the vehicle. Adopting an extensionalist account of temporal consciousness and focusing on experiences of the specious present, we can estimate that the shortest relevant vehicle events last at least one hundred milliseconds. After all, for extensionalists, content/character duration tracks vehicle duration and even a single millisecond flash of light is experienced by subjects as lasting for 100-400 ms, so we can infer that the events that act as vehicles of such experiences last at least that long. Longer estimates of the length of a specious present have it clocked in at over one second so it may be that visual experiences with minimal phenomenal duration never have vehicles

that are shorter than that.⁶⁰ By focusing on visual consciousness, we can be most confident that the relevant events most centrally involve processes in the occipital lobe. To argue for EVC, I will be considering how such processes are coupled to changes in the environment, specifically to unfolding changes in the ambient optic array. Attempting to screen off the ambient optic array as causal scaffolding and look at the internal activity in isolation results in an artificial or gerrymandered understanding of the relevant internal activity. Especially for phenomenally rich experiences with lengthy vehicles,⁶¹ once decoupled from their dynamic context, we cannot assume that internal activity taken in isolation would be sufficient on its own to constitutively ground them.

Toward the end of the second chapter, I discussed a number of theories of visual perception that stand at odds with more orthodox cognitivist accounts. These theories are united in their rejection of the claims that the primary function of visual perception is to produce representations of the world for subsequent cognition and that there are always systems sufficient for realizing visual experiences within the head. In defending the crucial second premise in my argument for EVC, I will emphasize each of these points of divergence separately. By focusing on its functional role, Wilson's (1994, 2004, 2010a) work provides resources for appreciating why only systems that sometimes extend beyond the body of their subject are likely capable of playing the right, action-guiding, role. The Wilson-inspired case is somewhat conservative in that it is compatible with the viability and

⁶⁰ The 100-400 ms experience of a single millisecond flash would only be a part of an experience with minimal phenomenal duration in that case; the flash's coming into being and fading out of existence would also be essential parts. See, e.g. Nisly-Nagele and Wasserman (2001) (cited in Dainton (2010, n. 44)), for more on divergences between the duration of stimulus display and experienced visual persistence.

⁶¹ Episodes of normal looking might plausibly be demarcated by blink rate, i.e. as lasting 1-10 seconds, though given saccadic suppression it would be mistaken to assume that the brain gets continuous visual input throughout the time between blinks.

need for representational explanations, decompositional strategies for localizing vehicles, and retaining pre-theoretic bodily boundaries for subjects. By focusing more directly on the nature the systems sufficient to realize experiences of various sorts, Chemero's (2009) work nicely synthesizes the case for taking the relevant minimal system's boundaries to be flexible and sometimes to spatially extend beyond the body. Combined with a more thorough anti-representationalism, this work allows us to appreciate why complex nonlinear coupling in some cases blocks the possibility of finding more local vehicles within a system. After developing this main line of argument by defending a premise about dynamical entanglement, emphasizing the two distinct strategies for establishing it just mentioned, I then consider why bodily duplicates can and would have different experiences in cases where one is completely decoupled from any environment whatsoever. Introduce randomness anywhere into the grounding vehicular event and there will no longer be something internal that functions or synergizes in the way it otherwise would.

To back up for a moment, let us consider the two other versions of externalism about consciousness that were discussed in the second chapter: wide intentionalism and disjunctive direct realism. While both entail that various types of conscious states and processes are individuated in part by relations to the environment, only the second implies vehicle externalism about visual experience. The first is compatible with either an internalist or an externalist account of the vehicles of conscious thoughts and experiences. Disjunctive direct realism, however, implies something quite specific about the vehicles of veridical visual experiences, which EVC more generally does not. Namely, disjunctivism implies that the vehicles of veridical visual experiences extend to include the seen object or

property itself. Suppose a subject consciously sees the Andromeda galaxy. In order for that to happen, according to the disjunctive direct realist, the subject must be appropriately related to something about 2.5 million light years away. It follows that the vehicle of that experience extends not just beyond that subject's body but to long before the subject was born. My version of EVC does not imply anything like this. For the light surrounding the subject to even potentially be part of the vehicle of an experience it must be capable of interacting with the relevant brain processes within the relevant time frame. The relevant coupling must be bi-directional rather than one-way and for that to happen it must be at least within a light cone surrounding the subject's eyes, i.e. the distance light can travel from them during the duration of the vehicle event. Even light from the sun takes over eight minutes to reach earth, which is far too long to make sense of it providing feedback for things that happen within a subject's brain, as opposed to exerting a constant causal influence.⁶² In fact, though light itself travels incredibly fast, the relatively short length of most vehicular events will limit all but relatively nearby events in the local environment from even being candidate constitutive grounds. So, the coupling based case for EVC has the resources to limit the maximal spatial extent a vehicle might have in a principled way. As a version of externalism about consciousness, this might be regarded as a point in its favor over disjunctivist views.

⁶² If there are visual experiences that last a very long time, e.g. the length of basketball game or the whole day, candidate vehicles will still exclude very distant things which can and do nevertheless interact with us by exerting one-way causal influence such as Andromeda.

5.1 An Argument for EVC: Temporal Extension plus Dynamical Entanglement

Though he ultimately finds the argument unsound, Clark (2009, 975-980) charitably reconstructs an argument for ECM, which he calls the dynamical entanglement argument, by distilling some of the more promising themes developed at various places by Hurley, Noë, Thompson, etc. Before discussing the dynamical entanglement argument, Clark presents two other often made arguments but finds them less impressive.

The first of these deals with cases of cortical/neural deference, where the relevance to experience activity in one part of the brain has is at least causally sensitive to dynamical sensorimotor feedback from the environment. Some much discussed examples include experiments on subjects using sensory substitution devices and spatial or color inversion lens. Hurley and Noë (2003), for instance, have argued that the variability of neural correlates we observe in these cases supports EVC and ECM. They suggest that it is the extended sensorimotor feedback loops in their entirety that determine the relevance of brain events to experience, not something about those brain events taken in isolation. Clark's (2009, 970) objection to the variable neural correlates argument for ECM is that the evidence "leaves open the possibility that embodied activity is just a causal precondition of the setting or re-setting of parameters in neural populations." His point is that in cases where different types of activity in different parts of the brain correlate with separate occurrences of conscious events of the same type there still may be an internal sufficient condition in each case.⁶³ Even when the internal brain events are quite different, the

⁶³ As Thompson and Cosmelli (2011, 180, n. 43) observe, "What look like variable neural correlates could still share some higher-order invariance, which might then qualify as a supervenience base for the experience."

extended sensorimotor dynamics that re-parameterize the brain are just causes. Since this will not be the main argument I advance, I will not comment further on its plausibility.

The second argument Clark considers is Noë's (2004, 2006, and 2007) virtual representation argument for ECM, which attempts to account for phenomenal presence entirely in terms of sensorimotor access. The dynamic sensorimotor theorist's explanatory strategy here draws on the way it is able to give a non-illusory account of change blindness. Potentialities structure the character of experience rather than just internal representations. When Noë (2004) writes "you cannot factor experience into an occurrent and a merely potential part" and that "experiential presence is virtual all the way in," his point is just that perceptual experiences only ever supervene on temporally extended events, never punctate ones. The state of a system, whether internal or extended, at a single moment never corresponds to what it is like for that system at that moment by itself. Clark is at a loss as to what to make of this argument, but I think the best interpretation of it is one where it supports TEV directly and EVC only indirectly.

In Clark's (2008, 975-980) articulation of the dynamical entanglement argument, he notes that a starting point for all involved is a rejection of what Hurley (1998) calls the Input-Output Picture, the view that personal-level distinctions among perception, cognition, and action map linearly onto processing activities within the brain.⁶⁴ Rather, proponents of the dynamical entanglement argument hold that "motor processing and perceptual uptake each unfold courtesy of a mass of ongoing looping interactions" (975). The continuous

⁶⁴ Wilson (2010a) calls this the 'flow through' model of how visual perception fulfills its action-guiding functions. Clark (2008, 251, n. 17) notes that something like the Input-Output picture or flow through model mistakenly suggests that "mind... is that which mediates perception and action, and the body just *is* the place where perception and action meet." To understand why this is problematic on purely phenomenological grounds, Clark appeals to Dennett's (1978, 310-323) well-known thought experiment.

reciprocal causation we find among such looping interactions is part of the reason why neural elements cannot simply be understood in isolation at some point along a linear hierarchy.

The next step, according to Clark, roughly amounts to TEV. If we reject the notion that visual experiences ever supervene on punctate states of subjects, we are left with the idea that only events that contain temporally extended processes ever suffice for visual experience. TEV, recall, is the thesis that some visual experiences are constitutively grounded by non-instantaneous events and no visual experiences are constitutively grounded by instantaneous events involving a subject's body. The universal condition in TEV is not strictly necessary, but it is plausible and important. After all, as I argued last chapter, the only plausible versions of the cinematic and retentional models (as accounts of content and character) end up being versions of TEV. Whether conscious contents flow continuously or are packaged discretely, they are always carried by temporally extended vehicles.

But even if TEV is correct, because the relevant grounding events must either persist or evolve over time in some specific way, say, Clark notes that this fails to show that temporally extended vehicles ever spatially extend beyond the brain. The next step is to argue that in at least some cases the only possible way for the right "kind of 'signature' temporal evolution of neural states" to occur is by the brain actually being embedded in a body and environment from which it receives constant feedback (979). The real challenge is to make sense of what this impossibility claim amounts to. After all, could not just the same temporal evolution of neural states occur randomly? If so, would not that suffice for

experiential duplication in all cases? To make the dynamical entanglement argument for ECM work, the answer must be ‘no’.

Here is my reconstruction of the dynamical entanglement argument specifically as it applies to conscious vision.

P1. In the explanation of visual experience, the brain internal parts of the temporally extended events that constitutively ground visual experiences often cannot be decoupled from parts of the non-bodily world.

P2. If event A is a constitutive ground of event E and event B cannot be decoupled from A in the explanation of E, then B is also a constitutive ground of E.

C. Some visual experiences are constitutively grounded by events that include parts of the non-bodily world.

The conclusion is the extended visual consciousness thesis or EVC. P2 draws a connection between coupling and grounding. For A and B to be coupled with respect to explaining E is for interactions between parts of A and parts of B to contribute to causing and sustaining E. In cases where just A is a constitutive ground of E and A is coupled to B it is also the case that B can be decoupled from A with respect to grounding E. In cases where they can't be decoupled, A isn't a full constitutive ground on its own. When we are dealing with a particular explananda and two events cannot be separated or decoupled relative to each other's role in explaining and non-causally determining it, both count as partial constitutive grounds. The full grounds, in such cases, will include both. It is important that we relativize this claim to a particular explananda. After all, it will not generally be metaphysically impossible to suppose just A or just B to occur on its own, given some characterization of each. But what's being claimed is that in such cases, A and B won't be of the relevant grounding type without each other. In the case at hand, our explananda are visual

experiences with relatively rich phenomenal characters. Relative to providing token-explanations of such experiences, the grounding roles played by brain events and the environmental events to which they are actively coupled cannot be understood in isolation.

The crucial premise here is P1. What is the modal strength of ‘cannot’ in P1? If we consider how and why the relevant non-instantaneous event occurs, what we find is an unfolding with an internal causal structure which could not remain in place without bi-directional coupling to the environment. Again, it is not that somehow in isolation we cannot understand how this activity could occur on its own, but rather that once it is detached from its ordinary dynamic context what we are dealing with is some kind of type-breaking gerrymandered randomness. In other words, we have no basis for insisting that entirely on its own it is of the relevant type capable of playing its ordinary vehicular role. Let’s consider two ways to go about arguing for P1.

Despite the fact that Clark and Chalmers restrict EM to unconscious mental phenomena, Wilson (2010a, 277) argues that “perceptual domains in general seem to be pre-adapted (as it were) as likely domains for which the extended mind thesis is defensible.” By comparing vision to bat echolocation and the navigational systems of electric fish, he argues that they are all best seen as external sensory systems. But if Wilson is right that these perceptual states and processes sometimes have extended vehicles, why should we think there’s always a conscious part that remains inside? Once one acknowledges that some occurrent perceptual processes are extended, it seems unprincipled to insist that the conscious aspect of those states constitutively depends only on what happens inside the subject’s head.

In explaining why his argument for EVC does not commit the coupling-constitution fallacy, Wilson writes (2010a, 287): “it begins with a claim about a function of vision and how that function is achieved, and concludes with a claim about not the character of any component of the resulting system but with the character of that system itself.” His point is that any real-world system capable of carrying out the action-guiding function of vision will require ‘integrative coupling’ between body and environment, and ‘require’ not merely as input or scaffolding. Rather, integrating with parts of the environment will sometimes be the only way for an agent to successfully see in the ways in necessary for ongoing successful action in the world.

Consider Wilson’s (2010a, 283) argument:

1. The function of some visual processes is to guide action via visual information.
 2. A primary way to achieve that function is through the active embodiment of visual processing (in a fairly strong sense).
 3. Visual processes are actively embodied (in that same fairly strong sense) just if in their normal operation in natural environments, these processes are coupled with bodily activities so as to form an *integrated system* with *functional gain*. But
 4. Visual processes that are actively embodied, in this sense, are also extended.
- Therefore,
5. Some visual processes, and the visual systems those processes physically constitute, are extended.

The first two premises of the argument tell us something about the functions of vision and how such functions are ordinarily achieved. Systems constituted by organisms acting in an environment would be capable of functioning in ways a disembodied brain or body lacking an environment would not. 3-5 are supposed to flesh out the implications of 1 and 2. On the step from 3 to 4, Wilson (2010a, 287) draws the comparison to bats and eels: “While their extended sensory systems are realized, in part, by sonic and electromagnetic fields that they generate through their bodily movements, our extended visual systems are realized, in

part, by optic flow fields that we generate through *our* bodily movements.” Of course we do not generate light as we move around, rather our movements change how light in the ambient optic array reflects off our bodies and surfaces in the surrounding environment. Key to this case then is re-conceptualizing flow changes in the ambient optic array to be sensory fields generated by the perceiving organism’s movements, including quite subtle eye and head movements. Conceptualizing seeing as primarily a type of action-guiding looking in this way encourages one to focus on the overall interactive relationship between brain, body, and world, rather than exclusively on giving a functional decomposition of just what happens within the brain. The reason is that if sensing is a kind of action or doing that unfolds over time, the functional role of some movements is to structure optic flow patterns in order to further drive internal motor activity in various ways. Just as turbo-driven engines use the exhaust they produce to go faster and thereby exploit outputs in an integrative way to enhance overall function, so too do visual systems within a subject’s brain functionally exploit its movements.

Significantly, Wilson’s function-based argument is compatible with standard mechanistic accounts of constitutive relevance like Craver’s (2007a, 2007b) mutual manipulability account. He writes: “The fact that there is a functionally gainful, integratively coupled system is compatible with the existence of identifiable parts, each with its own integrity and functions, and with the decomposition of that integrative system into those functional parts” (285). This makes Wilson’s case for EVC more continuous with standard explanatory practices in the cognitive sciences than the next I will consider. “Cannot” in P1 connotes ‘impossible to function the right way’ for Wilson; not ‘impossible to exist as the

relevant type of system'. Functioning the right way requires an active and continuous coupling between brain and environment and couldn't be achieved without both coupled components present. Even if we can assign identifiable component functions to each part, the whole requires all parts, or functional equivalents of them. Though such functioning can be preserved in an artificial environment, it could equally be preserved if, say, with a functional duplicate of the cortex made of silicon. But this threatens neither EVC nor the claim that brain events are in fact partial constitutive grounds of visual experiences. One reason I wish to avoid resting my case for P1 entirely on this sort of consideration is that such appeals functional reasoning, especially when it comes to consciousness, strikes many as tendentious. So it would worthwhile to say why decoupling brain from environment won't always leave a brain internal ground in tact that doesn't appeal overtly to functional considerations about conscious vision.

Let's consider a different strategy for defending P1 due to Silberstein and Chemero (2012). To say that the environment 'cannot' be decoupled, according to this strategy, is to say that the subject cannot remain the same type of system if decoupled. This may sound outlandish, clearly the organism can continue to exist if we take it out of its particular niche, but the point is rather that it could not exist as the same 'synergy', even if all the internal events, by some cosmic coincidence, happened to be preserved. What is a synergy? A synergy is a system whose states and processes are a result of interaction dominant coupling among its components. It is thus said to be 'softly assembled' meaning that it attains some kind of temporary functional cohesion. What's important about such systems, at present, is that they generally resist decompositional explanations in which the behavior of the whole

can be accounted for in terms of well-understood functions of their parts along with relatively straightforward interactive relations among parts. If the minimal system to which we can accurately ascribe visual experiences to is ever a synergy that includes components in the environment, EVC will follow. To be clear, it is intrinsic brain activity that can be said to softly assemble the relevant synergy via bi-directional coupling. Yet without a niche, in many cases, it would be impossible for the relevant synergy to come into existence and persist long enough to constitutively ground phenomenally rich visual experiences, e.g. ones with dynamic, affordance-laden, contents.

Dynamical systems theories of cognition have long repudiated computational models of cognitive processing in favor of models employing differential equations in which the state of the agent and the state of the environment are inter-defined. Thus instead of thinking of them as two distinct but causally interacting systems, we gain a perspective in which we have a single, relatively closed, system. The boundaries of this system are in constant flux; ones that are centered on but by no means delimited by brain and body. In acts of coordination, including visuomotor coordination, a moving acting body is coupled together by intrinsic neural activity with the ambient optic array in order to softly assemble a synergy. This becomes especially plausible when the systems in question are highly nonlinear. Since nonlinear systems resist decompositional explanations, synergies cannot be treated as a collection of uncoupled parts that interact in well-defined ways.

If we're dealing with a synergy though, how do we go about localizing the system's properties at all? One promising suggestion that Silberstein and Chemero (2012) develop, following up on Van Orden, Holden, and Turvey (2003), is based on the idea that

interaction dominant systems exhibit pink noise. Pink noise, or $1/f$ noise, occurs when the signal has equal energy per octave. In other words, its energy per Hz, or power spectrum density, is inversely proportional to its frequency. Visible light in this power spectrum looks pink. In one study aimed at investigating Heidegger's (1962) phenomenological notion of 'readiness-to-hand', Dotov, Nie, and Chemero (2010) found that pink noise could be detected at a computer mouse when subjects are playing a computer game and it is functioning normally. While playing the game, however, the connection between the mouse and the monitor displaying the game was temporarily disrupted at irregular intervals. When this happened, pink noise could no longer be detected in the movement of the mouse. When the mouse is ready-to-hand, subjects are unaware of the hand-mouse interface and we see evidence of an interaction-dominant system that extends beyond the body. During the periods where mouse-monitor connectivity was disrupted, subjects became conscious of the mouse itself and no longer synergized with it. So for short-lived interaction dominant systems, where decompositional explanatory strategies are unlikely to succeed, measuring for pink noise can help indicate the extent of extension.

5.2 Why Bodily Duplicates Decoupled from Any Environment Are Insufficient to Constitutively Ground Most Pedestrian Visual Experiences

In the forward to Clark (2008), Chalmers (2008, xiv-xv) writes:

An argument for extended consciousness would require twins with different states of consciousness: Olga and Twin Olga are internal duplicates, but what it is like to be Olga differs from what it is like to be Twin Olga. But no matter how hard one tries to construct an Otto-style story that works like this, the story does not seem to succeed. Perhaps part of the reason is that the physical basis of consciousness requires direct access to information on an extremely high bandwidth... But our

low-bandwidth conscious connection to the environment seems to have the wrong form as it stands.

I will return to consider whether bandwidth considerations are sufficient to screen off the ambient optic array in all cases shortly, but now it is time to address the issue of duplication. If only events involving the environment are sufficient to constitutively ground some types of experiences on some occasions, then it should be possible to alter experience simply by altering what happens outside of the subject's brain and body. The fact that this seems intuitively implausible is connected to the well-known brain-in-a-vat objection. Would not a brain-in-a-vat be capable of having all the same conscious visual experiences as you or I without any environment at all? Upon considering the objection, and bearing in mind that we are concerned with vehicles rather than content or character, it should become clear that it has much less force than meets the eye. After all, supervenience thought experiments involving brains in vats usually take it for granted that the brain-in-a-vat is appropriately coupled to something non-biological. Provided that there is something external to the brain (and body) that is coupled to it in the same way a real environment ordinarily is, this fails to threaten EVC.

Nevertheless, the positive challenge Chalmers raises has force. If EVC is true, it should be possible to say how changing something about the environment can alone change the character of someone's experience. After all, changing what was written in Otto's notebook was sufficient to change his dispositional belief; Twin Otto believes the MoMA is a few blocks from where Otto believes it is. Nevertheless, the ease of constructing a twin case for Otto seems to rest mainly on the fact that dispositional mental states can have instantaneous constitutive grounds. When we are dealing with the

temporally extended vehicles of occurrent mental events, the burden to make it intuitive that experience would be different in some particular specifiable way dissipates. All that's important is that it is plausible there would be some difference.

Let's consider why this is by looking the other way. Suppose Martha is seeing red. Twin-Martha is almost an exact duplicate of Martha. In fact everything about Twin-Martha's brain and body, and even the world Twin-Martha inhabits, are exactly the same as Martha's, with one exception. Whereas activity in Martha's V4 corresponds to the sort of activity usually found in subjects who are actually seeing red, activity in Twin-Martha's V4 instead corresponds to activity that would ordinarily be part of what constitutively grounds yellowish experiences. Once again, everything else about Twin-Martha, including what happens everywhere else in her brain is exactly the same as Martha's. To make the case coherent, we may assume we are dealing with something that is merely metaphysical possible (though nomologically impossible); or perhaps even a nomologically possible cosmic coincidence. Either way, it does not matter. We are dealing with a situation that could only arise in some very distant possible world. It seems fair to say we have no good reason to believe that Twin-Martha sees red in this scenario, so there would be difference between her visual experiences and Martha's. But it also seems unfair to demand that we say exactly what her experience would be like. There's certainly no good reason, given what else is happening in her brain, to think that she's seeing yellow. Perhaps she sees in black and white, or perhaps not at all.

Analogously, if the ambient environment is totally decoupled from Twin Olga's brain in the way it would need to be for isolating an external difference without an internal

difference (i.e. it's not simply changed from a real world to a simulacrum or alternative realizer), it's fair to insist there would be a difference in Olga and Twin Olga's visual experiences without there being the burden of saying exactly what that difference would be. If Olga's ambient optical environment is a forest and she's enjoying rich, affordance-laden, visual experiences, but Twin Olga's ambient optical environment is a city, then, despite being bodily duplicates (by some cosmic coincidence), Twin Olga's experiences will likely be different. Even if the events right at the periphery of Twin Olga's body were randomly the same as Olga's, and even if her visual experiences were more forest-like than city-like, she would not be phenomenally identical to Olga since she would no longer be engaged in the same acts of sensorimotor coordination with her environment.

The phenomenal characters of visual experiences include affordances and in other ways are individuated by effects downstream from early visual processing, e.g. in part by efferent motor commands, in part by comparisons to re-afferent feedback caused by bodily movements and exogenous environmental changes. Visual experiences with characters sufficient to explain successful engagements in an environment often have vehicles in which early visual processing is actively and intricately coupled to changes in the ambient optic array. When both hook-up points, motor output and sensory input, are stipulated to be unplugged or decoupled from brain internal activity, we can no longer assume the brain internal activity by itself is playing its ordinary realizing or grounding role.

Recall the challenge Chalmers issued for proponents of ECM and his suggested diagnosis. Taking up this line and citing Eliasmith (2008, 150) on how the dynamics of activity within the brain is qualitatively different than the dynamics between brain activity

and embodied action in an environment, Clark (2009, 986) objects to ECM, “if indeed the physical machinery of conscious experience requires fast timescale operations and processing, and the non-neural body acts as a low pass filter preventing external (and internal, muscular, etc.) signals from directly entering into such operations and processing, then such signals are fit to play only a causal role.” But as Thompson and Cosmelli (2011, 177) observe: “[t]he time it takes for visual stimulation to pass through the lens of the moving eye and reach the first stages of visual processing is a fraction of the time it takes for neural systems to build up any correlated activity.” The point is not that there are no relevant interfaces marked by sense organs, it is that the flow of information across pre-established interfaces can sometimes create or softly-assemble newer, wider, problem solving wholes. The time it takes for light to traverse the ambient optic array and for retinal stimulation to be relayed to the occipital lobe is quick enough to establish a high-bandwidth connection between what is outside the body and what is inside the head.

5.3 Conclusion

EVC implies ECM, which in turn implies EM. So, if what I’ve argued is correct, this is a striking vindication of active vehicle externalism or the extended mind thesis. While there may be other ways to argue for EVC, e.g. by appeal to neural deference or perceptual presence as access, I think that the dynamical entanglement argument is sound and provides a sufficient reason for thinking that the vehicles of some of our visual experiences are not confined to our heads. When dealing with visual experiences with temporally extended grounds, especially lengthy ones, the role of the non-bodily world

cannot be relegated to the causal background. The causal interactions between brain, body, and world become equally important to constitutively grounding such conscious visual experiences as causal interactions anywhere within the brain.

REFERENCES

- Adams, F. & Aizawa, K. (2001). The bounds of cognition. *Philosophical Psychology*, *14*(1), 43-64.
- Adams, F. & Aizawa, K. (2008). *The bounds of cognition*. Malden, MA: Blackwell Publishing.
- Adams, Z. & Maher, C. (2012). Cognitive spread: Under what conditions does the mind extend beyond the body? *European Journal of Philosophy*. DOI: 10.1111/j.1468-0378.2012.00543.x.
- Aizawa, K. (2010). Consciousness: don't give up on the brain. *Royal Institute of Philosophy Supplement*, *85*(67), 263-284.
- Austin, J. (1962). *Sense and sensibilia*. Oxford: Oxford University Press.
- Bach-y-Rita, P. (1972). *Brain mechanisms in sensory stimulation*. New York: Academic Press.
- Baker, L. (2009). Persons and the extended mind thesis. *Zygon*, *44*(3), 642-658.
- Bechtel, W. & Richardson, R. (1993/2010). *Discovering complexity: Decomposition and localization as strategies in scientific research*. Princeton: Princeton University Press.
- Beer, R. (1995). A dynamical systems perspective on agent-environment interactions. *Artificial Intelligence*, *72*, 173-215.
- Block, N. (1986). Advertisement for a semantics for psychology. *Midwest Studies in Philosophy*, *10*(1), 615-678.
- Block, N. (1990). Inverted earth. *Philosophical Perspectives*, *4*, 53-79.
- Bloom, P. (2005). *Descartes' baby: How the science of child development explains what makes us human*. New York: Basic Books.
- Burge, T. (1986). Individualism and psychology. *Philosophical Review*, *95*, 3-45.
- Burge, T. (1993). Mind-body causation and explanatory practice. In J. Heil & A. Mele (Eds.), *Mental Causation*. New York: Oxford University Press.
- Burge, T. (2010). *Origins of objectivity*. New York: Oxford University Press.

- Byrne, A. & Logue, H. (2008). *Disjunctivism: Contemporary readings*. Cambridge, MA: MIT Press.
- Byrne, A. & Tye, M. (2006). Qualia ain't in the head. *Noûs*, 40(2), 241-255.
- Carruthers, P. (2000). *Phenomenal consciousness: A naturalistic theory*. Cambridge, UK: Cambridge University Press.
- Chalmers, D. (1996). *The conscious mind: In search of a fundamental theory*. New York: Oxford University Press.
- Chalmers, D. (2004). The representational character of experience. In B. Leiter (Ed.), *The Future for Philosophy*. New York: Oxford University Press.
- Chalmers, D. (2008). Forward to *Supersizing the mind* by Andy Clark. New York: Oxford University Press.
- Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge, MA: MIT Press.
- Chomsky, N. (1959). Reviews: *Verbal Behavior* by B.F. Skinner. *Language*, 35(1), 26-58.
- Clark, A. (1989). *Microcognition: Philosophy, cognitive science, and parallel distributed processing*. Cambridge, MA: MIT Press.
- Clark, A. (1997). *Being there: Putting brain, body, and world together again*. Cambridge, MA: MIT Press.
- Clark, A. (1998). Time and mind. *Journal of Philosophy*, 95(7), 354-376.
- Clark, A. (1999). Visual awareness and visuomotor action. *Journal of Consciousness Studies*, 6(11-12), 1-18.
- Clark, A. (2003). *Natural-born cyborgs: Minds, technologies, and the future of human intelligence*. New York: Oxford University Press.
- Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. New York: Oxford University Press.
- Clark, A. (2009). Spreading the joy? Why the machinery of consciousness is (probably) still in the head. *Mind*, 118(472), 963-993.
- Clark, A. (2010a). Memento's revenge: The extended mind, extended. In R. Menary (Ed.), *The Extended Mind*. Cambridge, MA: MIT Press.

- Clark, A. (2010b). Coupling, constitution, and the cognitive kind: A reply to Adams and Aizawa. In R. Menary (Ed.), *The Extended Mind*. Cambridge, MA: MIT Press.
- Clark, A. (2012). Dreaming the whole cat: Generative models, predictive processing, and the enactivist conception of perceptual experience. *Mind*, *121*(483), 753-771.
- Clark, A. & Chalmers, D. (1998). The extended mind. *Analysis*, *58*(1), 7-19.
- Clark, M. & Liggins, D. (2012). Recent work on grounding. *Analysis*, *72*(4), 812-823.
- Correia, F. & Schnieder, B. (2012). *Metaphysical grounding: Understanding the structure of reality*. Cambridge, UK: Cambridge University Press.
- Cosmelli, D., Lachaux, J., & Thompson, E. (2007). Neurodynamical approaches to consciousness. In P. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *The Cambridge Handbook of Consciousness*. New York: Cambridge University Press.
- Craver, C. (2007a). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. Oxford: Clarendon Press.
- Craver, C. (2007b). Constitutive explanatory relevance. *Journal of Philosophical Research*, *32*, 3-20.
- Crick, F. & Koch, C. (2003). A framework for consciousness. *Nature Neuroscience*, *6*(2), 119-126.
- Dainton, B. (2010). Temporal consciousness. In E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*.
<http://plato.stanford.edu/archives/fall2010/entries/consciousness-temporal/>.
- Davidson, D. (1969). The individuation of events. In N. Rescher (Ed.), *Essays in Honor of Carl G. Hempel*. Dordrecht: Reidel.
- Davidson, D. (1987). Knowing one's own mind. *Proceeding of the American Philosophical Association*, *60*(3), 441-458.
- Davies, M. (1991). Individualism and perceptual content. *Mind*, *100*(399), 461-484.
- Dennett, D. (1978). *Brainstorms: Philosophical essays on mind and psychology*. Montgomery, VT: Bradford Books.
- Dennett, D. (1991). *Consciousness explained*. Boston, MA: Little, Brown.
- Dewey, J. (1896). The reflex arc concept in psychology. *Psychological Review*, *3*(4), 357-

370.

Dewey, J. (1916). *Democracy and education*. New York: Macmillan.

Dotov, D., Nie, L., & Chemero, A. (2010). A demonstration of the transition from readiness-to-hand to unreadiness-to-hand. *PLoS ONE*, *5*(3).
DOI:10.1371/journal.pone.0009433.

Dretske, F. (1981). *Knowledge and the flow of information*. Oxford: Blackwell Publishing.

Dretske, F. (1995). *Naturalizing the mind*. Cambridge, MA: MIT Press.

Dreyfus, H. (1982). Introduction. In H. Dreyfus & H. Harrison (Eds.), *Husserl, Intentionality, and Cognitive Science*. Cambridge, MA: MIT Press.

Eagleman, D. & Sejnowski, T. (2000). Motion integration and postdiction in visual awareness. *Science*, *287*(5460), 2036-2038.

Egan, F. (1992). Individualism, computation, and perceptual content. *Mind*, *101*(403), 443-459.

Eliasmith, C. (2008). Dynamics, control, and cognition. In P. Robbins & M. Aydede (Eds.), *Cambridge Handbook of Situated Cognition*. Cambridge, UK: Cambridge University Press.

Field, H. (1978). Mental representation. *Erkenntnis*, *13*(1), 9-61.

Fine, K. (1995). Ontological dependence. *Proceedings of the Aristotelian Society*, *95*, 269 - 290.

Fine, K. (2001). The question of realism. *The Philosophers' Imprint*, *1*(2), 1-30.

Fodor, J. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.

Fodor, J. (1980). Methodological solipsism considered as a research strategy in cognitive psychology. *Behavioral and Brain Sciences*, *3*(1), 63-73.

Fodor, J. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.

Fodor, J. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind*. Cambridge, MA: MIT Press.

Fodor, J. (1994). *The elm and the expert: Mentalese and its semantics*. Cambridge, MA:

MIT Press.

Fodor, J. (2009). Where is my mind? Review of Andy Clark's *Supersizing the Mind*. *London Review of Books*, 31(3), 13-15.

Fodor, J. & Pylyshyn, Z. (1981). How direct is visual perception? Some reflections on Gibson's 'ecological approach'. *Cognition*, 9(2), 139-196.

Gallagher, S. (2005). *How the body shapes the mind*. Oxford: Oxford University Press.

Gallagher, S. & Varela, F. (2003). Redrawing the map and resetting the time: Phenomenology and the cognitive sciences. *Canadian journal of philosophy*, 29, 93-132.

Gangopadhyay, N. & Kiverstein, J. (2009). Enactivism and the unity of perception and action. *Topoi*, 28(1), 63-73.

Gibbs, R. (2001). Intentions as emergent products of social interactions. In B. Malle, L. Moses, & D. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition*. Cambridge, MA: MIT Press.

Gibson, J. (1966) *The senses considered as perceptual systems*. Boston: Houghton Mifflin.

Gibson, J. (1972). A theory of direct visual perception. In J. Royce & W. Rozenboom (Eds.). *The Psychology of Knowing*. New York: Gordon & Breach.

Gibson, J. (1979/1986). *The ecological approach to visual perception*. New York: Houghton-Mifflin.

Goldman, A. (1976). Discrimination and perceptual knowledge. *Journal of Philosophy*, 73(20), 771-791.

Grice, H. (1957). Meaning. *The philosophical review*, 66(3), 377-388.

Grice, H. (1968). Utterer's meaning, sentence-meaning, and word-meaning. *Foundations of language*, 4(3), 225-242.

Grice, H. (1969). Utterer's meaning and intention. *The philosophical review*, 78(2), 147-177.

Grush, R. (2003). In defense of some 'Cartesian' assumptions concerning the brain and its operation. *Biology and Philosophy*, 18(1), 53-92.

Grush, R. (2006). How to, and how not to, bridge computational cognitive neuroscience and Husserlian phenomenology of time consciousness. *Synthese*, 153(3), 417-450.

- Haddock, A. & Macpherson, F. (2008). *Disjunctivism: Perception, action, knowledge*. Oxford: Oxford University Press.
- Hafed, Z. & Krauzlis R. (2006). Ongoing eye movements constrain visual perception. *Nature Neuroscience*, 9(11), 1449-1457
- Harman, G. (1990). The intrinsic quality of experience. *Philosophical Perspectives*, 4, 31-52.
- Haugeland, J. (1998). Mind embodied and embedded. In *Having Thought: Essays in the Metaphysics of Mind*. Cambridge, MA: Harvard University Press.
- Hefft, H. (2001). *Ecological psychology in context: James Gibson, Roger Barker, and the legacy of William James's radical empiricism*. Mahwah, NJ: L. Erlbaum.
- Heidegger, M. (1962). *Being and time*. J. Macquarrie & E. Robinson (Trans.), London: SCM Press.
- Held, R. (1965). Plasticity in sensory-motor systems. *Scientific American*, 213(5), 84-94.
- Hurley, S. (1998). *Consciousness in action*. Cambridge, MA: Harvard University Press.
- Hurley, S. (2001). Perception and action: Alternative views. *Synthese*, 129(1), 3-40.
- Hurley, S. (2010). Varieties of externalism. In R. Menary (Ed.), *The Extended Mind*. Cambridge, MA: MIT Press.
- Hurley, S. & Noë, A. (2003). Neural plasticity and consciousness. *Biology and Philosophy*, 18(1), 131-168.
- Husserl, E. (1991). On the phenomenology of the consciousness of internal time (1893-1917). J. Brough (Trans.), *Husserliana: Edmund Husserl - Collected Works (Book 4)*. Dordrecht, Netherlands: Kluwer.
- James, W. (1890). *The principles of psychology*, New York: Dover.
- James, W. (1997). What pragmatism means. In R. Coles (Ed.), *Selected Writings*. New York: Book-of-the-Month Club.
- Kaplan, D. (2012). How to demarcate the boundaries of cognition. *Biology and Philosophy*, 27(4), 545-570.
- Kelly, S. (2005). Temporal awareness. In D. Woodruff Smith & A. Thomasson (Eds.),

- Phenomenology and Philosophy of Mind*. Oxford: Oxford University Press.
- Kelso, S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. Cambridge, MA: MIT Press.
- Kim, J. (1993). *Supervenience and mind: Selected philosophical essays*. Cambridge, UK: Cambridge University Press.
- Kirsh, D. & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, 18(4), 513-549.
- Koch, C. (2004). *The quest for consciousness*. New York: Roberts and Co.
- Kriegel, U. (2013). *Phenomenal intentionality*. New York: Oxford University Press.
- Kripke, S. (1980). *Naming and necessity*. Cambridge, MA: Harvard University Press.
- Ladyman, J. & Ross D. (2007). *Everything must go: Metaphysics naturalized*. Oxford: Oxford University Press.
- Lakoff, G. & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. New York: Basic books.
- Lewis, D. (1986). *The plurality of worlds*. London: Blackwell Publishing.
- Lewis, D. & Langton, R. (1998). Defining 'intrinsic'. *Philosophy and Phenomenological Research*, 58(2), 333-345.
- Loar, B. (1981). *Meaning in mind*. Cambridge, UK: Cambridge University Press.
- Lowe, E. (1996). *Subjects of experience*. Cambridge UK: Cambridge University Press.
- Lowe, E. (2010). Ontological dependence. In E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/archives/spr2010/entries/dependence-ontological/>.
- Lutz, A. & Thompson, E. (2003). Neurophenomenology: Integrating subjective experience and brain dynamics in the neuroscience of consciousness. *Journal of Consciousness Studies*, 10(9-10), 31-52
- Machamer, P., Darden, L., & Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science*, 67(1), 1-25.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York: W.H. Freeman and Company.

- Maturana, H. & Varela, F. (1980). *Autopoiesis and cognition: The realization of the living*. Dordrecht/Boston: D. Riedel.
- Maturana, H. & Varela, F. (1987). *The tree of knowledge*. Boston: New Science Library.
- McGinn, C. (1982). The structure of content. In A. Woodfield (Ed.), *Thought and Object: Essays on Intentionality*. Oxford: Oxford University Press.
- McGinn, C. (1989). *Mental Content*. Oxford: Blackwell Publishing.
- McGinn, C. (1995). Consciousness and space. *Journal of Consciousness Studies*, 2(3), 220-230.
- McLuhan, M. (1964). *Understanding media: The extensions of man*. New York: McGraw Hill.
- Menary, R. (2006). Attacking the bounds of cognition. *Philosophical Psychology*, 19(3), 329-344.
- Menary, R. (2007). *Cognitive integration: Attacking the bounds of cognition*. Basingstoke, Hampshire, UK: Palgrave Macmillan.
- Menary, R. (2010). Cognitive integration and the extended mind. In R. Menary (Ed.), *The Extended Mind*. Cambridge, MA: MIT Press.
- Metzinger, T. (2003). *Being no one*. Cambridge, MA: MIT Press.
- Millikan, R. (1984). *Language, thought, and other biological categories*. Cambridge, MA: MIT Press.
- Millikan, R. (1989). Biosemantics. *Journal of Philosophy*, 86(6), 281-297.
- Millikan, R. (1993). *White queen psychology and other essays for Alice*. Cambridge, MA: MIT Press.
- Millikan, R. (2004). *Varieties of meaning*. Cambridge, MA: MIT press.
- Moore, G. (1922). *Philosophical Studies*. London: Trench, Trubner, & Co.
- Mossio, M. & Taraborelli, D. (2008). Action-dependent perceptual invariants: From ecological to sensorimotor approaches. *Consciousness and cognition*, 17(4), 1324-1340.
- Müller, J. (1843/2003). *Elements of Müller's Physiology*. W. Baly (Trans.), Bristol:

Thoemmes Press.

- Myin, E. & O'Regan, J. (2002). Perceptual consciousness, access to modality and skill theories. A way to naturalize phenomenology? *Journal of Consciousness Studies*, 9(1), 27-46.
- Newell, A. & Simon, H. (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the ACM*, 19(3), 113-126.
- Nisly-Nagele, S. & Wasserman, G. (2001). Dissociating sensory and cognitive contributions to visual persistence I. Photoreceptor response duration as a function of flash intensity, adaptation state and candidate code. *Biological Cybernetics*, 85, 167-83.
- Noë, A. (2004). *Action in perception*. Cambridge, MA: The MIT Press.
- Noë, A. (2006). Experience without the head. In T. Gendler & J. Hawthorne (Eds.), *Perceptual Experience*. Oxford: Oxford University Press.
- Noë, A. (2007). Magic realism and the limits of intelligibility: What makes us conscious? *Philosophical Perspectives*, 21(1), 457-474.
- Noë, A. (2009). *Out of our heads: Why you are not your brain, and other lessons from the biology of consciousness*. New York: Hill and Wang.
- Noë, A. & O'Regan J. (2002). On the brain-basis of visual consciousness: a sensorimotor account. In A. Noë & E. Thompson (Eds.), *Vision and Mind: Selected Readings in the Philosophy of Perception*. Cambridge, MA: MIT Press.
- Noë A. & Thompson, E. (2004). Are there neural correlates of consciousness? *Journal of Consciousness Studies*, 11, 3-28.
- Nolan, D. (2013). Hyperintensional metaphysics. *Philosophical Studies*. DOI: 10.1007/s11098-013-0251-2.
- Nozick, R., (1981). *Philosophical Explanations*, Cambridge MA: Belknap Press.
- O'Regan, J. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology*, 46(3), 461-488.
- O'Regan, J. (2010). Explaining what people say about sensory qualia. In N. Gangopadhyay, M. Madary, & F. Spicer (Eds.), *Perception, Action, and Consciousness: Sensorimotor Dynamics and Two Visual Systems*. Oxford: Oxford University Press.
- O'Regan, J. (2011). *Why red doesn't sound like a bell: Understanding the feel of*

- consciousness*. Oxford: Oxford University Press.
- O'Regan, J. & Noë, A. (2001a). A sensorimotor approach to vision and visual consciousness. *Behavioral and Brain Sciences* 24(5), 939-1031.
- O'Regan, J. & Noë, A. (2001b). What is it like to see: A sensorimotor theory of perceptual experience. *Synthese*, 129(1), 79-103.
- Olson, E. (2011). The extended self. *Minds and Machines*, 21(4), 481-495.
- Papineau, D. (1987). *Reality and representation*. Oxford: Blackwell Publishing
- Pascal, F. & O'Regan, J. (2008). Commentary on Mossio and Taraborelli: Is the enactive approach really sensorimotor? *Consciousness and Cognition*, 17(4), 1341-1342.
- Pautz, A. (2006). Sensory awareness is not a wide physical relation: An empirical argument against externalist intentionalism. *Noûs*, 40(2), 205-240.
- Prinz J (2006). Putting the brakes on enactive perception. *Psyche* 12(1), 1-19.
- Pearl, J. (2000). *Causality: Models, reasoning and inference*. Cambridge, MA: MIT press.
- Putnam, H. (1973). Meaning and reference. *Journal of Philosophy*, 70(19), 699-711.
- Putnam, H. (1975). The meaning of 'meaning'. *Minnesota Studies in the Philosophy of Science*, 7, 131-193.
- Putnam, H. (1982). *Reason, truth and history*. Cambridge, UK: Cambridge University Press.
- Putnam, H. (1999). *The threefold cord: Mind, body and world*. New York: Columbia University Press.
- Pylyshyn, Z. (2001). Seeing, acting, and knowing. *Behavioral and Brain Sciences*, 24(5), 999.
- Rey, G. (1998). A narrow representationalist account of qualitative experience. In J. Tomberlin (Ed.), *Philosophical Perspectives (Volume 12: Language, Mind, and Ontology)*, Atascadero, CA: Ridgeview Publishing.
- Rockwell, W. (2005). *Neither brain nor ghost: A nondualist alternative to the mind-brain identity theory*. Cambridge, MA: The MIT Press.
- Ross, D. & Ladyman, J. (2010). The alleged coupling-constitution fallacy and the mature

- sciences. In R. Menary (Ed.), *The Extended Mind*. Cambridge, MA: MIT Press.
- Rowlands, M. (1999). *The body in mind: Understanding cognitive processes*. Cambridge, MA: Cambridge University Press.
- Rowlands, M. (2003). *Externalism: Putting mind and world back together again*. Chesham, Bucks: Acumen/McGill-Queen's University Press.
- Rowlands, M. (2009). Enactivism and the extended mind. *Topoi*, 28(1), 53-62.
- Rowlands, M. (2010). *The new science of the mind: From extended mind to embodied phenomenology*. Cambridge, MA: MIT 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*. Oxford: Oxford University Press.
- Russell, B. (1913/1984). On the experience of time. In *The Collected Papers of Bertrand Russell*, vol. 7, London: Allen and Unwin.
- Ryle, G. (1949). *The concept of mind*. London: Hutchinson.
- Ryle, G. (1954). *Dilemmas*. Cambridge, UK: Cambridge University Press.
- Schaffer, J. (2009). On what grounds what. In D. Chalmers, D. Manley & R. Wasserman, (Eds.), *Metametaphysics: New Essays on the Foundations of Ontology*. Oxford: Oxford University Press.
- Scholl, B. & Simons, D. (2001). Change blindness, Gibson and the sensorimotor theory of vision. *Behavioral and Brain Sciences*, 24(5), 1004-1005.
- Searle, J. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge, UK: Cambridge University Press.
- Segal, G. (1989). Seeing what is not there. *Philosophical Review*, 98, 189-214.
- Shapiro, L. (1993). Content, kinds, and individualism in Marr's theory of vision. *Philosophical Review*, 102(4), 489-513.
- Shapiro, L. (2004). *The mind incarnate*. Cambridge, MA: MIT press.
- Shapiro, L. (2011). *Embodied Cognition*. New York: Routledge.

- Silberstein, M. & Chemero, A. (2012). Complexity and extended phenomenological-cognitive systems. *Topics in cognitive science*, 4(1), 35-50.
- Simon, H. (1969). *The sciences of the artificial*. Cambridge, MA: MIT Press.
- Sporns, O. (2011). *Networks of the brain*. Cambridge, MA: MIT Press.
- Sterelny, K. (2004). Externalism, epistemic artefacts and the extended mind. In R. Schantz (Ed.), *The Externalist Challenge*. New York: de Gruyter.
- Sutton, J. (2010). Exograms and interdisciplinarity: History, the extended mind, and the civilizing process. In R. Menary (Ed.), *The Extended Mind*. Cambridge, MA: MIT Press.
- Taraborelli, D. & Mossio, M. (2008). On the relation between the enactive and the sensorimotor approach to perception. *Consciousness and Cognition*, 17(4), 1343-1344.
- Thau, M. (2002). *Consciousness and Cognition*, Oxford: Oxford University Press.
- Thompson, E. (1995). *Colour Vision*. London: Routledge.
- Thompson, E. (2007). *Mind in Life*. Cambridge, MA: Harvard University Press.
- Thompson, E. & Cosmelli, D. (2011). Brain in a vat or body in a world? Brainbound versus enactive views of experience. *Philosophical Topics*, 39(1), 163-180.
- Thompson, E. & Varela, F. (2001). Radical embodiment: Neural dynamics and consciousness. *Trends in Cognitive Science*, 5(10), 418-425.
- Tye, M. (1995). *Ten problems of consciousness: A representational theory of the phenomenal mind*. Cambridge, MA: MIT Press.
- Tye, M. (2000). *Consciousness, color, and content*. Cambridge, MA: MIT Press.
- Tye, M. (2002). Representationalism and the transparency of experience. *Noûs*, 36(1), 137-151.
- Van Gelder, T. (1995). What might cognition be, if not computation? *Journal of Philosophy*, 92(7), 345-81.
- van Orden, G., Holden, J., & Turvey, M. (2003). Self-organization of cognitive performance. *Journal of Experimental Psychology: General*, 132, 331-351.

- VanRullen, R. & Koch, C. (2003). Is perception discrete or continuous? *Trends in Cognitive Science*, 7(5), 207-213.
- Varela, F. (1999). The specious present: A neurophenomenology of time consciousness. In J. Petitot, F. Varela, B. Pachoud, & J. Roy (Eds.), *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science*. Stanford: Stanford University Press.
- Varela, F., Thompson, E., & Rosch, E. (1991). *The embodied mind*. Cambridge, MA: MIT Press.
- Varela, F. & Maturana, H. (1973). De máquinas y seres vivos: Una teoría sobre la organización biológica. *Santiago de Chile: Editorial Universitaria*.
- Velliste, M., Perel, S., Spalding, M., Whitford, A., & Schwartz, A. (2008). Cortical control of a prosthetic arm for self-feeding. *Nature*, 453(7198), 1098-1101.
- Von der Malsburg, C., Phillips, W., & Singer, W. (2010). *Dynamic coordination in the brain: from neurons to mind*. Cambridge, MA: The MIT Press.
- Ward, D. (2012). Enjoying the spread: Conscious externalism reconsidered. *Mind*, 121(483), 731-751.
- Wexler, M. & Van Boxtel, J. (2005). Depth perception by the active observer. *Trends in Cognitive Science*, 9, 431- 438.
- Wilson, R. (1994). Wide computationalism. *Mind*, 103, 351-372.
- Wilson, R. (1995). *Cartesian psychology and physical minds*. New York: Cambridge University Press.
- Wilson, R. (2004), *Boundaries of the mind*. New York: Cambridge University Press.
- Wilson, R. (2005). *Genes and the agents of life*. New York: Cambridge University Press.
- Wilson, R. (2010a). Extended vision. In N. Gangopadhyay, M. Madary, & F. Spicer (Eds.), *Perception, Action, and Consciousness: Sensorimotor Dynamics and Two Visual Systems*. Oxford: Oxford University Press.
- Wilson, R. (2010b). Meaning making and the mind of the externalist. In R. Menary (Ed.), *The Extended Mind*. Cambridge, MA: MIT Press.
- Williamson, T. (2006). Can cognition be factorized into internal and external components? In R. Stainton (Ed.), *Contemporary Debates in Cognitive Science*. Malden, MA:

Blackwell Publishing.

Wittgenstein, L. (1953). *Philosophical investigations*. G. Anscombe (Trans.), Oxford: Blackwell Publishing.

Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford: Oxford University Press.