10

THE BOUNDARIES OF THE MIND[1]

Katalin Farkas

1. The Cartesian conception of the mind

The subject of mental processes or mental states is usually assumed to be an individual, and hence the boundaries of mental features – in a strict or metaphorical sense – are naturally regarded as reaching no further than the boundaries of the individual. This chapter addresses various philosophical developments in the 20th and 21st century that questioned this natural assumption. I will frame this discussion by first presenting a historically influential commitment to the individualistic nature of the mental in Descartes' theory. I identify various elements in the Cartesian conception of the mind that were subsequently criticized and rejected by various externalist theories, advocates of the extended mind hypothesis and defenders of embodied cognition. Then I will indicate the main trends in these critiques.

Descartes' work was partly a response to developments in natural science in the 17th century, and one of his goals was to provide a theoretical-philosophical foundation for modern science which rejected Aristotelian natural philosophy. Descartes was not the last philosopher who hoped to make a lasting contribution by providing a theory that integrates philosophy with modern science. Ever since the 17th century, there has been an occasionally reoccurring anxiety in philosophy – famously expressed for example by Kant in the *Critique of Pure Reason* – about the fact that while the sciences appear to make great progress, philosophy in comparison seems to make very little, if any progress. Various remedies have been suggested, for example by offering a methodology or grounding for philosophy that is either imported from the sciences, or is comparable to the objectivity and explanatory power of scientific method. This will be one of the persisting themes in this chapter. In section 2, we will see how the attempt to ground philosophical theories by semantics offers the hope of progress and questions the Cartesian boundaries of the mind at the same time. In section 3, we will see how naturalism about the mind also leads to reconsidering the issue of boundaries.

Descartes' considerations about the nature of the mind in his Second Meditation have had a profound effect on the development of philosophy in the Western tradition.[2] In the First Meditation, Descartes considers the possibility of being deceived by an evil demon. The starting point is the simple observation that things

can appear different from the way they are. Extending the gap between appearance and reality to its extreme, it is possible, contends Descartes, that the world I take to be around me does not exist, that even my body does not exist, but all appearances of a world and my body are results of the manipulations of an evil demon. If I were the victim of the demon, things would appear exactly the same as they do now, but in reality, they would be very different. The modern version of this scenario of radical deception is usually known as the "brain-in-a-vat" or "the Matrix" scenario, where the stimuli arriving to our brain from external objects or the rest of our body are replaced by an elaborate machinery of virtual reality.

Even in this case of radical deception, Descartes argues, I would still be a thinking thing; moreover, a thinker who would have exactly the same mental features as I do now. The sky and earth may not exist, but I would still *believe* they do; the light I seem to see and the noise I seem to hear may not be there, but I would still *feel* that they are. Two points are worth emphasizing here. First, even though the demon scenario allows for the non-existence of my body, Descartes explicitly says that, at this point, he is still agnostic about the existence of a mind separable from the body. So the resulting conception is not committed to dualism about mind and body. Second, there is an important epistemic asymmetry between my mind and the rest of the world: namely, the possibility of being deceived by a demon, or being a brain in a vat, threatens, at least prima facie, my knowledge of the rest of the world, but it does not threaten my knowledge of the content of my mental states. This does not necessarily mean that I am omniscient or infallible about the nature of my mind; but it does mean that at least in *this comparison* – i.e. vulnerability to a threat from the demon scenario – the mind fares *better* than the body. The title of the Second Meditation is fittingly "The nature of the human mind, and how it is better known than the body".

Eventually, by the end of the *Meditations*, Descartes arrives at the conclusion that the mind is indeed distinct from the body. At the same time, Descartes was very interested in the body's contribution to our mental life, and his view can be summarized (with some simplification) as follows. There are mental phenomena, most notably, sensory (perceptual and bodily) experiences, and emotions, which are caused by the brain. Nerves from our sense-organs and throughout the body carry stimuli as far as the pineal gland in the brain. The gland has a distinctive state for each type of sensory or affective experience and causes our immaterial mind to undergo that experience. How this causal connection works between two entirely different substances is of course one of the greatest puzzles for a Cartesian dualist, but this is not the topic of this chapter (see Chapters 2 and 7). It is more interesting for our purposes that apparently Descartes thought that non-sensory or non-affective cognition (for example, pure theoretical and practical reasoning) need not, or indeed could not, involve the body in such an intimate way (Farkas 2005). So there is a separation between what we may call "pure" cognition, on the one hand, and sensory and affective mental states, on the other. This separation, in itself, is not Descartes' invention. Aristotle draws a similar distinction between rational and non-rational parts of the soul.

I provide this sketch of Descartes' conception of the mind because much of the following discussion will be usefully understood by the various critiques' deep or superficial disagreement with certain elements of the Cartesian conception. In one way or another, many philosophers in the 20th and 21st century objected to the idea that a solitary mind deceived by an evil demon or evil scientist can have the same mental features as we do; or to the idea that the physical basis of mental phenomena can be restricted to the brain; or to the idea that pure cognition is largely independent of the body. This chapter focuses on developments in the analytic philosophical tradition (for a critique of the Cartesian conception in continental philosophy, see, for example, Hubert Dreyfus's commentary on Heidegger's *Being and Time* (Dreyfus 1991). I will not attempt to reconstruct all pros and cons in the debates I mention; for state-of-the-art summaries of these issues, it is worth consulting, for example, Lau and Deutsch (Wilson and Foglia (2011). Instead, I shall try to trace some broad historical tendencies that influenced philosophical thinking on the boundaries of the mind.

2. "Externalism" or "anti-individualism"

2.1 The semantic tradition

Gottlob Frege is often identified as one of the first and most influential figures in the history of analytic philosophy.[3] Frege proposed the first systematic modern theory of semantics, that is, a theory of how the semantic values (the truth and reference) of linguistic expressions are determined. Semantics and symbolic logic have then gone on to become two of the greatest success stories in the history of analytic philosophy, and no doubt this is the reason why Frege is regarded as one of the founders of the tradition, despite the fact that he was a German thinker, deeply rooted in a philosophical tradition that was quite different from the English-speaking empiricism that forms a more congenial historical background to analytic philosophy. I mentioned above the occasionally reoccurring anxiety in philosophy about the apparent lack of progress in the several thousand years' history of the subject. Semantics and logic have been seen by many as finally offering the prospect for real progress in philosophy, not just in semantic theory itself, but also as a tool to get a better grip on a wide range of philosophical issues. We can find a vigorous expression of this sentiment, for example, in Timothy Williamson's 2004 paper "Must Do Better". Williamson's idea of progress in philosophy is well expressed in these complimentary words about Michael Dummett's contribution to the realism/anti-realism debate:

> Instead of shouting slogans at each other, Dummett's realist and anti-realist would busy themselves in developing systematic compositional semantic theories of the appropriate type, which could then be judged and compared by something like scientific standards.
>
> (Williamson 2006, 179)

258

In what we may call "the semantic tradition", semantic theories and notions, like possible worlds, quantifiers, operators, domains of quantifications, connectives, and so on, have shaped and still shape the understanding of an astonishing range of philosophical topics, including as diverse issues as ontological realism (Chalmers et al. 2009), the analysis of knowledge and its relation to evidence (Williamson 2000), the metaphysics of modality (Kripke 1972), the nature of belief and desire (Richard 1990), the issue of cognitivism versus non-cognitivism in metaethics (Schroeder 2010), and many more.

One of the basic instruments in a semantic theory is to attribute a certain feature to each linguistic expression which Frege called "sense". *Sense determines reference* (or semantic value in general) at least in that sameness of sense entails sameness of reference; this commitment is crucial, for the whole point of the theory is to identify a feature of expressions that accounts for their semantic values. Though this gives, in itself, very little idea of what "senses" are, it is prima facie plausible to identify sense with the meaning of a linguistic expression. Frege held that shared meaning is possible only if senses are neither mental nor physical, but rather belong to a third realm of beings to which thinking subjects have the same access. The default assumption in the subsequent history of semantics tended to follow this way of thinking in assuming that "propositions" (the senses of declarative sentences) and "concepts" (the senses of sub-sentential expressions, the constituent of propositions) are abstract entities.

Proper names pose a particular problem for the idea that expressions have a sense or a meaning which determines their reference. Saul Kripke's *Naming and Necessity*, published first in 1972, was to become one of the most influential philosophical works of the 20th century, offering one of the best examples of how a focus on the semantic features of language can contribute to a whole range of philosophical problems, including metaphysics and the philosophy of mind. Kripke argued that names contribute their referent to complex expressions directly, without the mediation of senses. Hence what is expressed by names (the object it denotes) is often to be found outside thinking subjects. Kripke argued that similar considerations apply to natural kind terms denoting biological species or chemical kinds, like "gold" or "water" or "tiger".

The sense/reference framework has another important consequence for names, indexicals and natural kind terms. Parallel to Kripke, Hilary Putnam developed in his papers in the seventies a theory of natural kind terms similar to the one presented in *Naming and Necessity*. In his 1975 paper "The Meaning of 'Meaning'" (Putnam 1975a), Putnam invites us to consider the consequences of this theory with the help of a thought experiment. Imagine a planet called Twin Earth which is an exact replica of Earth, with qualitatively identical counterparts of all Earthly inhabitants, including our protagonist, Oscar. The one difference between Earth and Twin Earth is that the liquid they call "water" is in fact a different chemical compound, and hence when Oscar and Twin Oscar talk about "water", they refer to two different kinds.

A similar phenomenon arises in the case of names and indexical expressions. When Oscar and Twin Oscar use the name "Aristotle", they refer to different

individuals: Oscar refers to Aristotle, and Twin Oscar to Twin Aristotle (as we would put it). Other influential cases – based also on expressions other than names and natural kind terms – are discussed in the work of Tyler Burge (Burge 1979). For example, Burge presents two linguistic communities which use the word "arthritis" in slightly different ways, but focuses on two subjects, who have the same views on arthritis. These two subjects, just like Oscar and Twin Oscar, are internally identical, but refer to different things. As we saw, one option in the case of names is to say that they are directly referring expressions, without a sense. But suppose it's implausible to say that some linguistic expressions which contribute to intelligent discourse are devoid of sense or meaning. Sense determines reference: sameness of sense entails sameness of reference, and hence difference of reference entails difference of sense. Therefore, the sense or meaning of "Aristotle" and "water" are different for Oscar and Twin Oscar (and similarly for Burge's protagonists).

Many of our linguistic expressions refer to things outside us. Hence it is obvious that at least some semantic features, namely, the references of many expressions constitutively depend on things outside thinking subjects. This is uncontroversial. If meaning is to determine reference, and two internally identical subjects – like Oscar and Twin Oscar – can refer to different things, then it seems that meanings also have to depend on external factors. It is important to see that this conclusion need not depend essentially on the Kripke/Putnam theory of natural kinds, or on alleged intuitions about the reference of the term "water". As long as we accept that two internally identical subjects can refer to different things by the use of names or indexicals (which is quite difficult to deny; indeed, we can stipulate such use of names), the requirement that sense determines reference entails that these subjects do not share their meanings either, despite their internal sameness. "Meanings ain't in the head", as Putnam famously declared.

2.2 The semantic conception of intentionality

The claim that *meanings* are not in the head is perhaps not that surprising after all. For example, it should be fairly uncontroversial that linguistic meaning depends on things outside an individual, namely on linguistic conventions being accepted by others in her linguistic community. Or it could depend on abstract senses that exist in the third realm beyond mental and physical beings. What about the idea that certain semantic features – other than reference – in my idiolect depend on things outside me? Even that is less than shocking. Semantic theories usually offer models which often include inevitable simplifications – this is especially true of formal semantic theories. All sorts of things can model all sorts of things. We can start with a model where the sense of a sentence is an ordered n-tuple of the senses of its constituents. Then we could be persuaded that names refer directly, without the mediation of a sense, and propose instead that the sense of sentences containing names should be modeled by an n-tuple that includes the reference, rather than the sense of the name. The point I want to stress is that in a way, this move is easy:

n-tuples, whether consisting "senses" or "references" can serve as good models; their usefulness will depend on the explanatory power of the model.

Putnam formulated the original claim about meanings, but the thesis assumed its real significance and became relevant to our current topic when it was broadened to include mental contents. One of the first to make this move was Tyler Burge (Burge 1979).[4] The general argument could proceed as follows. We start with noting that mental states like beliefs, or entertained thoughts, can be true or false. Moreover, not only beliefs, but also desires, perceptual experiences, certain emotions, and other mental states are *about* things in the world: they exhibit intentionality, or the mind's direction upon objects (see Chapter 8 on Intentionality). The next, crucial move is to understand intentional directedness on the analogy of semantic reference. This is not implausible: a belief is about things, which also serve as referents of expressions we use to express the belief. We might think that fundamentally the same idea is involved when a word refers to a thing, or an idea is about the same thing. This move offers the possibility of importing the conceptual tools of semantic theory into understanding the nature of mental states: we attribute a semantic content to a belief which is the same as the semantic content of the sentence we use to express the belief.

The semantic content of sentences, hence of beliefs, is nothing other than the sense or meaning of a sentence. We already know that meaning is outside the head: so the content of beliefs is also outside the head. Similar considerations will apply to other instances of intentional directedness. Hence some mental features are constitutively determined by things outside a thinking subject. This is the view known as *externalism* or *anti-individualism* about mental content, and it entails the rejection of the *internalist* Cartesian conception of the mind sketched in section 1. A solitary brain-in-a-vat or a subject deceived by an evil demon cannot have all the same mental states as we have, if the objects to which we refer don't exist in their world. Eliminating the world outside does not leave the inner world of thought intact, as Descartes believed. If we want thought to be about the world, then even what it is *possible* to think will depend on how things are external to us.

Lines of resistance to the externalist conclusion open accordingly. The requirement that sense determines reference leads to the externalist conclusion only if we assume that sense alone determines reference; for if sense plus something else determines reference, then from different references we cannot infer different senses (Farkas 2008, ch.7). Starting from the 2000s, there has been an intense debate in semantics about the role of context in determining truth-value (see Preyer and Peter 2007 for a representative cross-section of the debate), and a number of theories were developed which give up the principle that sense alone determines reference (MacFarlane 2005). Thus one could retain the semantic conception of intentionality and still resist the externalist conclusion. Alternatively, one could begin to question the idea that intentional directedness should be understood on the model of semantic reference (Crane 2014).

2.3 Two-dimensional views

A semantic theory is a theory of how the semantic value (truth, reference) of expressions is determined. The feature of expressions that determines their reference is variously called "meaning" or "sense" or "content". Some philosophers have suggested that meanings or contents have two different aspects that cannot be explained by a single notion. Just as in the case of semantic externalism, these "two-dimensional semantic theories" or "dual content theories" can be put forward about meanings (and, in this case, also about types of necessity), rather than about the mind. Chalmers (2006) gives a thorough overview of the various versions of this "two-dimensional" approach to contents.

David Kaplan, who is often regarded as the first to present a systematic two-dimensional framework (Kaplan 1977), distinguished between the character and the content of indexical expressions. On this view, the character remains constant for each use of the first person pronoun "I", and it expresses something like "the speaker of this utterance". However, the contents expressed by first-person sentences are different for different speakers: they are singular propositions which constitutively contain the subject of the utterance. Kaplan considers a Putnam-like scenario of Castor and Pollux, identical twins who are stipulated to have qualitatively identical internal states. Kaplan holds that the cognitive or psychological states (he uses the terms interchangeably) of the twins are exactly the same, even though they express different singular propositions when each says "My brother was born before me" (Kaplan 1977, 535). Kaplan follows Putnam's original formulation the lesson of the Twin Earth story: semantics contents do, but psychological states don't determine reference. As mentioned before, this is a type of externalism, but not externalism about the mind.

Just like in the case of the Twin Earth story, the two-dimensional framework was subsequently modified, so that both dimensions were brought into the mental realm. The idea is that internally identical subjects in different environments (like Castor and Pollux, or Oscar and Twin Oscar, or me and my brain-in-a-vat counterpart) are similar in some mental respect, but different in another. For each pair, their mental states share their "narrow" contents, but differ in some of their "broad" contents (Fodor 1987).

The two-dimensional view has been seen by many as a judicious compromise between externalist and internalist views (Chalmers 2002). The broad content of mental states accounts for some of our practices in attributing mental states. For example, if we say that Castor and Pollux believe different things when they each think that "My brother was born before me", it is tempting to say that the difference in beliefs is a mental difference. At the same time, we can see why Kaplan was inclined to say that the twins are psychologically alike. If Castor and Pollux are both convinced that they are second-born, this may prompt similar actions. When we think about how the world appears from the subjective point of view, or how to explain actions in terms of the subject's mental states, it is tempting to discover mental *similarities* among internally identical agents.

262

Two-dimensional views, therefore, do not see broad and narrow features in competition: both contents can be attributed to the same mental state, for different purposes. However, someone could raise the following worry. When I consciously think to myself that "Water is wet" or "My brother was born before me", it does not seem at all that I am entertaining two thoughts on each occasion, with two different contents. With each conscious act of thinking, there seems to be only one thing that I grasp. So it is not very clear how the two different contents are present in my mind. One natural answer on behalf of the two-dimensional theory is to return to the observation, made earlier, that the semantic approach to mental content provides *models*, where operations on a structure of abstract entities model some or other function that contents are supposed to do – determine reference, or explain action, as the case may be. Mental contents can be modeled by sets of possible worlds, centered world, diagonal propositions, and so on, and the theory need not make a claim on the experienced reality of our mental life.

2.4 The extent of inner space

So far we have seen externalist theories which handle the issue of mental content through what we could broadly describe as the "modeling approach". In this section, we shall look at a rather different externalist view, defended in the works of Gareth Evans (Evans 1982) and John McDowell, which addresses the phenomenological-psychological reality of content. I shall use McDowell's 1986 paper "Singular Thought and the Extent of Inner Space" as a representative of this approach, since it explicitly tackles the issue of the boundaries of the mind.

McDowell's target in this paper is what he calls the "fully Cartesian" picture of the mind. On this picture, the mental realm is what is left after entertaining the possibility of radical deception, and consists of transparently accessible and infallibly known facts. Privileged access to the inner realm makes access to the rest of the world correspondingly problematic, opening a gap between the inner and outer world that is very difficult to bridge. The Cartesian conception puts "subjectivity's very possession of an objective environment in question" (McDowell 1986, 237). On McDowell's view, this is not simply the epistemological anxiety about the possible non-existence or radical different nature of the world; rather, the threat is that we cannot explain how thinking about a mind-independent world is so much as possible.

McDowell's answer is based on transforming Bertrand Russell's notion of acquaintance (Russell 1917). Acquaintance is an epistemic-psychological relation in which we stand to objects when we are directly aware of them, without an intermediary process of inference or knowledge of truths. When a particular is an object of acquaintance, it is a constituent of the singular proposition that forms the content of a judgment. McDowell makes it clear that, for Russell, unlike, for example, for Kaplan, singular propositions are not merely part of semantics, but they are intended as a "distinctive kind of configuration in psychological reality" (McDowell 1986, 228). However, for Russell, the psychological reality of

263

singular propositions comes with a serious restriction on the range of objects that can enter into such propositions. Suppose – as McDowell himself would subsequently propose – that we allow external objects to be objects of acquaintance, and constituents of singular propositions. Since we can be mistaken about the presence of an external object, this move would open the possibility of an *illusion* of entertaining a singular proposition. This was unacceptable for Russell.

We can regain possession of the world if we allow that some of our mental states constitutively involve external objects, and thus "we are compelled to picture the inner and outer realms as interpenetrating, not separated from one another by the characteristically Cartesian divide" (McDowell 1986, 241). Some of the object-involving states are subjectively indistinguishable from states which do not involve objects, for example in a veridical perception and the matching perfect hallucination. On the fully Cartesian conception, the mental nature of subjectively indistinguishable states must be the same; otherwise we could mistake one mental state for another. McDowell wants to resist this move by giving up the claim that we are infallible about all aspects of mental states, and he accepts that mental states of very different nature can give rise to the same appearance:

> Short of the fully Cartesian picture, the infallibly knowable fact – its seeming to one that things are thus and so – can be taken disjunctively, as constituted either by the fact that things are manifestly thus and so or by the fact that that merely seems to be the case.
>
> (McDowell 1986, 242)

McDowell, like many others influenced by Oxford philosophy in the second half of the 20th century, can be seen as responding to a question raised prominently in P. F. Strawson's work: how can we explain the very idea of a subject possessing the experience of an objective environment (Strawson 1959, Chapter 2). One area where the question received an especially great amount of attention was perception. It seems a fundamental phenomenological fact about perception that it appears to present a mind-independent world. How is this possible? In the empiricist tradition, it was customary to view perceptual experiences as sensations, that is, as modifications of a subject's consciousness – but this leaves the fundamental phenomenological fact unexplained.

The "disjunctive" theory of perception was developed from the 1980s partly to answer this challenge (see also Chapter 4, Theories of Perception). The basic idea is hinted in the quote from McDowell (McDowell 1986, 242) : when something appears to be the case, it could *either* an object-involving fact manifesting itself in experience, *or* an indistinguishable mere appearance – hence the name "disjunctivism". The fact that some experiences constitutively involve an external object is meant to explain how we can make sense of the idea that perception presents a mind-independent world.[5] The important point for our purposes is that the mental nature of object-involving and non-object involving experiences are radically different, and this difference is due to facts external to the subject: the presence or

absence of an object of perception. Disjunctivism and the relational views about experience are therefore forms of externalism about mental features.

Though Evans and McDowell have been as deeply influenced by the semantic theories of Frege and Russell as many of the externalist philosophers previously mentioned, their approach to the mind is somewhat different from those we placed in the "semantic tradition". It does not seem that they are attempting to offer models for various possible functions of mental features. Instead, they take as a starting point, and as psychologically real, the phenomenologically fundamental features of our thinking and experience: for example, the fact that we seem to be in possession of a conception of an objective world. Perhaps relatedly, McDowell has little interest in a scientistic conception of the mind. On the contrary, one of his main inspirations is Wilfrid Sellars's idea that a proper understanding of the mind is not possible in a merely causal-naturalistic explanatory framework (McDowell 1994). Hence the type of externalism that is motivated by the kind of considerations Evans and McDowell put forward does not fit into either of the broad trends we describe in this chapter: the modeling approaches presented in sections 2.1–2.3, or the naturalistic theories of section 3.

2.5 The boundaries of privileged access

According to externalists, Descartes was wrong in claiming that internally identical subjects always have the same mental features. It's been argued that this entails that Descartes was also wrong in claiming an epistemic privilege to the mental realm (see Brown 2004 for various aspects of the debate). More precisely, compared to internalism, externalism limits – according to this argument – the scope of privileged first-person knowledge of mental features. Indeed, it was claimed (Farkas 2008) that the restriction of privileged access is not a consequence, but rather a defining feature of externalist views. Arguably, this attitude was discernible, for example, in McDowell's considerations quoted in section 2.4: McDowell identifies the privileged epistemic status of mental facts as the central tenet of the fully Cartesian picture, and argues that by giving up this claim, we open the way towards making some of the external world constitutive of our mental states.[6]

It is interesting to mention here another influential view on the limitations of self-knowledge, represented in the work of Sigmund Freud. The two views are of course very different both in their theses and their motivations. But they can be both seen as undermining a central tenet of the Cartesian conception of the mind: in a manner of speaking, while externalists want to extend the boundary outwards, Freud suggested that the boundaries of the mind lie much deeper than the shallow heights reached by straightforward reflective awareness. This is the realm of the unconscious.

We have not addressed Freud in this chapter because Freud and psychoanalysis have had remarkably little effect on mainstream analytic philosophy, compared, for example, to continental philosophy, where Freud had much more of an influence, and compared to the rest of intellectual life and culture.[7] Unconscious

mental states are the subject of much theorizing in analytic philosophy of mind, but the disciplines that analytic philosophers consult about these states are either cognitive psychology (for example, in studying the sub-personal states involved in learning or perceptual processes), or social psychology (for studying phenomena like implicit bias). Nonetheless, the broad influence of Freud's ideas probably contributed to undermining the Cartesian conception which is the main target of externalist views popular in the analytic philosophical tradition and which is also a target of many naturalist views of the mind discussed in the next section.

3. Naturalism about the mind

3.1 The functionalist-computationalist view

In Descartes' time, a philosopher interested in the study of the mind freely included anatomical, psychological, and biological considerations in his works. As with many other disciplines, the various sciences of the mind have subsequently become autonomous and separate from philosophy. Yet developments in the sciences continued to exercise a profound effect on philosophy, and this is especially apparent in analytic philosophy of mind in the 20th century. I have already alluded to the recurrent desire to place philosophy on a secure methodology that would offer a chance of progress comparable to progress in science. There have been two further, distinguishable, though often co-existent, manifestations of this impact.

The first is the doctrine of naturalism or physicalism. In the first half of the 20th century, especially after the development of the theories of relativity and quantum mechanics, it seemed that physics could offer an explanation of the world that was unparalleled in scope, depth, evidence, and explanatory power. The other natural sciences, even if not obviously reducible to physics, have links to physics that promised a unified theory of the world. This inspired a physicalist-naturalist program: aiming at an account of all mental phenomena that is compatible with the thesis that everything that exists figures in the theories of natural sciences. Though materialism was a notable philosophical position already in ancient philosophy, and from the 17th century onwards, it is characteristic of philosophy of mind in the second half of the 20th century that physicalism (both for defenders and opponents) dominated the agenda in a way it never had before (see Chapters 2, 3, 7 in this volume).

The second impact of the sciences is an increased interest in philosophy to approach the study of the mind from an interdisciplinary standpoint. This means actively trying to figure out how empirical results in psychology, evolutionary theory, or neuroscience may affect our philosophical theories of the mind. The two approaches are compatible, but independent: one could believe in a naturalist ontology but still rely mainly on the (broadly speaking) speculative methods of philosophy and not pay much attention to empirical work.[8] Or someone could reject a physicalist or naturalist ontology (together with a somewhat surprising

number of natural scientists),[9] yet invest a lot of time in empirical studies of the mind in the hope that they illuminate our philosophical theories.

Starting in the late 1950s and early 1960s, the 20th century witnessed the development of a robust and sophisticated defense of a naturalist conception of the mind, and the emergence of cognitive science, an interdisciplinary approach to the workings of the mind (see Chapter 11, the Rise of Cognitive Science). One prominent idea both in naturalist theories and in cognitive science is that the mind could be understood on the analogy of computers (inspired by Turing's groundbreaking work for example in Turing 1950).[10] According to the classic functionalist-computational picture of the mind, our core cognitive processes can be understood as programs that manipulate certain representations stored in the brain. The central computing unit is connected to the periphery in two ways: our nervous system transduces stimuli that arrive from the world to our sensory organs, and also conveys certain tasks to be solved (for example by signaling that the body needs nutrition). The information arriving to the central system is used to build an inner representation of the world, which then helps the system to solve the tasks posed for it (for example, by computing the navigations needed to reach food). The solutions are then translated to action commands, which are communicated to another component of the periphery, the motor system (the result being that the organism moves towards the source of food).

One crucial point here is that in this so-called "sense-think-act cycle", the central thinking module, which is sandwiched between the periphery of perception and action, is conceived as running a highly abstract program; that is, a program that can be realized by very different physical mechanisms, and would be, in principle, compatible with a large variety of inputs and outputs, fashioned for all sorts of sensory organs and all sorts of bodies to be moved. This idea actually has very old roots: we saw a similar conception being present already in Aristotle and Descartes, in the view that pure cognition is independent both of the body, and the sensory-affective aspect of our mental life.

Descartes thought that everything material must work on mechanistic principles, and he simply could not imagine how a programmed mechanism could account for the creativity of human thought. Therefore, he held that the immaterial soul must be the home for rational cognition. Descartes' argument is based on some empirical-scientific assumptions that are clearly superseded today. As we shall see in section 3.4, the classic computationalist picture has come under increased criticism. But even if the picture needs correction, we should not underestimate the significance of having a conception of mental processes, the computational-functionalist conception, which makes sense – in a way that's ~~completely~~ consistent with a naturalist world-view – of something that completely baffled Descartes and others for centuries (see Rey 1997, Chapter 2).

There are at least three boundary issues raised by naturalism and the functionalist-computational theory of the mind. First, naturalist accounts of representations usually rely on an external individuation of mental content. Second, functionalism about the mind has the possible consequence that the physical basis of the mind

extends beyond the boundaries of our organic body. Third, the classical computational conception was criticized for not taking into account the essential role of the "periphery" in cognitive processes. The resulting views – externalism about content, the extended mind hypothesis, and various versions of embodied cognition – assert the dependence of the mind on things outside the brain or the subject in rather different ways. I shall explain each of them in turn.

3.2 Naturalist reduction of content

Jerry Fodor has developed one of the most powerful and sophisticated defenses of the computational theory of the mind, through his landmark publications starting in the 1970s (Fodor 1975; 1987). Part of this project is to account for the fact that mental states seem to be about things in the world (see Chapter 8, Intentionality). As Fodor famously put it:

> I suppose sooner or later the physicists will complete the catalogue they've been compiling of the ultimate and irreducible properties of things. When they do, the likes of *spin*, *charm*, and *charge* will perhaps appear on their list. But *aboutness* surely won't: intentionality simply doesn't go that deep.

> (Fodor 1987, 97)

Aboutness in the computational theory of mind is approached through the notion of representation. The basis of the analogy with computer programs is that certain symbols, called "representations" are stored in the brain, and the machine – the mind, the brain – manipulates these symbols on the basis of their physical shapes (that is the only kind of feature detectable by a physical symbol-manipulator). The symbols represent various items in our environment, and this makes it possible for us to think and reason about them. If I believe that cats like cream, and I want give my cat a treat, I will decide to get some cream for him. But what makes it the case that a certain thought, realized by a symbol in the brain, represents cats, rather than, say, dogs? One plausible answer from a physicalist point of view is that there is some sort of causal or nomological connection between the presence of things in the world, and the presence of the representation. The crudest form of this would be the following: seeing a cat causes a certain brain-state which, in virtue of being caused by a cat, represents cats (or this particular cat). This very crude version is unsatisfactory, partly because it cannot account for misrepresentation: if all tokenings represent their causes, no tokening will ever be mistaken. Accounts that try to base representation on some sort of lawful connections or evolved functions face better prospects.

On a nomological account (Fodor 1987), symbols represent things whose presence has a lawful correlation with the symbol. This account can employ a certain version of the two-dimensional framework mentioned in section 2.3. Computations

on symbols in the brain can explain one part of cognition and the narrow aspect of content; the lawful connections with the world can explain the broad aspect.

The Twin Earth scenario illustrates the consequences of the nomological view for the individuation of content. Oscar's "water" representations are nomologically correlated with the presence of H2O, Twin Oscar's "water" representations with the presence of XYZ. The content of their representations are different. The nomological view entails externalism about content as it was defined in section 2. An even more emphatic illustration of this comes from considering a certain version of the brain-in-a-vat scenario. In the virtual reality of brains-in-vats, presumably some complex part of the computer program is responsible for the tokening of representations in the brains, hence these are the prime candidates for standing in a lawful correlation with the presence of representations in the brain. The somewhat surprising conclusion is that brains-in-vats have actually mostly true beliefs about the world: since they represent bits of the computer program, and their world does consist of bits of the computer program, they are mostly right.

Starting in the 1980s, several authors (Dretske 1981; Millikan 1984; Papineau 1987) proposed alternative naturalist-reductive theories which analyze representation in terms of evolved functions. These views also entail externalism: if we imagined an artificial or accidentally created replica of a human being, her representational states would be different, since they would lack an evolutionary history (Davidson 1987). Naturalist theories are therefore usually committed to some form of externalism about mental content.

3.3 Extended mind

Most naturalist-physicalist theories of the mind rely on some version of functionalism, broadly understood. Central to this conception is the idea that what makes something a particular mental state depends on the role it plays in a cognitive system, but not on the physical constitution of the piece of machinery that realizes this role. This multiple realizability is often illustrated by stating that a certain belief or desire, or feeling pain, can be realized by neural states of quite different character in different species of animals, and possibly even by a creature who is made of inorganic material – as long as there is an isomorphism in the functional role the state plays in a larger system (Putnam 1960).

This raises the question of how much liberty we can take with the realization of functional roles. Andy Clark and David Chalmers proposed the following thought experiment (Clark and Chalmers 1998). Inga is an ordinary person who lives in New York and wants to visit the Museum of Modern Art. She recalls that the museum is on 53rd Street and sets off. Contrast her case with that of Otto, who suffers from long-term memory loss, and therefore enters all important information into a notebook that he carries with him and consults all the time. When Otto wants to visit the Museum of Modern Art, he looks up the address in his notebook, and sets off.

269

It is usually agreed that Inga has the belief that the museum is on 53rd Street even prior to recalling this information. On a functionalist theory, this state is defined in terms of having a certain functional role in Inga's cognitive system: taking into account her other mental states (for example a desire to visit a museum), it responds to certain inputs with certain outputs (for example setting off towards the museum). Clark and Chalmers argue that the information stored in Otto's notebook has exactly the same role in Otto's cognitive system, and therefore we should attribute the same belief to Otto, even before he consults his notebook. The fact that the notebook is to be found outside Otto's organic body is irrelevant here. Clark and Chalmers offer the following general consideration to support their claim:

> If, as we confront some task, a part of the world functions as a process which, *were it done in the head*, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world is (so we claim) part of the cognitive process.
>
> (Clark and Chalmers 1998, 8)

The principle has subsequently become known as "the Parity Principle", and it is a consequence of a functionalist view of mental states. It is instructive to compare Otto and Inga with Ned Block's famous example in which he imagines the population of China realizing a system that is functionally equivalent to Block's brain: such a system would arguably lack mental states altogether, hence, Block argued, there has to be more to mentality than playing a functional role (Block 1980). But, unlike the Chinese network, both Inga and Otto are *bona fide* cognitive agents with mental states, so it makes perfectly good sense to ask whether they possess a particular mental state (ie. the belief that the museum is on 53rd Street) or not.

If functionalism is correct, this translates to the question of whether any state plays the appropriate role in their cognitive system. Functionalism asks us to disregard physical realization, for example the difference between brain tissue and pages of a notebook. Moreover, as Mark Sprevak has convincingly argued (Sprevak 2009), the spirit of functionalism also asks us to disregard the micro-functional differences that undoubtedly exist between Inga's and Otto's access to the relevant information. The macro-functional roles of the information stored in Inga's relevant brain-state and Otto's notebook are arguably the same. Therefore it seems we have to conclude that Otto also has the belief that the museum is on 53rd Street. The argument extends to any type of mental state which can be plausibly accounted for in terms of functional roles: for example, standing states like intentions or desires. It is less obvious whether, or how, the argument extends to episodes in the stream of consciousness whose identity arguably depends on their conscious or phenomenal character. Both Clark and Chalmers are inclined to think that consciousness does not extend in the way standing mental states do (Chalmers 2008; Clark 2010).[11]

The term "extended mind" suggests that the notebook is actually part of Otto's mind, or, put in more functionalist terms, the notebook is part of the physical

reality that serves as a realizer for the functional states we attribute to Otto. This seems to be a dependence on things outside the subject that is different from the dependence in externalist theories discussed in section 2. We will return to this question in section 4.1.

3.4 Embodied cognition

The topic of "embodied" or "situated" cognition has received a lot of attention in cognitive science starting in the 1990s. The debates we have discussed so far have, by and large, been directly motivated by questions arising within philosophy: about the nature of intentionality, the proper account of representation, the philosophical theory of perception, the mind-body problem. In contrast, much of the historical and current background for discussing embodied cognition is found outside the usual disciplinary boundaries of philosophy. Part of the task of philosophers in this debate is, and has been, to distill the philosophical significance of certain ideas coming from the more empirical disciplines.

Robert Wilson and Lucia Foglia (2011) identify several key historical influences on the formation of embodied cognitive science, including three books that they regard as the first landmark publications in the area. In *Metaphors We Live By*, George Lakoff and Mark Johnson argued that many central cognitive processes are influenced by the use of metaphors, which, in turn, are deeply influenced by the kind of bodies human beings have and use in interacting with the world (Lakoff and Johnson 1980). Francisco Varela, in his book *The Embodied Mind*, co-authored with Evan Thompson and Eleanor Rosch, proposed an "enactive" program for cognitive science that questions "the assumption . . . that cognition consists of the representation of a world that is independent of our perceptual and cognitive capacities by a cognitive system that exists independent of the world" (Varela, Thompson and Rosch 1991, xx) Instead, cognitive structures emerge from the organism's interaction with the world: from the way sensory stimulation systematically changes as organisms and objects move (called "sensorimotor patterns" or "sensorimotor dependencies"). By using the term "embodied", the authors aim to highlight "first, that cognition depends upon the kinds of experience that come from having a body with various sensorimotor capacities, and second, that these individual sensorimotor capacities are themselves embedded in a more encompassing biological, psychological, and cultural context" (Varela et al. 1991, 173–174). As the quote also makes it clear, the scope of this thesis – unlike the extended mind thesis of the precious section – is meant to include conscious experiences, as well as standing states like beliefs.[12]

Continuing the list of key influences on embodied cognitive science, in robotics, Rodney Brooks (Brooks 1991a; 1991b) reported successful work in building robots based on a view of computational intelligence which did not rely on "independent information processing units which must interface with each other via representations. Instead, the intelligent system is decomposed into independent and parallel activity producers which all interface directly to the world

271

through perception and action" (Brooks 1991a, 139). Brooks's insights feature prominently in Andy Clark's 1997 *Being There: Putting Brain, Body, and World Together Again*. Further historical influences mentioned by Wilson and Foglia include James Gibson's ecological view of perception (Gibson 1979), work on dynamical systems, and the dynamicist theory of cognitive development (Thelen and Smith 1994). Finally, a somewhat unexpected source of inspiration is to be found in the phenomenological philosophical tradition of the first part of the 20th century; in particular in the works of Husserl, Sartre, Heidegger, and Merleau-Ponty (see Chapter 1, The Phenomenological Tradition).

As these summaries of the origins of the embodied conception already indicate, the ideas coming together under the label of "embodied cognition" are rather diverse, and there isn't one single and specific thesis that represents the movement. We will now touch upon some of its central themes.

One dominant motif is a phenomenological reflection on mind and cognition, prominent in the works of philosophers like Merleau-Ponty and Heidegger. Phenomenologists point out that our view of the world is formed in close, and often unreflective, active interactions with things in the world, and those interactions, in turn, are essentially shaped by the nature of our body and sensory system. The resulting view is often styled as "anti-Cartesian", but it is instructive to see how it relates exactly to Descartes' actual views. Undoubtedly, this picture has a different flavor than the highly reflexive, contemplative, and abstract viewpoint that is often felt to be advocated by Descartes. There are many more, often metaphorical statements that aim to illustrate this contrast between "involvement" and "detachment", for example in the claim that once we get phenomenology right, we realize "that to perceive is to be in an interactive relationship with the world, not to be in an internal state that happens to be caused by the external world" (Cosmelli and Thompson 2011, 165). Arguably, many of these contrasts reveal an important difference in emphasis, but they hardly amount to serious doctrinal differences. After all, an interactive relationship is perfectly compatible with the idea that we have internal states caused by the world.

Second, though Descartes believed that the mind is an immaterial substance separable from the body, he fully acknowledged that as far as *actual* human psychology goes, sense-experience and the involvement of the body are crucial elements. "I am not merely present in my body as a sailor is present in a ship" (Descartes 1984, vol.2, 81), Descartes famously remarked.

As I mentioned in section 1, Descartes also had a partly empirical hypothesis about the role of the body in different mental processes. He thought that while brain states serve as proximal causes for sensory and affective mental states, the brain is not involved in pure thinking; instead, thinking takes place in the immaterial soul. Very few people today accept the existence of Cartesian immaterial substances, but it has been argued that the sharp Cartesian divide between pure thinking and the rest of mental phenomena survives in functionalist theories. As described in section 3.1, these theories hold that paradigmatic forms of cognition consist of centrally executed operations on highly abstract representations,

sandwiched, but only contingently, between the input of a sensory system and the output to a motor system. A number of findings about the nature of cognition have been thought to undermine the classical computational picture.

Margaret Wilson (2002) provides a useful overview of some of these findings. First, cognition is situated: we often carry out cognitive tasks while perceptual information keeps coming in, and action is performed with a possible impact on the environment, both of which continuously modify the task at hand. Second, cognition is often under time-pressure that prevents building an abstract internal model of the world. Third, we off-load cognitive work on our body and on the environment: rotating the shapes in Tetris on the screen rather than in our head, we share the burden of holding and manipulating information with our environment. Fourth, much of cognition is directly for action, rather than for representation. All these ideas put considerable pressure on the claim that cognition is realized exclusively or even mainly by the kind of processes that figure in the classical computational view.

4. Varieties of "externalism"

4.1 Different conceptions of boundaries

The views presented so far all concern the boundaries of the mind in some sense. In this section, I will briefly compare these different senses. Susan Hurley distinguished between "what" and "how" externalism (Hurley 2010). "What" externalist theories (which include the theories we discussed in section 2) claim that some personal-level mental features – for example, content or phenomenal quality – are explained by external factors. Extended and embodied views belong to "how" externalism, on which external features "explain how the processes or mechanisms work that enable mental states" of specific types (Hurley 2010, 101). In earlier work (Hurley 1998), Hurley called how-externalism "vehicle" externalism, because it involves claims about the realizers or vehicles of mental states. In contrast, externalists of section 2 hold that we individuate certain mental states partly by their relations to their intentional objects – but this view need not say anything about vehicles. (I have refrained from calling embodied and embedded views "externalist", reserving this term only for what Hurley calls "what-externalism".)

Further differences between the come to light if we consider the brain-in-a-vat scenario. In his review of Alva Noë's book *Action in Perception*, Ned Block illustrates his disagreement with Noë's enactive approach (which is a type of embodied view, on our classification) by speculating about a solitary brain-in-a-vat who, through an unlikely but not impossible chance fluctuation of particles, comes to existence in exactly the same physical state as the brain of an embodied human being (Block 2005). Block takes Noë to hold that the brain would not have the same experience as its embodied counterpart, and he argues that this is is implausible, because the body and the environment are merely causes of the experience, but not constitutive parts of the realizer of the experience.[13]

Block claims supervenience on the brain specifically for *experiences*, but not for all mental states. Though he is an internalist about experience, like the majority of philosophers of mind, he is an externalist about mental contents (the issue of extended mind is not addressed in the review). This way of presenting the matter suggests that all the different debates discussed in this chapter can be formulated in terms of which mental features do or do not supervene the brain or the body. However, Evan Thompson and Diego Cosmelli (2011) argued that setting up the opposition this way doesn't get to the heart of the enactive-embodied approach. Their interest is not the purely philosophical question about "the minimal metaphysical supervenience base" of experiences, but rather an explanatory framework for interdisciplinary research – for example, for research in neuroscience on the neural correlates of consciousness. On this latter approach, an interesting question is the bioengineering task of keeping a brain alive and functioning in a vat, and of providing stimuli that match our environment. Thompson and Cosmelli investigate this question in some detail, and find that the task is absolutely formidable, and the only way it could be done is to build something like a body for the brain and place it in an appropriate environment. They conclude:

> In the range of possible situations relevant to the explanatory framework of the neuroscience of consciousness, the brain in a vat thought experiment, strictly speaking, doesn't seem possible (because the envatted brain turns out to be an embodied brain after all).
>
> (Cosmelli and Thompson 2011, 173)

The extended mind hypothesis (Section 3.3) is usually classified together with, or even as one of the possible embodied views (Section 3.4), because of the apparent shared interest in the realization of cognitive processes (in other words, because of answering a how, rather than a what-question, to use Hurley's terminology). In fact, the motivations of the two views are rather different. The extended mind hypothesis, as explained above, is a consequence of the functionalist view that only functional roles matter and the nature of the physical realizer don't. Clearly, Otto's notebook could be replaced by a computer, by a tape recorder, by any kind of device that was capable of holding the abstract representations stored in Otto's notebook. This is quite alien to the spirit of embodied views, which emphasize the dependence of cognition on the particular shape of our bodies and the on the contingent variation of sensory stimulation with our interactions with the world. The two views are not incompatible, but, arguably, they limit each other's scope. States that are especially suitable for extension tend to employ multiply realizable, abstract representations – so these states are not strongly embodied. In contrast, embodied processes that depend on a contingent bodily setup are likely to resist extension (see also Clark 2008, part III).

4.2 Current state

Externalism, as defined in section 2, has become something like the orthodoxy in contemporary analytic philosophy of mind. Most philosophers who write on relevant topics are externalist at least about some mental features; that is, they hold that these features depend on factors external to a thinking subject.[14] Many are externalists about (at least some aspects of) content. Defenders of disjunctivist and relational views of perception are externalists about experiences. A further proposal which generated significant interest and debate was Timothy Williamson's claim that knowing is a state of mind (Williamson 2000). Since on this view, an ordinary knowing subject and her ignorant brain-in-a-vat counterpart have different mental states, this is also a form of externalism. In recent years, there has been some revival of an internalist defense (Farkas 2008; Mendola 2008), but this position remains in the minority.

The extended mind hypothesis, as discussed in section 3.3, remains a controversial thesis in philosophy, with committed defenders, committed opponents, and many agnostics. The findings described in section 3.4 give strong support to the claim cognitive processes significantly involve sensory input, motor output and interaction with the environment, so embodied cognitive science remains a robust research program. As we have seen in section 4.1 above, the focus of this program is often on empirical questions that go outside the usual disciplinary bounds of philosophy. But while the results strongly suggest that not all cognition is offline, abstract and representational, as Margaret Wilson (2002) notes, the same phenomena do not show that *no* cognition is performed in such a way. In his review of Varela, Thompson, and Rosch's *Embodied Mind* Daniel Dennett asked whether the enactive program was really revolutionary or rather a welcome shift in emphasis (Dennett 1993, 122). Dennett thought it was too soon to answer the question in 1993, and it is not obvious that the matter has been settled since then.

Notes

1 I would like to thank Tim Crane, Amy Kind, Philip Walsh, and Jeff Yoshimi for valuable comments on an earlier draft. Research for this chapter was supported by the Hungarian Scientific Research Fund, grant no. OTKA K-112542.
2 Descartes' *Meditations* are included in the second volume of his selected philosophical writings (Descartes 1984).
3 Frege's most important writings, including "On Sinn and Bedeutung" (1892) and "The Thought" (1918) are collected in Beaney 1997.
4 Putnam notes in a foreword to Pessin and Goldberg 1996 that at the time of writing "The Meaning of "Meaning"", he was not sure what the Twin Earth story entailed with respect to the mind (as opposed to meanings). But subsequently, he was persuaded by Burge and John McDowell that mental states also depend on factors outside us.
5 For defense of the disjunctive theory of perception, see McDowell 1982, Martin 2004. John Campbell is not a disjunctivist, but he also defends a form of externalism about perceptual experiences as a response to the question of how we can possess an objective environment; see his contribution in Campbell and Cassam 2014.

6 Brie Gertler (2012) assesses different detailed definitions of externalism, and concludes that there is no univocal thesis of externalism and internalism.

7 Here is an illustration: at the time of writing this chapter in October 2014, there are around 1,500 entries in the *Stanford Encyclopedia of Philosophy*, which is the most widely consulted internet reference work in analytic philosophy. Ninety-two of these entries (mostly on continental or feminist philosophy) refer to Freud, whereas, for example, Hilary Putnam is mentioned in 230 entries. The SEP entry on the "Philosophy of Psychiatry" does not contain a single reference to Freud's work. The term "psychoanalysis" is mentioned in 76 documents. In contrast, the number of documents that mention "cognitive science" is 175, "artificial intelligence" 119, "quantum mechanics" 138.

8 This would be true for example of U. T. Place, J.J.C. Smart, and David Armstrong, who published influential work defending physicalism in the 1950s and 60s; see Place 1956, Smart 1953 and Armstrong 1968.

9 I don't have statistics on the philosophical views of scientists, but it is interesting to note that most Nobel Prize winners in the 20th century who did research on the brain expressed some view in writing on the mind body-problem, and, with one exception, they were not physicalists. Charles Scott Sherrington (Nobel Prize 1932) held a "double aspect" theory; John Eccles (Nobel Prize 1963) was a dualist, Gerald Edelman (Nobel Prize 1972) defended non-reductive biologism, and Roger Wolcott Sperry (Nobel Prize 1981) defended a type of emergentism. The exception is Francis Crick (Nobel Prize 1962), who was an ardent physicalist.

10 Classic defenses of an early version of functionalism can be found in Putnam 1960 and 1967.

11 For various pros and cons in the debate, see the papers collected in Menary 2010.

12 For another development of the enactive conception, see Noë 2005.

13 For this type of criticism of the extended mind view, see also Adams and Aizawa (2010).

14 According to the 2009 survey conducted by PhilPapers, 51 percent of respondents accept or lean toward externalism, 20 percent accept or lean towards internalism, and 29 percent indicated "Other". However, we should note that many philosophers who accept dual content theories call themselves internalists, because they recognize some sort of narrow content in the contested cases. This hides the fact that they accept externalism about *some* mental features. For example, David Chalmers and Terry Horgan, philosophers who argued for a robust notion of narrow content, both claim they accept internalism, even though they both think that there is also an aspect of mental content which is broad. Horgan makes this clear in a comment: "I hold that the most fundamental kind of mental content is internalist (and phenomenally constituted), but that some thought-constituents also have a form of intentionality that constitutively depends in part on internal/external linkages" (PhilPapers Survey).

Bibliography

Adams, Fred and Ken Aizawa (2010). "Defending the Bounds of Cognition," in Menary (ed.): 67–80.

Armstrong, D. M. (1968). *A Materialist Theory of the Mind*. London: Routledge & Kegan Paul.

Beaney, Michael (ed.). (1997). *The Frege Reader*. Oxford: Blackwell Publishing.

Block, Ned (1980). "Troubles with Functionalism," in Block, Ned (ed.) *Readings in Philosophy of Psychology*. Cambridge, MA: Harvard University Press: 268–305.

Block, Ned (2005). "Review of *Action in Perception* by Alva Noë," *The Journal of Philosophy*, 102/5: 259–272.

Brown, Jessica (2004). *Anti-individualism and Knowledge*. Cambridge, MA: MIT Press.

Brooks, Rodney (1991a). "Intelligence without Representation," *Artificial Intelligence*, 47: 139–159.

Brooks, Rodney (1991b). "Intelligence without Reason," *Proceedings of 12th International Joint Conference on Artificial Intelligence*, Sydney, Australia: 569–595.

Burge, Tyler (1979). "Individualism and the Mental". Reprinted in Foundations of Mind: Philosophical Essays. Oxford: Clarendon Press, 2007: 100–150.

Campbell, John and Cassam, Quassim (2014). *Berkeley's Puzzle*. Oxford: Oxford University Press.

Chalmers, David (2002). "The Components of Content", in Chalmers, David (ed.), *Philosophy of Mind: Classical and Contemporary Readings*. Oxford: Oxford University Press: 608–633.

Chalmers, David (2006). "The Foundations of Two-dimensional Semantics", in Carpintero, M. Garcia, and Macia, J. (eds.), *Two-dimensional Semantics: Foundations and Applications*. Oxford: Oxford University Press: 55–140.

Chalmers, David (2008). "Foreword," in Clark (ed.): ix– xxix.

Chalmers, David, Manley, David and Wasserman, Ryan (eds.). (2009). *Metametaphysics*. Oxford: Clarendon Press.Clark, Andy (1997). *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT Press.

Clark, Andy (2008). *Supersizing the Mind: Embodiment, Action and Cognitive Extension*. Oxford: Oxford University Press.

Clark, Andy (2010). "Extended Mind Redux: A Response," http://opinionator.blogs.nytimes.com/2010/12/14/extended-mind-redux-a-response/.

Clark, Andy and Chalmers, David (1998). "The Extended Mind," *Analysis*, 58: 7–19.

Cosmelli, Diego and Thompson, Evan (2011). "Brain in a Vat or Body in a World: Brainbound versus Enactive Views of Experience," *Philosophical Topics,* 39: 163–180.

Crane, Tim (2014). *Aspects of Psychologism*. Cambridge, MA: Harvard University Press.

Davidson, Donald (1987). "Knowing One's Own Mind", in *Proceedings and Addresses of the American Philosophical Association*, 60: 441–458.

Dennett, Daniel (1993). "Review of *The Embodied Mind: Cognitive Science and Human Experience* by Francisco J. Varela; Evan Thompson and Eleanor Rosch," *The American Journal of Psychology*, 106/1: 121–126.

Descartes, René (1984). *Philosophical Writings of René Descartes*. 3 volumes. Edited and translated by J. Cottingham, R. Stoothof, D. Murdoch and A. Kenny. Cambridge: Cambridge University Press.

Dretske, Fred (1981). *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press.

Dreyfus, Hubert L. (1991). *Being in the World: Commentary on Heidegger's "Being and Time", Division 1*. Cambridge, MA: MIT Press.

Evans, Gareth (1982). *The Varieties of Reference*. Edited by John McDowell, Oxford: Clarendon Press.

Farkas, Katalin (2005). "The unity of Descartes's Thought," *History of Philosophy Quarterly*, 22/1: 17–30.

Farkas, Katalin (2008). *The Subject's Point of View*. Oxford: Oxford University Press.

Fodor, Jerry (1975). *The Language of Thought*. New York: Thomas Crowell.

Fodor, Jerry (1987). *Psychosemantics*. Cambridge, MA: Bradford Books.

Gertler, Brie (2012). "Understanding the Internalism-Externalism Debate: What is the Boundary of the Thinker?" *Philosophical Perspectives: Philosophy of Mind:* 51–75.

Gibson, J.J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

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

Hurley, Susan (2010). "The Varieties of Externalism," in Menary (ed.): 101–153.

Kaplan, David (1977). "Demonstratives," in Almog, J., Perry, J. and Wettstein, H. (eds.), *Themes from Kaplan*. Oxford: Oxford University Press, 1989: 481–563.

Kripke, Saul (1972). "Naming and Necessity," in Davidson, D. and Harman, G. (eds.), *Semantics of Natural Language*, 2nd edition. Dordrecht: Reidel Publishing Co.: 253–355.

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

Lau, Joe and Deutsch, Max (2014). "Externalism About Mental Content," *The Stanford Encyclopedia of Philosophy* (Summer 2014 Edition), Zalta, Edward N. (ed.). http://plato.stanford.edu/archives/sum2014/entries/content-externalism/.

MacFarlane, John (2005). "Making Sense of Relative Truth," *Proceedings of the Aristotelian Society*, 105: 321–339.

Martin, M.G.F. (2004). "The Limits of Self-Awareness," *Philosophical Studies*, 120: 37–89.

McDowell, John (1982). "Criteria, Defeasibility and Knowledge," Reprinted in McDowell 1998: 369–394.

McDowell, John (1986). "Singular Thought and the Extent of Inner Space," Reprinted in McDowell 1998: 228–259.

McDowell, John (1994). *Mind and World*. Cambridge, MA: Harvard University Press.

McDowell, John (1998). *Meaning, Knowledge and Reality*. Cambridge, MA: Harvard University Press.

Menary, Richard (ed.). (2010). *The Extended Mind*. Cambridge, MA: MIT Press.

Mendola, Joseph (2008). *Anti-Externalism*. Oxford: Oxford University Press.

Millikan, Ruth (1984). *Language, Thought and Other Biological Categories*. Cambridge, MA: MIT Press.

Noë, Alva (2005). *Action* in *Perception*. Cambridge, MA: MIT Press.

Papineau, David (1987). *Reality and Representation*, Oxford: Basil Blackwell.

Pessin, Andrew and Goldberg, Sanford (eds.). (1996). *The Twin Earth Chronicles*. Armonk, NY and London: M. E. Sharpe.

PhilPapers Survey (2009). http://philpapers.org/surveys/.

Place, U.T (1956). "Is Consciousness a Brain Process?" *British Journal of Psychology*, 47: 44–50.

Preyer, Gerhard and Peter, Georg (eds.). (2007). *Context-Sensitivity and Semantic Minimalism*. Oxford: Oxford University Press.

Putnam, Hilary (1960). "Minds and Machines," Reprinted in Putnam 1975b, 362–385.

Putnam, Hilary (1967). "The Nature of Mental States," Reprinted in Putnam 1975b, 429–440.

Putnam, Hilary (1975a). "The Meaning of 'Meaning,'" Reprinted in Putnam 1975b, 215–271.

Putnam, Hilary (1975b). *Mind, Language, and Reality*, Cambridge: Cambridge University Press.

Richard, Mark (1990). *Propositional Attitudes: An Essay on Thoughts and How We Ascribe Them*. Cambridge: Cambridge University Press.

Rey, G. (1997). *Contemporary Philosophy of Mind: A Contentiously Classical Approach*. Oxford: Basil Blackwell.

Russell, Bertrand (1917). "Knowledge by Acquaintance and Knowledge by Description," *Proceedings of the Aristotelian Society*, 1910–1911: 108–128.

Schroeder, Mark (2010). *Being For: Evaluating the Semantic Program of Expressivism*. Oxford: Oxford University Press.

Smart, J.J.C. (1953). "Sensations and Brain Processes," *Philosophical Review*, 68: 141–156.

Sprevak, Mark (2009). "Extended Cognition and Functionalism," *Journal of Philosophy*, 106(9): 503–527.

Strawson, P. F. (1959). *Individuals*. London: Methuen.

Thelen, E., and Smith, L. B. (1994). *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, MA: MIT Press.

Turing, Alan (1950). "Computing Machinery and Intelligence," *Mind*, 59: 433–460.

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

Williamson, Timothy (2000). *Knowledge and Its Limits*. Oxford: Oxford University Press.

Williamson, Timothy (2006). "Must Do Better," in Greenough, P. and Lynch, M. (eds.), *Truth and Realism*. Oxford: Oxford University Press: 177–187.

Wilson, Margaret (2002). "Six Views of Embodied Cognition," *Psychonomic Bulletin and Review*, 9: 625–636.

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