Non-reductivism and the Metaphilosophy of Mind

GIUSEPPINA D’ORO
Keele University

PAUL GILADI
University College Dublin

ALEXIS PAPAZOGLOU
Royal Holloway

Introduction

While in the latter half of the twentieth century non-reductivism has enjoyed greater popularity than either reductivism or eliminativism, most forms of non-reductivism in analytic philosophy of mind have been articulated from the same naturalist platform as reductivism and eliminativism. In what follows, we will consider some of the metaphilosophical assumptions that have dominated analytic philosophy of mind, and how they gave rise to the central question that the best-known forms of non-reductivism available have sought to answer, namely: how can mind fit within nature? As we shall see, some of the key problems which have preoccupied non-reductivists, such as those of epiphenomenalism, explanatory exclusion and causal overdetermination, arise largely because of an implicit commitment to a cluster of metaphilosophical assumptions that have not been relinquished as the reductivist orthodoxy which dominated the early days of analytic philosophy of mind gave way to a new non-reductivist consensus. Sometimes philosophical progress is made not when new answers to the same questions are found, but when different questions are asked, questions which rely on a different set of metaphilosophical assumptions (Guess 2017). It is our goal to introduce forms of non-reductivism that have questioned the fruitfulness of the question “how can mind fit within nature?” and which have taken a different approach to the so-called “placement” problem: rather than trying to solve it, by seeking to show how mind can fit into nature within the constraints of a metaphilosophical picture inherited from reductivism, they have put pressure on the philosophical assumptions that have given rise to the question of the place of mind in nature.

We will begin by considering forms of non-reductivism that operate in the same philosophical paradigm as reductivism and examine how they endeavour to address the problems of epiphenomenalism, causal overdetermination and explanatory exclusion. We shall argue that the difficulties encountered by non-reductivist positions such as token physicalism, explanatory gap theory and multiple realization functionalism are the inevitable consequence of the metaphilosophical paradigm they have inherited from their reductivist predecessors. Having discussed the background
metaphilosophical assumptions that give rise to the question “how can mind fit within nature?”, and the difficulties encountered by forms of non-reductivism which seek to accommodate the mind in the natural world, we will turn to consider forms of non-reductivism which question the very idea that the task of the philosophy of mind is to solve the location problem. The view that the location problem is the central problem of the philosophy of mind has been questioned in (at least) two different ways. On the one hand the problem has been dismissed as arising from a mistaken conception of the relation between method and metaphysics. If ontological questions are internal to methodological ones, then the subject matter of a form of inquiry is determined by its method, a method that will have been devised to answer the questions characteristic of that form of inquiry. Further, if there are no external ontological questions, then it would make no sense to ask how can mind fit within nature, since, on this view, nature is the explanandum of natural science, not an inquiry-independent reality into which mind must somehow fit. This way of dissolving the placement problem ultimately entails that the reality investigated by science is the correlative of a certain form of inference, which is characteristically deemed to be nomological, and that it is a mistake to assume that science uncovers inquiry-independent truths. This form of non-reductivism is to be found in the neo-Kantian tradition of Windelband and Rickert, the hermeneutic tradition of Dilthey and Gadamer, and is informed by a Kantian rejection of pre-critical, dogmatic metaphysics. This view of the relation between method and metaphysics also underpins the methodologically motivated non-reductivism that was invoked to defend the autonomy of a humanistic historiography and social science in the 1960s and reached mainstream analytic philosophy through the work of W.H. Dray (1957, 1958, 1963, 1980) and Peter Winch (1958). These methodologically-motivated forms of non-reductivism are articulated in different philosophical traditions and cut across the continental/analytic divide, but they share the view that method determines subject matter and that ontological questions are internal to forms of inquiry. The view that the location problem is the central problem of the philosophy of mind has also, and more recently, been questioned within a different philosophical framework, that of liberal naturalism, by philosophers such as McDowell, who has sought to undermine the location problem in a slightly different way. Rather than reversing the order of priority between ontology and epistemology, McDowell has argued for a liberalization of the ontological picture of nature operative within the reductivist paradigm. What needs to be done, on this view, is to reject the mechanistic picture of reality that is inherited from science. Once this is done the location problem subsides precisely because it thrives upon a barren and disenchanted conception of nature. Although these strategies are very different, they share the view that a more promising form of non-reductivism must challenge some deep metaphilosophical assumptions that have been left untouched as the reductivist consensus gave way to a new non-reductivist orthodoxy. Having considered forms of non-reductivism that challenge the metaphilosophical assumptions that give rise to the location problem, we will ask whether attempts to defend the autonomy of the mental that have relinquished the metaphilosophical commitments which give rise to the question “how does mind fit into nature?” fare any better. While an editorial introduction would not be the place to answer
this question fully, we hope that outlining different ways of defending the autonomy of the mental, and connecting them to the usual non-reductivist solutions, will succeed in broadening the items on the agenda of the philosophy of mind to make room for philosophical positions which do not share the same metaphilosophical platform as standard forms of non-reductivism.

Non-reductivism: the usual suspects

There was a time, in the early days of analytic philosophy of mind, when the consensus was generally reductivist. The mind-brain identity theory, advocated by Place (1956) Smart (1959) and Feigl (1958), had aimed to solve the problem of the causal efficacy of the mental (how can mind have causal efficacy?) by arguing that mental states are nomologically reducible to physical states. If mental states are reducible to physical states, so the argument went, then mental states are not “causal danglers” (the term used to denote properties with no causal powers) since physical states are (paradigmatically) causally efficacious. The days since this reductivist consensus reigned have long past and the agenda of much contemporary analytic philosophy of mind has been not to reduce the mental to the physical but rather to argue against the possibility of any such reduction. Yet while the consensus may have changed, the underpinning metaphilosophy has lived on, and the moves available to the non-reductivist have tended to be constrained by one or more of the following cluster of assumptions that have gone largely unchallenged:

a) that to explain is causally to explain;

b) that real causal relations are captured by explanations at the basic level of physics, which for this reason are seen to be ontologically basic;

c) that the explanations at the basic level of physics are complete.

Given these background assumptions, the logical distance separating non-reductivism from reductivism in the philosophy of mind may be justifiably compared to that separating compatibilism from determinism in the free will debate: compatibilists do not reject determinism; they seek rather to defend the possibility of freedom within a causal/deterministic framework. Likewise, non-reductivists have sought to defend the autonomy of the mental while accepting one or more of the three assumptions listed above, assumptions that they have inherited from the reductivist paradigm. These assumptions determine the questions that non-reductivists address, the kind of moves they see themselves able to make, and ultimately the problems which they think need solving. The best-known forms of non-reductivism, for example, tend to wrestle with the twin problems of epiphenomenalism and explanatory exclusion. Such problems arise because if it is assumed that the explanations of physics are complete and capture real causal relations, then explanations in the special sciences cannot have the status of genuine causal laws. For if they did, they would compete with physical explanations. Let us consider
then how the most widespread forms of non-reductivism have navigated the assumptions that they have inherited from the reductivist paradigm.

Non-reductive physicalism sought to retain the main advantage of reductive physicalism (its ability to explain how mind can have causal efficacy) by arguing that mental properties are identical with physical properties whilst deeming impossible the kind of nomological mapping of the mental onto the physical advocated by reductive physicalists. This, non-reductive physicalists believed, enabled them to maintain both that mental properties are causally efficacious in virtue of being token identical with physical properties, and to deny reduction. Yet, since its introduction, non-reductive physicalism has come under attack either for reintroducing the problem of causal danglers, which reductive physicalism claimed to have solved, or for avoiding this problem only at the price of facing another potentially more serious threat, that of explanatory exclusion.¹ For, so the objection goes, if mental states are token identical with physical states, and physical states are causally efficacious, then mental states are causally efficacious qua physical states. Mental states are therefore epiphenomenal qua mental; if they were not epiphenomenal qua mental, then psychological explanations would compete with those at the level of physics. One can see how the three assumptions mentioned above constrain the options available to the non-reductive physicalist. Given that non-reductive physicalists accept a) that to explain is causally to explain, b) that the explanations at the level of physics capture ontologically real relations and c) that explanations at the level of physics are complete explanations, the problem of explanatory exclusion can be avoided only at the price of accepting that the explanations at the higher levels do not interfere with those at the lower level, i.e. by denying that mental states enter into causal relations. This is not to say that non-reductive physicalists explicitly endorse epiphenomenalism with regard to the mental, but that epiphenomenalism is an unwanted consequence of non-reductive physicalism.

Some non-reductivists have been concerned less with providing an answer to the question “how can mind have causal efficacy?” and more preoccupied by the consideration that the phenomenon of consciousness is recalcitrant to physical explanation. How could something like the sharp taste of a lemon or the beauty of a sunset be explained in physical terms? Different conclusions, either of an epistemic or of an ontological nature, have been drawn from the consideration that so-called qualia resist physical description. Explanatory gap theorists such as Nagel (1979), Levine (1983) and McGinn (1997) have argued that the phenomenon of consciousness is impervious to physical explanation, not because the mind is not physical and amenable to causal explanation, but because there is a distinctively first-person mode of access to mental states which makes us cognitively closed or unable to understand how physicalism could be true. Constrained in this way, explanatory gap theory is strictly speaking not an argument against the reductivist claim that physics can provide complete causal explanations of what there is, including the mind. It is aimed rather at accounting for why the mind/body problem persists despite the progress of natural science by invoking our inability to comprehend how physicalism could

¹ This problem has been well documented by J. Kim (1989; 1990; 1995) and E. Sosa (1984 and 1993).
be true, much perhaps in the way in which Leibniz appealed to the finitude of the human intellect to account for the human inability to grasp that things are as they are out of necessity. Rather than articulate a metaphysical challenge to the claim that b) causal relations are captured by explanations at the level of physics and c) that these explanations are complete, explanatory gap theory appeals to epistemic dualism (or dualism of access) to support, rather than undermine, the completeness of explanations at the basic level of physics. Understood as an *epistemic* problem generated by two very different modes of access (one first-personal and one third-personal), the mind/body problem will outlive a fully developed neuroscience. Rather as for Leibniz the epistemic gap between human and divine knowledge poses no threat to the claim that if a proposition is true then the predicate is either overtly or covertly contained in the subject, so for explanatory gap theory the existence of an unbridgeable epistemic gap constitutes no threat to the completeness of physics.

Others (Jackson 1986) have wanted to draw more robust ontological conclusions from the recalcitrance of phenomenal concepts to be explained in physical terms. They have argued that the inability of physicalism to account for the phenomenon of consciousness is indicative not of an epistemic failure to comprehend how physicalism could be true, as explanatory gap theorists maintain, but of the falsity of physicalism. Jackson’s famous argument about Mary, the neuroscientist locked in a black and white room, aimed to show that even if one could, like Mary, know all the neurophysiological facts about colour, one would still lack knowledge of what colours such as red look like. And if there is some knowledge that Mary would acquire on exiting the black and white room (knowledge of what red looks like), then physicalism must be false because there is a certain kind of knowledge, knowledge of phenomenal concepts, that remains beyond the ken of a complete neuroscience. Jackson’s argument established that if Mary’s knowledge is incomplete, then even if the explanations at the level of physics can tell us everything about the physical, they cannot provide a complete explanation of reality, as physicalists claim: physicalism is therefore false, not merely incomprehensible from the human point of view, as explanatory gap theorists argue. Jackson’s argument rejects one of the cluster of assumptions that belong to the reductivist paradigm, namely that c) physical explanations provide complete explanations of reality. Yet even forms of non-reductivism such as Jackson’s, which draw more substantive ontological conclusions from the inability of physicalism to explain the phenomenal properties of objects, stop short of challenging the assumptions that a) to explain is causally to explain and that b) real causal relations are captured by explanations at the basic level of physics. Rather than jeopardizing the assumptions that b) only explanations at the basic level of physics capture real causal relations, Jackson endorses epiphenomenalism about qualia. Qualia exist *and* cannot be comprehended from a naturalistic standpoint; but since they do no causal work they do not interfere with causal explanations at the basic level. Epiphenomenalism is therefore seen not as a problem, but as a solution to the problem of causal overdetermination.

Other forms of non-reductivism have moved away in various ways from the metaphilosophical assumptions that govern the reductivist paradigm. These, as we have seen, were that a) to explain is
causally to explain, b) the idea that genuine causal explanations operate at the basic level of physics, and c) that the explanations at the basic level of physics are complete. Multiple realization functionalists such as Fodor (1989 and 1997) have rejected b) that only the explanations offered by physics express genuine causal relations. They have argued that the explanations of the special sciences (chemistry, biology and psychology) are not merely useful for pragmatic purposes, but have the status of genuine causal laws. Fodor (1974) mobilises these considerations in support of an argument for the autonomy of special sciences such as psychology, sociology and economics.² Fodor (1989, 64), for example, has claimed that causal laws are generalizations which specify nomologically sufficient conditions for the occurrence of an event. On this understanding of a causal law, there is no reason for doubting that the laws of the special sciences are genuine causal laws. The difference between the special sciences and physics is that the laws of physics, being strict, deductively entail the explanandum, whereas the laws of the special sciences, being non-strict, entail it only probabilistically. But the fact that the laws of the special sciences are probabilistic does not necessitate the conclusion that whereas the generalizations of the special sciences are useful for pragmatic purposes only, those of physics capture genuine causal relations which hold between events independently of our explanatory goals. Explanations in the special sciences, including psychology, express genuine causal laws, not generalizations which are useful for mere pragmatic purposes. Multiple realisation functionalism, Fodor argues, provides an effective antidote to what he describes as a recent outbreak of “epiphobia” among non-reductivists, or “the fear that one is turning into an epiphenomenalist” (Fodor 1989, 59) because it rejects the widespread view that a defence of the autonomy of the mental rests on either a covert or explicit commitment to epiphenomenalism.³ Since the laws of the special sciences are genuine causal laws, there is no reason to doubt the causal efficacy of the mental qua mental. It is unclear however, that multiple realization functionalism is better equipped than non-reductive physicalism to navigate the straights of epiphenomenalism and explanatory exclusion. By arguing that explanations in the special sciences have the status of genuine causal laws, multiple realization functionalists must commit themselves to the view that there are as many layers of reality as there are levels of explanation in order to avoid causal over-determination: thus while they certainly avoid epiphenomenalism and explanatory exclusion, they do so at the price of ontological profligacy.⁴

**Broadening the non-reductivist agenda**

While multiple realization functionalists challenged the assumption that b) it is only explanations at the level of physics that capture real causal relations, they never really questioned the assumption a) that to explain is to explain causally. Functionalists (whether they do or do not allow for multiple realizability)

---

² On this see also Block (1997a and 1997b).
³ For the epiphenomenalist objection see Honderich (1982), Smith (1984) and Hutto (1999).
⁴ For this charge see Heil (1999, 2003a, 2003b)
define the mind in terms of its causal roles. The defence of the irreducibility of the mental that one finds in multiple realization functionalism is thus a very far cry from the kind of defence of the autonomy of humanistic explanations which one found, for example, in W.H. Dray’s (1957, 1958, 1963, 1980) response to Hempel’s (1942) argument for methodological unity in the sciences or in Peter Winch’s (1958) articulation of a humanistically oriented social science. As some might remember, W.H. Dray’s defence of the autonomy of history against Hempel’s call for methodological unity in the sciences was premised on the assumption that explanations in the human sciences are rational, not nomological, and thus that the distinction between humanistic and scientific explanation is a distinction between different kinds of inference, not a distinction in the degree of accuracy allowed by their empirical predictions. This older form of non-reductivism was premised on the assumption that the domain of culture, unlike that of nature, is beyond the reach of causal explanations because understanding other cultures requires understanding agents as responding to norms rather than as following laws of nature. Unless such a distinction is made, the actions of the believer who genuflects at the altar would have to be explained in the same way as the behaviour of the sunflower that turns towards the sun. The distinction between the study of nature and culture, on which this older form of non-reductivism rests its defence of the autonomy of the human sciences, is therefore fundamentally different from the distinction on which multiple realization functionalists such as Fodor rest their defence of the autonomy of the mental. Functionalism may not be reductivist from an ontological point of view, but it is unapologetically reductivist from a methodological perspective. Fodor’s distinction between physics and the rest of the special sciences is arguably an up-to-date version of Mill’s distinction between the exact and the inexact sciences. (Mill 1843, book VI, chapters 3 and 4). In Fodor’s view, as indeed Mill’s, the explanations of psychology differ from those of physics merely because the empirical laws of psychology, unlike those of physics, are hedged by ceteris paribus clauses, not because they are sui generis rational explanations with a different logical form from the explanations of natural science. It is precisely the assumption that to explain is to explain causally that was challenged by the methodological non-reductivism defended by Collingwood (1944), Dray (1957) and Winch (1958) on the grounds that explanations which appeal to norms are different in kind and not merely in degree from explanations which appeal to empirical regularities.

From the list of the usual suspects, the form of contemporary non-reductivism which comes closest to challenging the assumption a) that ‘to explain’ is merely a shorthand for ‘to explain causally’ is Davidson’s (1980) Anomalous Monism, which claims that the domain of the mental has a normative dimension which is absent from the domain of nature (Davidson 1963). The motivation for Anomalous Monism was not just to explain how the mental could be causally efficacious (as in the case of more standard forms of non-reductive physicalism) but also to argue in defence of a methodological rift.

---

5 Some functionalists are also reductivists (Armstrong 1968; Lewis 1972; Kim 1998) because they argue that the causal roles which define the mental are played by physical states.
between the *Geisteswissenschaften* and the *Naturwissenschaften*. There is, in Davidson, an echo of an older methodological divide between the human and the natural sciences which underpinned the defence of the disciplinary autonomy of history articulated by W.H. Dray or the defence of a humanistically oriented social science to be found in Winch’s *The Idea of a Social Science*. The echo of this older debate is, however, only a distant and weak one, since the metaphilosophical picture which underpins Davidson’s Anomalous Monism allows for the possibility of a neutral monism that would be problematic from the standpoint of forms of non-reductivism which are premised on the view that method determines subject matter and consequently that ontological questions are internal to methodological ones. The latter would deny rather than endorse the view that there could be such a thing as ontological neutrality. The methodological distinction between the structure of explanation in the human and natural sciences (which underpinned Dray’s and Winch’s defence of a humanistically oriented philosophy of history and social science) was based on a very distinctive metaphilosophy, one which treated causation as an *intensional* relation holding between the *explanans* and the *explananda* characteristic of any given form of inquiry, not as an *extensional* relation holding between particulars independently of how they are described within a given form of inquiry.\(^6\) They would have therefore rejected the idea of an ontologically neutral monism, *not* because they were committed to ontological dualism or pluralism, but because they denied the possibility of an inquiry-independent subject matter which is on the other hand presupposed by metaphysics as traditionally conceived.

For both Dray and Winch the differences between the different methodological assumptions governing the human and the natural sciences are brought to the fore when the focus shifts from the present to the past. The principle of the uniformity of nature, which is essential for the articulation of nomological explanations is clearly inadequate to investigate the actions of individuals who do not share the same intensional context as that of the investigator. It is in vain, for example, to expect that a historian will be able to make sense of the behaviour of agents who, at the time of the Black Death, rubbed chicken skin on their swollen lymph nodes by seeking to understand their actions in the light of medical knowledge that was not available at the time. The scientific investigation of nature operates under the presupposition of uniformity (scientists must presuppose that the plague is caused today by the same bacillus that caused it in medieval England or else they could not test for it); historians and cultural anthropologists, by contrast, must be open to the possibility that the norms to which agents respond change over time on pain of failing to understand what they are studying. Since the differences between the methodological assumptions which govern the human and natural sciences are brought into relief when one casts one eyes upon the past, it was no coincidence that the methodological dispute between reductivists and non-reductivists had been originally fought on the terrain of the philosophy of history and cultural anthropology rather than on that of the philosophy of mind. And once the debate

\(^6\) On Davidson’s “extensional reply” to the charge of epiphenomenalism see Hutto (1999), chapter 6.
between reductivists and non-reductivists moved away from the philosophy of history and social science it lost its methodological edge.

The methodologically-informed non-reductivism which was advocated by Dray and Winch, which still resonates, even if only very faintly, in Davidson’s Anomalous Monism, was quickly left behind as the metaphilosophical assumptions which underpinned it became more and more unpopular towards the end of the twentieth century. As the ontological backlash against the linguistic turn ushered a return to a pre-Kantian conception of metaphysics, the defence of the autonomy of the special sciences also changed its face. As we have seen, the challenge for non-reductivists became that of accounting for how explanations at the higher levels of the special sciences, chemistry, biology, psychology and so on, could be reconciled with the view that the explanations at the level of physics are ontologically basic. The fundamental problem of the philosophy of mind became what Huw Priced has coined as the “Placement Problem”, and the question to be addressed by non-reductivism became “how can mind fit within nature?” Shining the metaphilosophical light on the philosophy of mind allows one to see that certain problems, such as those of epiphenomenalism and explanatory exclusion, belong to a conceptual framework which inherits one or more of a cluster of assumptions from the reductivist paradigm. The fundamental idea behind this special issue is to rethink the philosophy of mind outside the standard metaphilosophical box, to broaden the possibilities open to non-reductivism.

Beyond the placement problem

The metaphilosophical assumptions that give rise to the Placement Problem have been challenged in two different ways.

One important challenge to the paradigm that gives rise to the Placement Problem has come from philosophers operating under the label “liberal naturalism”. Liberal naturalism is arguably a loose label which is applied to philosophers such as Huw Price (2004 and 2013) and John McDowell (1966 and 1998) who have very different philosophical orientations. According to John McDowell, the Placement Problem arises because of the overly restrictive conception of nature that non-reductivists have inherited from the reductivist paradigm. McDowell, like Davidson, is moved by the consideration that the form of intelligibility operative in the human sciences, that of the logical space of reasons, is irreducible to the form of intelligibility which governs the natural sciences, i.e. law-like explanations. But he also believes that the tension between the space of reasons and that of naturalistic vocabulary is a relatively recent problem which has arisen because of the impoverished conception of nature inherited from the scientific revolution. Nature, McDowell argues, should instead encompass both first nature (the object of investigation of the natural sciences) as well as second nature, our culturally acquired ability to become responsive to reasons. While second nature remains beyond the ken of scientific explanations, it is a part of nature, more generously conceived, in that it involves nothing supernatural.7

7 For a critical appraisal of the liberal naturalist strategy see Papazoglou (2012).
We should, according to this picture, think of the space of reasons, our second nature, as part of nature because it belongs to our human mode of living and, as McDowell notes, ‘what is specifically human is surely natural (the idea of the human is the idea of what pertains to a certain species of animals.)’ (McDowell 1996, 77) Huw Price’s liberal naturalism, on the other hand, argues that the Placement Problem arises out of a representationalist assumption, namely the idea that our vocabularies, of science as well as every-day speech, philosophy etc., aim to represent how things are in the world. According to the representationalist thesis there is a word-object relationship that our language, both scientific and non-scientific, aims to get right. Price calls the kind of scientific naturalism that is guilty of the representationalist assumption object naturalism. Price argues that if we think instead of the statements of science, ethics, aesthetics etc. as parts of different vocabularies, as different tools with different purposes (often non-representational) the placement problems that scientific naturalism faces disappear. Price then proposes a different version of naturalism, one free from the tacit representationalist thesis, one that he calls subject naturalism: “According to this second view, philosophy needs to begin with what science tells us about ourselves. Science tells us that we humans are natural creatures, and if the claims and ambitions of philosophy conflict with this view, then philosophy needs to give way. This is naturalism in the sense of Hume…and arguably Nietzsche.” (Price 2004, 73) For Price, then, the question about terms and forms of explanation to do with the mental or the human sciences become problems about human linguistic behaviour, more of an anthropological/genealogical project about how these vocabularies come about and what role they play within human life, rather than about where these items fit within the furniture of the world, a question of metaphysics.

Liberal naturalists then either argue in support of a more generous ontological account of nature, like McDowell, one which allows space for the mind and the subject, or, like Price, reject the representationalist/realist accounts of language, and argue for a pragmatist account of the seemingly problematic terms in question, thus avoiding engaging in the metaphysical struggle that reductivism and anti-reductivism engage in. For the liberal naturalists, then, twentieth century non-reductivism is a non-starter since it operates under the fundamental presuppositions of reductivism, whether about the ontology of nature or the relationship of language to the world, the very assumptions that are the source of the Placement Problem. But Liberal naturalists do not argue against naturalism as such, but simply for a more liberal naturalism, one with a human face, that allows us, in the case of McDowell to reject supernaturalist metaphysics, and in the case of Price, to continue seeing philosophy as essentially continuous with the natural sciences.

8 “The popular kind of naturalism...exists in both ontological and epistemological keys. As an ontological doctrine, it is the view that in some important sense, all there is is the world studied by science. As an epistemological doctrine it is the view that all genuine knowledge is scientific knowledge.” (Price 2004, 73)

9 In (Price 2013) Price allows that some vocabularies do indeed have a representational purpose, but that can only be revealed if we first treat language’s function as being broader than just representational.
Another important challenge to the philosophical paradigm that gives rise to the Placement Problem come from philosophers working in what may be very loosely called a Kantian idealist tradition who have questioned the layered view of the sciences (with physics at the base) with which much twentieth-century non-reductivism has operated and rejected the conception of metaphysics which underpins it. These philosophers have argued instead for a bifurcated view of the sciences underpinned by the metaphilosophical view that the task of metaphysics is not to describe the most general structures of reality, but to elucidate the presuppositions or principles which govern different forms of knowledge. Within this conception of metaphysics, the fundamental task of a non-reductive philosophy of mind is to articulate the distinctive methodology of those forms of inquiry which are concerned with culture rather than nature, not to explain the phenomenon of consciousness or account for how the mental can be causally efficacious. Once the underlying metaphysics is changed and causal relations are understood as explanatory relations serving the different investigative goals of the human and the natural sciences, the problems which besiege most contemporary non-reductivism in the philosophy of mind no longer arise, since they are the product of a layered understanding of the sciences premised upon a conception of metaphysics as the study of the most fundamental structures of reality.

It is in this spirit that philosophers such as Wilhelm Windelband and Heinrich Rickert focused on the nature of judgment in the Kulturwissenschaften and Naturwissenschaften. Windelband (1894/1980) argued that the Kulturwissenschaften are ideographic rather than nomothetic because they are primarily concerned not with the formulation of general laws but with singular judgments about past facts. So whereas the physicist is interested in ironing out the differences between different material objects, and finding out what they all have in common, and how they all behave under certain conditions, the historian is interested in unique historical events or figures because they are exactly different from any other such events or figures. The two distinct methodological approaches are linked to the ends of prediction in the case of ‘nomological science’, and fine detail and nuance in the case of ‘ideographic’ science. Similarly, Rickert (1902) starts with the Kantian thought that we are presented by the world with an infinite manifold of sensation which we need to organise and simplify in some way, because we are finite ourselves. Depending on how we choose to synthesize the manifold of sensation, reality presents itself as nature insofar as we treat it by universalizing methods, and reality presents itself as history insofar as we deal with it by individualizing methods. Therefore “Empirical reality becomes nature when we conceive of it with reference to the general, it becomes history when we conceive it with reference to the distinctive and the individual.” Unsurprisingly, in this neo-Kantian tradition, the problem to be addressed by non-reductivism was taken to be neither the problem of mental causation, nor the problem of consciousness, but the problem of articulating the distinctive methodologies of the human and the natural sciences. It in this spirit too that Dilthey defended the methodological distinction between Erklären and Verstehen and argued for a distinctive empathetic method for the human sciences.

In the context of a Kant-inspired metaphysics that extended the Copernican turn to the special sciences, this problem was to be addressed through a systematic elucidation of the a priori principles operative
in different spheres of knowledge. This methodological approach to the defence of the autonomy of the mental lives on in the conceptual idealism of Michael Oakeshott (1933, 1983) and R.G. Collingwood (1944, 1999). In fact it was the latter’s arguments for explanatory pluralism that were mobilized by Dray in the debate with Hempel. However, the original inspiration remains Kant, who argues that the account of reality as nature that physics puts forward presupposes a set of first principles that define the domain of investigation as that which is law-like from the outset. Physics does not discover that the world is law-like through experience, but already approaches the world of experience assuming it conforms to law-like descriptions: “Nature is the existence of things, insofar as that existence is determined according to universal laws.” (1783, §14).

Where does this methodological approach to the defence of the autonomy of the mental, which takes its cue from the Kantian idealist tradition, sit in relation to the cluster of assumptions that tend to constrain the possible moves available to the most widespread forms of non-reductivism? Methodological non-reductivism rejects the assumption a) that ‘to explain’ is shorthand for ‘to explain causally’. The form of explanation that is operative in the human sciences is variously described as ideographic (Windelband), empathetic (Dilthey) or rational (Dray); while there are very significant differences between the alternative suggested, all these forms of non-reductivism argue that the explanations operative in the human sciences are different in kind and not merely in degree from the nomological inferences which they take to be operative in natural science. Methodological non-reductivism also rejects the assumption b) that explanations at the basic level of physics are ontologically basic because they alone describe real causal relations that hold between events independently of the method through which they are investigated: these methodological non-reductivists took the main role of philosophical analysis to be not that of investigating the most general structures of reality, but that of systematically uncovering the a priori structures which govern different forms of knowledge. While this Kantian-inspired methodological non-reductivism rejects assumptions a) that to explain is causally explain and b) that explanations at the level of physics are ontologically basic, it does not deny assumption c) that explanations at the level of physics are complete. The latter is not denied provided that the principle of causal closure is recast as a presupposition or heuristic principle of the scientific investigation of nature, one of those a priori principles that is the task of philosophy to elucidate. Understood in this way, the principle of causal closure governs the way in which phenomena are rendered intelligible within natural science and no longer entails that the form of intelligibility at work in physics captures extensional relations which hold among events independently of the method through which they are explained or made intelligible. To claim that explanations at the level of physics are ontologically basic would amount to claiming that there is one form of knowledge that captures things as they are in themselves, and this would be tantamount to undoing the Copernican revolution.

Once the principle of the causal closure of physics is understood within the framework of a metaphilosophy in which method determines subject matter, it no longer gives rise to the twin problems
of epiphenomenalism and explanatory exclusion. These are problems which belong to a form of non-reductivism that has inherited too many assumptions from the reductivist paradigm. There are very significant differences between philosophers such as Rickert and Windelband who were working in the neo-Kantian tradition sensu stricto, the hermeneutic tradition of Dilthey, and the defence of the autonomy of history and social science articulated by Dray and Winch in analytic philosophy of history and social science; however, these philosophers all share a rejection of the methodological monism that is unproblematically assumed by multiple realization functionalists, and a rejection of the priority of ontology over epistemology that governs the conception of metaphysics underlying non-reductive physicalism. Since they deny the claim that there is a level of explanation which is ontologically basic, they have no the need to appeal to supervenience to account for the relation holding between the basic level of physics and the higher level of psychology. The relation of supervenience, which plays such a pivotal role in forms of non-reductivism committed to a layered model of the sciences, has no significant job to do in a metaphysics which understands nature as the correlative of scientific method.

While both liberal naturalism and the methodological non-reductivism which one finds in the post-Kantian idealist tradition advocate a shift in the metaphilosophical assumptions that govern the best-known forms of non-reductivism, there are significant differences between which of those assumptions they target. Neither McDowell nor Price operate in what may be loosely called the neo-Kantian idealist tradition. Their solutions to the Placement Problem does not involve reversing the order of logical priority holding between epistemology and ontology, or advocating a Kantian metaphysics in which method determines subject matter. But while the differences between these two approaches are significant, both attempts have in common the view that the questions which non-reductivists tried to answer are the wrong questions to ask, and that one should seek to re-orientate the questions to be addressed by the philosophy of mind rather than trying to find solutions to the problems defined by a philosophical agenda which was largely dictated by reductivists. Both approaches, therefore, agree that the Placement Problem is a problem that belongs to a certain way of doing philosophy and that once the metaphilosophical assumptions which belong to that conception of the role and character of philosophical analysis are abandoned, the problems that it attempted to solve no longer arise: what needs to be done is to change the items for discussion on the philosophical agenda.

There is something, however, that the liberal naturalism of McDowell, at least, has in common with the Kantian idealist tradition. His defence of the autonomy of the mental is more clearly focussed on the need to explain the relation between the space of reasons and that of nature, or law-like explanations. To this extent, liberal naturalism, at least in McDowell’s formulation, operates in the shadow of Kant’s distinction between theoretical and practical reason. There is a clear acknowledgement in McDowell’s liberal naturalism that the natural and human sciences are different ways of making things intelligible and that trying to understand one in terms of the other inevitably leads to changing the subject matter under discussion. This problematic is not, as we have seen, at the forefront of the standard forms of non-reductivism in the philosophy of mind which tend to be
preoccupied by question as to whether the mental is epiphenomenal or causally efficacious and, if the latter, how can such a claim be made good without abandoning one or more of the three assumptions we have discussed. With the exception of Anomalous Monism, forms of non-reductivism articulated from a naturalistic platform did not acknowledge the existence of a robust methodological divide between the Geisteswissenschaften and the Naturwissenschaften. Multiple realization functionalism assumes without further ado that the main difference between physics and the rest of the special sciences is that the generalizations at work in these different domains are formulated in different vocabularies, those of physics, biology, chemistry, psychology, etc., not that they make use of different kinds of inference. It is the nature of the inference at work in the Geisteswissenschaften and the Naturwissenschaften on the other hand, that tends to preoccupy McDowell’s liberal naturalism.

**Between naturalism and supernaturalism: is there a third way?**

What we have here are clearly forms of non-reductivism that break away from the naturalist or at least the hard-naturalist mould in which much non-reductivism has tended to operate. But can forms of non-reductivism which deviate from the naturalist/hard-naturalist mould succeed where non-reductive physicalism, explanatory gap theory, epiphenomenal dualism and multiple realization functionalism have failed? Some have claimed that it is not just forms of non-reductivism which seek to articulate a defence of the autonomy of the mental against the metaphysical background of a hard naturalism which fail in their task. Sebastian Gardner (2007), for example, has recently argued that there is no logical space to be occupied between (hard, scientific) naturalism and supernaturalism and thus no platform for the claims of neo-Kantian idealists and liberal naturalists alike.10 Orthodox or hard naturalism, as we have seen, is rejected both by liberal naturalists and by idealists of a Kantian bent. The former inflate the concept of nature with which hard naturalism works. The latter deflate the ontological ambitions of more traditional forms of idealism and accommodate a plurality of perspectives within a Kantian conception of metaphysics that prioritizes epistemology over ontology. Neither, so Gardner argues, succeeds in delivering a stable position which does not ultimately collapse into the polar opposites of hard naturalism and supernaturalism. The soft naturalist who seeks to introduce a plurality of perspectives within the constraints of a relaxed naturalist/monist ontology will ultimately have to explain how the different perspectival conceptions of reality it allows for are accountable to the harder reality captured by the scientific perspective. The attempt to find an intermediate position between hard naturalism and supernaturalism within the idealist tradition does not, in Gardner’s view, fare any better. In what we have loosely described as the neo-Kantian tradition, the accusation that the rejection of naturalism leads to a commitment to supernaturalism is avoided by arguing that nature is not a thing in-itself, but the correlative of the method of the

---

10 For a defence of liberal naturalism against this sort of accusation see De Caro and Voltolini (2010) and Giladi (2014).
Geisteswissenschaften. The Geisteswissenschaften and Naturwissenschaften have, therefore, different explananda that are known through the distinctive inferences at work in the forms of inquiry to which they belong. Being is not known qua being but as the distinctive reality that is disclosed by the sciences of nature and mind. Gardner objects that by making subject matter relative to method this idealist solution weakens the distinction between knowledge and reality: reality is no longer understood as what is mind-independently real but as the inquiry-dependent subject matter of a form of investigation. Once the distinction between knowledge and reality is weakened two possibilities are open: either a) to accept that the defence of methodological pluralism is “merely” epistemological and entails a commitment to a form of subjective idealism or b) to endorse a form of full blown ontological idealism which rejects the existence of a robust distinction between knowledge and reality, thought and being. If one acts on the first horn of this dilemma, then naturalism is not defeated since a pluralism that is merely methodological/epistemological leaves the door open for the claim that being qua being may be captured by the methodological practices of natural science. If one acts on the second horn of the dilemma, then one commits ontologically to a form of idealism that is much more robust than the one defended in the neo-Kantian tradition and fails to avoid supernaturalism.

Gardner’s conclusion, however, may be overly pessimistic. It could be argued for example, that given the tight reciprocal relation holding between method and subject matter in any form of inquiry it is impossible for anyone to explain anything unless they are committed to a set of a priori principles that give rise to the questions characteristic of a form of inquiry. If so, the very idea of comparing inquiry-dependent subject matter with inquiry-independent reality is in fact illusory. And if such a comparison genuinely is illusory, then the claim that method determines subject matter need not slide into a commitment to a form of subjective idealism that is epistemically humble and agnostic concerning the ultimate structure of reality. It may, therefore, be possible to endorse a form of idealism that can accommodate the naturalist perspective as a perspective alongside that of the human sciences, without either a) subscribing to a form of subjective idealism that must remain agnostic concerning the ultimate structure of reality or b) making the heavy-duty ontological commitments which Gardener suggests are needed to defeat naturalism.11 Whether it is possible to steer this via media between naturalism and supernaturalism will largely depend on how effective the new kinds of non-reductivism developed either in the idealist or liberal naturalist tradition are in blocking the sort of questions that belong to the kind of metaphysics they want to leave behind.

There is hopefully something to be gained by broadening the discussion of non-reductivism in the philosophy of mind to include positions which cast doubt on what are perceived to be the important problems of the philosophy of mind (the problem of mental causation or the problem of consciousness, for example) and which urge a reconsideration of what the key questions are which a non-reductive

---

11 For a defence of the claim that idealism need not collapse into either subjective idealism or ontological idealism see D’Oro (2017 and forthcoming 2018).
philosophy of mind should be addressing. Shining the metaphilosophical light on the discussion of non-reductivism in the philosophy of mind shows, amongst other things, that while certain approaches may be more prevalent either within the analytic or the continental traditions, they are not the prerogative of either. There is, for example, more in common between the neo-Kantian positions of Rickert and Windelband and the non-reductivism defended by Winch and Dray in analytical philosophy of history than there is between Dray’s and Winch’s argument for the methodological autonomy of history and social science and the non-reductivism that has grown in the soil of the analytic philosophy of mind.

Before drawing this discussion to a close, we should point out that the forms of non-reductivism explored in this special issue are all committed to articulating a defence of the autonomy of the mental that does not involve an appeal to supernaturalism. That it is undesirable to reintroduce transcendent metaphysical entities in order to do justice to the mental is something that is assumed by liberal naturalists and post Kantian idealists alike. We have therefore not explored forms of ontological dualism which do not operate within that constraint. This is not to say that the assumption that it is desirable to avoid supernaturalism is completely unproblematic, and should not be put under any pressure, but rather to acknowledge that it is operative in all the papers selected for this special issue. While the view that a defence of the autonomy of the mental should be accomplished without appealing to transcendent entities has not been examined here, it should be clear from the preceding discussion that a rejection of physicalism does not automatically translate into a defence of the autonomy of the mental against certain reductionist assumptions. Jackson’s refutation of physicalism, as we have seen, leads to a form of epiphenomenal dualism that does not touch the assumption of the causal closure of the physical domain so dear to hard naturalists. Similarly Block’s multiple realization functionalism endorses rather than rejects the claim for methodological unity in the sciences. We also have not engaged with increasingly more popular forms of panpsychism, such as those of Galen Strawson (2006) or Philip Goff (2017) because they display a slightly too cosy relation with naturalism and it is unclear to what extent they deny that the mental can be explained scientifically. Be that as it may, all the contributors to this issue seek, albeit in different ways, to articulate a defence of the mental without introducing transcendent entities and they all agree in rejecting the claim that there is no middle ground between hard/scientific naturalism and supernaturalism.

Robert Smithson challenges the layered view of reality which has led some contemporary metaphysicists to advocate that we revise the common-sense ontology embedded in ordinary ways of speaking in line with scientific theory. He argues that revisionary metaphysicians are correct in identifying a conflict between the manifest and scientific image, but that they are wrong in inferring from the existence of a conflict between the scientific and the manifest image the falsity of the latter. The truth of our manifest judgments does not hinge on what the scientific image will turn out to be because manifest and scientific discourse are governed by different norms. This explains why manifest judgments about, say, the existence of tables, or the lateness of the bus, are never revised in the light of
scientific discoveries concerning the microscopic constitution of objects or quantum physics. This is not to say that manifest judgements are not revisable, but that the ways in which they are revised is answerable to their own domain-specific norms. Idealism here is invoked to explain both how there is a conflict between the scientific and the manifest images of reality and why this conflict need not be resolved by choosing one image at the expense of the other. Smithson develops this idealist position in the context of debates in contemporary analytic metaphysics but his argument for the autonomy of the manifest image closely parallels the view that given the tight relation holding between method and subject matter in any form of inquiry the attempt to capture the domain of culture in terms of natural scientific discourse only ends up changing the subject matter in question. There are clear affinities between Smithson’s idealist defence of the manifest image and the defence of ordinary objects articulated by philosophers such as Thomasson (2007) against the eliminativist tendencies of much scientific inspired revisionary metaphysics12 even though Thomasson does not see herself as working within the idealist tradition. But the reluctance of many philosophers to associate themselves with the idealist tradition may be more indicative of the fact that idealism has tended to be identified with an ontological claim which sees mind as being causally responsible for the world than the absence of any philosophical affinities between deflationary conceptions of metaphysics invoked in contemporary analytic metaphysics to curb the eliminativist tendencies of revisionary metaphysicians and ontologically more modest forms of idealism.

Plato Tse explores Fichte’s critical engagement with physicalism or what Fichte calls “dogmatic materialism”. The term “dogmatic materialism” indicates that the motivation for Fichte’s critique of physicalism is not the endorsement of supernaturalism, but a rejection of the realist assumptions which dominated a pre-critical conception of metaphysics: physicalism is an ontological claim which belongs to a pre-Kantian metaphysics that fails to acknowledge the presuppositions of knowledge and in this sense is a dogmatic ontological position. Fichte’s critique of physicalism arises from a rejection of any position that seeks to reverse the Copernican revolution by failing to recognize the norms which are constitutive of all “science”, including natural science. Fichte understands “science” as a body of knowledge whose object is logically rather than causally constituted by certain necessary acts of the mind. In this sense of “science”, the natural sciences owe their status as a body of knowledge to some implicit normative presuppositions that philosophy explicates. Causation, for example, is an explanatory norm of natural science. Idealism does not deny causation, but restricts its applicability to a particular body of knowledge, natural science. Since the task of philosophy is to explicate the normative presuppositions implicit in the sciences it is “the science of science” and as such cannot be continuous with any particular body of knowledge. The idealism on which Fichte rests his critique of dogmatic materialism therefore entails the rejection of any attempt to naturalize epistemology.

12 See for example Unger, P. (1979a and 1979b)
David Zapero presents the liberal naturalism of McDowell as appropriating the central insight of the Kantian tradition but dissociating it from idealism in order do justice to the ideas of truth and answerability to the way things are. In taking a distance from idealism McDowell sees liberal naturalism as safeguarding certain realist intuitions about the accountability of our claims to reality. McDowell’s liberal naturalism therefore has one foot in the Kantian tradition, in so far as it acknowledges there are different stances on reality (to which there correspond different forms of account giving), and one foot in the realist tradition, in so far it commits to a “minimal empiricism” according to which experience is the tribunal to which any form of account giving is answerable. Whereas Kant’s transcendental idealism identified this minimal empiricism with a form of lower-level empirical realism that contrasts with a higher level transcendental idealism, McDowell’s rejects the notion of the Ding-an-sich that transcendental idealism failed to completely eradicate and with it the corresponding notion of a form of objectivity which transcends all stances. Rather than doing justice to realist intuitions within the framework of transcendental idealism (by appealing to a form of internal/empirical realism) McDowell denounces the notion of objectivity attached to the thing-in-itself as a fantasy. Once the notion of objectivity as answerability to a notion of reality that transcends all stances is left behind, minimal empiricism can be understood independently of the apparatus of transcendental idealism and doing justice to the realist intuition that our stances must be answerable to reality no longer requires that the things to which our stances are answerable should “be purged of any link to the way we take up those things with our parochial, cognitive capacities”.

David Macarthur’s liberal naturalism has commonalities and differences from the idealist tradition. He suggests that the methodological distinction between the Naturwissenschaften and the Geisteswissenschaften that is invoked in the neo-Kantian tradition to defend the autonomy of the human sciences distorts the richness of methodological practices at work in the natural sciences, as well as the commonalities between the natural and social sciences. His liberal naturalism seeks to circumvent the placement problem by taking the second-person perspective as primary over both the first-person perspective of the human sciences and the third person perspective of physics (though it is worth noting that the second-person relation involves the first-person perspective with the reciprocal “I-thou” relationship), and focuses on the intersubjectively valid norms that govern all forms of inquiry. For Macarthur the scientific image is based on an abstraction from the normative “I-thou” relations which govern the community of investigators whose activities can be understood along the lines of what Wittgenstein called forms of life. Scientific intelligibility is therefore not fully captured by the impersonal, subject-independent third person perspective. It is on the contrary firmly embedded in an interpersonal space.

Like Macarthur, Paul Redding endorses a “recognitive” or second-person account of the mind but, unlike Macarthur, connects it to a form of inferentialism that denies the reducibility of intensional to extensional notions. It is the defence of the irreducibility of intensional to extensional notions that opens up the logical space in which a weaker form of idealism (one that does not collapse into a form
of supernaturalism) can operate. Redding makes this case by arguing that Hegel’s defence of the autonomy of the mental is to be understood in terms of a defence of the irreducibility of intensional to extensional notions that anticipates important developments in twentieth century logic. Extensionalism has normally gone hand-in-hand with metaphysical naturalism or materialism while intensionalism has normally gone hand-in-hand with idealism of an ontological nature. As a result, the defence of the irreducibility of intensional to extensional notions is normally considered to entail a commitment to ontologically robust forms of idealism. But Hegel’s defence of the irreducibility of intensional to extensional notions leads only to a weaker, non-ontologically committed form of idealism and may therefore be compatible with liberal naturalism. Hegel’s argument is that even particular (rather than universal) judgments such as “the rose is red”, whose truth appears to be a matter that can be settled extensionally are infected with intensional notions: a full understanding of their meaning requires locating them in a broader intensional framework where the judgment is part of a greater inferential totality that cannot be captured in extensional terms. “The rose is red”, for example, entails that the rose is not blue, that is it coloured etc. The claims “the rose is not blue”, “the rose is coloured” etc., do not describe real states of affairs; their meaning cannot therefore be captured in extensional terms. This account of the irreducibility of intensional to extensional notions subverts the Cartesian representationalist picture of the mind as an inner theatre and replaces it with an inferentialist account of judgment that sits more comfortably with Hegel’s “recognition” account of the mind in the Phenomenology of Spirit.

References

Raymond Geuss, R. Changing the Subject: Philosophy from Socrates to Adorno. Harvard University Press.
Kant, I. [1783] (2004). *Prolegomena to any Future Metaphysics That will be able to present itself as Science.* Oxford: Oxford University Press.


* This project publication and the special issue of which it is part was made possible through the support of a grant from the John Templeton Foundation. The opinions expressed in this publication are those of the author(s) and do not necessarily reflect the views of the John Templeton Foundation.