# THE MYSTERIANISM OF OWEN FLANAGAN'S NORMATIVE MIND SCIENCE

### by Mikael Leidenhag

Abstract. This article critically analyzes Owen Flanagan's physicalism and attempt at deriving ethical normativity from current neuroscience. It is argued that neurophysicalism, despite Flanagan's harsh critique of "the new mysterians," entails a form of mysterianism and that it fails to appropriately ground human mentality within physicalism. Flanagan seeks to bring spirituality and a physicalist ontology together by showing how it is possible to derive an account of the good life from science. This attempt is critiqued and it is shown that Flanagan fails to establish the consistency between ethical normativity and physicalism. Hence, another form of mysterianism seems to emerge within this normative mind science.

Keywords: consciousness; eudaimonia; Owen Flanagan; hard problem of consciousness; mysterianism; neurophysicalism

How to explain consciousness, its properties and capacities, has famously been referred to as a "hard problem." By taking the phenomenon of mind seriously, we have to tackle the deep issue of squaring our everyday presuppositions regarding human mentality with what is revealed by the natural sciences. Some have argued that such a project is doomed to fail and that the nature of consciousness is eternally beyond human conceptualization. Such "mysterianism" holds that creatures such as we are simply not cognitively equipped to deal with, or offer a systematic explanation for, the nature of consciousness. This principled agnosticism is strongly challenged by philosopher Owen Flanagan, who instead argues that the physical sciences, and neuroscience in particular, can provide a constructive naturalistic theory of mind. Through what Flanagan calls the "natural method" it is possible to solve the hard problem of consciousness. Moreover, if we can unveil the nature of mind we can shed light on what Flanagan dubs as the "really hard problem," namely how to find or place values in a fully natural world. The result is not only a naturalistic theory of the mind, but

Mikael Leidenhag is postdoctoral fellow in Theology and Science, School of Divinity, University of Edinburgh, Edinburgh. UK; e-mail: mleidenh@ed.ac.uk.

also a "normative mind science," a positive account of the good life, and a spiritual naturalism.

I seek in this article to critically evaluate Flanagan's science of mind, and his attempt to derive normative ideas from his naturalistic understanding of mind. More specifically, I will argue that Flanagan does not succeed in offering a physicalist explanation of consciousness. On the contrary, this physicalist position results in a version of mysterianism. Moreover, his metaphysical quietism regarding morality results in an additional ethical mysterianism. In the end, I suggest that Flanagan's twofold mysterianism bears witness to the deep philosophical problems of placing higher order features of reality, such as mind and values, within a physicalist framework.

#### EXPLAINING MIND THROUGH THE NATURAL METHOD

The major guiding point of Flanagan's project is that mind is a fully natural phenomenon that can be studied by the physical sciences. Flanagan, contrary to other physicalists, maintains that mind is objectively real and not an epiphenomenon that can be explained away as some "dispensable cog in the machine" (Flanagan 1992, 13). For us to grasp what makes us human, and in order to satisfactorily explain intelligent activity, the mind and its capacities are indeed essential categories. Flanagan resists "conscious inessentialism" and argues that there is no need for "quining consciousness" or to go down the same eliminativist path as Patricia Churchland. The verb "Quining," taken from Daniel Dennett, means to "deny resolutely the existence or importance of something seemingly real or significant, for example, the soul" (Flanagan 1992, 21). Flanagan argues that the mind, however poorly understood, is "robust enough to stand up and keep its ground against those who would quine it under" (Flanagan 1992, 28). Although Flanagan retains an objectivist view of consciousness, he still concedes and recognizes partly the validity of the "epiphenomenalist suspicion"; that is, it could be the case that consciousness is "a causally inconsequential byproduct, or side effect, of physical processes in our brains" (Flanagan 1991, 38). Epiphenomenalism has a point in that it objects to those accounts of consciousness that rely too heavily on the causal efficacy of human minds; causal accounts that stand in tension with physicalist interpretations of science. Nevertheless, epiphenomenalism is ultimately implausible and incoherent, according to Flanagan (1991, 39). Consciousness is real, and subjective awareness plays a role in our mental lives, although "exactly what role it plays, how important it is in fixing informational content, in what domains it is important, how it figures in remembering," and so on, are unsettled questions that need to be further addressed (Flanagan 1992, 151). They are matters to be settled in "an empirical court" (Flanagan 1992, 151).

Some naturalists, while conceding the realness of mind, maintain that a naturalistic explication of consciousness and its relationship to the brain is a pipe dream. It is one of the unsolvable problems that haunts and taunts philosophers and scientists alike. These naturalists are *nonconstructive* naturalists and they hold to, as Flanagan famously dubbed it, a form of *mysterianism*. The old mysterians were dualists who argued that mind cannot be fully understood because it operates according to nonnatural principles and is constituted by nonnatural properties. The new mysterians, however, are naturalists "with a kinky twist" who maintain that consciousness exists and is subject to natural laws (Flanagan 1991, 313). Yet, consciousness is characterized by a subjective aspect, or a first-person perspective, that cannot be articulated through, or be reduced to, naturalistic categories. Consciousness is fully natural but forever beyond scientific and naturalistic explication. This view is expressed to some extent by Thomas Nagel (1974), and is explicitly defended by the agnostic naturalist Colin McGinn (1989).

For McGinn, the link between consciousness and the brain cannot be rendered intelligible, because whatever observation we make about the brain is insufficient for capturing consciousness. Mere empirical observation regarding neurological firings in the brain structure does not explain, for example, the ontology of qualia and phenomenal properties, or how such phenomena can arise naturally and fit an evolutionary framework. The problem with this objection, Flanagan argues, is that it "holds excessively high standards of intelligibility" (1992, 112). It is common in science to infer unobservable phenomena from observable ones, in the same way as unobservable electrons are inferred from processes in a cloud chamber. Hence, the mysterianist retreat from explanation is unnecessary.

Given the high explanatory standards of the new mysterians it is, according to Flanagan, impossible to solve the hard problem of consciousness. Flanagan further argues that the problem of both dualistic and naturalistic explanations of the mind–brain relationship is that they treat the mind as a *thing*. Both reductive and non-reductive (i.e., emergent) naturalists have treated the mind as a thing, a some*thing* or an entity. Instead, we should, according to Flanagan, view mind as a functional state, analogues to how walking and breathing are functions of our bodies. Hence, mental "states are functional states and functional properties of the complex commerce we have with the outside world" (Flanagan 1991, 45). The brain is a "Darwinian machine, a device governed by principles of massively parallel processing and neuronal group selection" (Flanagan 1991, 323).

According to Flanagan, we should not adopt the pessimistic attitude of mysterianism. Instead, we work constructively on the basis of the "natural method." This method, being rather simple, means that we should listen carefully to what individuals report about how things seem, listen carefully to what psychologists and cognitive scientists have to say about mental life and the role of consciousness, and take into account the

current description of consciousness as offered by neuroscience. The goal is to synthesize these stories regarding the subjectivity of the mental and its relationship to the workings of the brain. While the mysterians and the anticonstructivists deny the validity of any method for explaining mindbody relations, Flanagan suggests a more optimistic approach that seeks to bring available sources of knowledge into dialogue with each other in order to reach an approximately true account of the origin and function of consciousness. As Flanagan says, "The object of the natural method is to see whether and to what extent the three stories [phenomenology, psychology/cognitive science, and neuroscience can be rendered coherent, meshed, and brought into reflective equilibrium, into a state where theory and data fit coherently together" (Flanagan 1996, 18). The phenomenal aspects are essential for explaining consciousness, but they do not exhaust the properties of the mind: "The hidden structure of conscious mental states includes their neural realization" (Flanagan 1996, 34). This is Flanagan's constructivist neurophysicalism. The natural method, according to Flanagan, gives constructive physicalism a competitive edge over the new mysterians' retreat from explanations.

### AGENCY, NATURALISM, AND THE MEANING OF LIFE

Agency is an indispensible part of folk psychology and is central for the function of many human practices. Notions of "agency" and "free action" are, of course, debated, but they generally signify "the ability to pay attention, the causal efficacy of conscious deliberation, reasons sensitivity, the capacity to act in accordance with desires, the capacity to consciously monitor and guide action" (Flanagan 1996, 58). While some naturalistic functionalists opt for the eliminativist route, Flanagan seeks to retain the reality of agency and make it compatible with a naturalistic outlook on physical reality and human creatures. By naturalism Flanagan means "the view that all phenomena are natural and subject to causal principles" (Flanagan 1996, 56). Thus, mind, morality, and other higher level phenomena are made of natural stuff and are explicable in terms of natural laws. There is no mysterianist escape route available for higher level phenomena, and the task for the naturalist is to show how the central ingredients of agency realism can be made coherent with causal principles and a neuroscientific description of human persons. This, as we have seen that Flanagan argues, is achievable through the natural method.

It is argued, by dualists and naturalistic eliminativists, that naturalism is unable to retain agential realism; that it inevitably ends up denying the reality of consciously initiated action. This kind of philosophical argument states that the naturalistic list of ingredients is too limited to support agency. Another more empirical argument takes its cue from the infamous experiment conducted by Benjamin Libet, suggesting that

science itself rules out free action (Libet 1985). Subjects were connected to electroencephalographs (EEG) which measured "the readiness potential" in the brain (the cortical area), which is assumed to subserve hand movement. Then the subjects were asked to spontaneously flex their hand when they felt like it. The experiment showed that, despite the lack of preplanning, the consciousness of an intention to flex the hand occurs 350 milliseconds "after the onset of the readiness potential and about 200 milliseconds before the muscle activation" (Flanagan 1996, 59–60). What this finding seemingly demonstrates is that the readiness potential precedes conscious intention, which in turn precedes the decision to move the hand. Libet concluded that the brain, unconsciously, initiated an act before the appearance of conscious intention. Hence, the conscious decision seems to be ontologically redundant and, therefore, epiphenomenal.

Is this now widely debated experiment a threat to Flanagan's naturalistic and functionalist account of agency? Flanagan finds the results of Libet's experiment rather uncontroversial. Only someone who starts off from Cartesian dualism would be surprised to find that brain processes are ontologically prior to conscious intentions and voluntary actions. For a naturalistic view of consciousness this is to be expected, according to Flanagan, and the correlation between neural processes and phenomenal happenings pose in no way a threat to the coherence of functionalism. Moreover, Flanagan suggests, only some neural activity is conscious. He writes, "All conscious processes occur in complex neural networks in which they both supervene on certain neural processes and are caused by and cause other mental processes... some of which are conscious but most of which are not" (Flanagan 1996, 60). Hence, the idea that unconscious behavior undermines a naturalistic account of agency can be challenged. The experiment does not conclusively establish that consciousness lacks a functional role. It could still be that conscious processes "serve as a middle link in a three-term chain," between unconscious brain processes and the physical activity involved in flexing the hand (Flanagan 1996, 62). Libet's experiment, therefore, provides the kind of results that one would expect if one believes that conscious processes supervene on neural processes, and that consciousness plays a *variable* role in different cognitive domains. (For other, more current, responses to Libet's findings see Sinnott-Armstrong and Nadel, 2010.)

Flanagan concludes that naturalism is able to accommodate the notion of conscious agents, and therefore naturalism is a sufficient ontology for normativity, ethics, and a vision of the good life.

Naturalism does not invite eliminativism, but can accommodate the findings of neuroscience while conceding the realness of conscious efficacy. He further argues that a refined account of agency sheds light on meaning and the good life. Indeed, he writes that "Agency, free-action, responsibility, and a meaningful life are not conceptual enemies but are, in fact,

required to make sense of each other" (Flanagan 1996, 63). It matters a lot what a person does in life; we leave "parts of ourselves in the world by having changed that world in directions that matter, that are positive" (Flanagan 1996, 10). This is a form of "naturalistic transcendence," and it "involves believing that there are selves, that we can in self-expression make a difference, and if we use our truth detectors and good detectors well, that difference might be positive, a contribution to the cosmos" (Flanagan 1996, 11). Flanagan concludes that this belief is to "have a kind of religion" (1996, 11). In this way, Flanagan singles out "agency" as a key element in his normative project. A robust notion of agency is crucial, because without it it becomes difficult to make sense of the ideas "that I am agent, that I am self-productive, and that I create or co-create some of the meaning my life has" (Flanagan 1996, 53). This is the first, initial step towards a normative mind science.

It is with his 2007 book *The Really Hard Problem: Meaning in a Material* World that Flanagan starts to more systematically outline the possibilities of a naturalistic spirituality. This, what he calls "project eudaimonia," seeks to derive a normative account of the good life based on, specifically, contemporary neuroscience. Naturalism is generally considered incompatible with the idea that there is objective meaning in the universe. That is, if naturalism is true, we are simply lumps of matter existing in an ultimately meaningless universe, and thus there is no real possibility for a normative mind science. Flanagan takes issue with this commonly held view, suggesting instead that meaning, perhaps paradoxically, comes from accepting a naturalistic account of nature and human personhood. Indeed, if one accepts the naturalistic story "and engages in realistic empirical appraisal of our nature and prospects, we have chances for learning what methods might reliably contribute to human flourishing" (Flanagan 2007, 4). Although naturalism itself poses no threat to meaning, scientism does. Scientism, according to Flanagan, "is the brash and overarching doctrine that everything worth saying or expressing can be said or expressed in scientific idiom" (2007, 22). Because meaning seems to go beyond strict scientific categories (meaning cannot, e.g., be expressed through the language of the hard sciences), it has to be denied. Scientism, however, is "descriptively false and normatively false" and not everything "worth expressing can or should be expressed scientifically" (Flanagan 2007, 23). It is important for the successfulness of Flanagan's project to establish the difference between the disenchanting spirit of scientism, on the one hand, and a naturalistic account of the human person, on the other. In order to arrive at a good balance between the spaces of meaning and science, the threat of scientism must be evaded.

What does it mean, then, to flourish, and how is it possible to ground a realist account of the good life in naturalism? Flanagan suggests that our ability to flourish depends on us partaking in *the spaces of meaning*, a notion

introduced by Nelson Goodman. It refers to how particular social objects contribute to the construction of our lived worlds. Through Goodman, Flanagan brings out the six spaces of meaning—art, science, technology, ethics, politics, spirituality—in which we make sense of things, orient our lives, and find ways of living meaningfully (Flanagan 2007, 12). Our particular "nature as social mammals requires us to find meaning in a culturally available Space of Meaning or not at all" (Flanagan 2007, 37). Flanagan's aim is to achieve balance between the space of science and the space of meaning, so as to arrive at a naturalistically informed spirituality that pays close attention to the findings of science.

A naturalized eudaimonia will descriptively seek to map out the nature, causes, and mechanisms of flourishing and normatively articulate why some ways of living are better than others. If the second task can be accomplished, then a normative mind science can get off the ground. Normative mind science is not a new thing, argues Flanagan, but can be traced back to ancient philosophy and Buddhism. It can be found in Aristotle's Nicomachean Ethics, which puts happiness in the context of living in line with reason and virtue (courage, justice, temperance, and wisdom). The project of eudaimonics is thoroughly empirical. One starts with a hypothesis about what "constitutes a good or healthy person" and "one asks questions about what causes and constituents contribute to or make up the well-functioning form" (Flanagan 2007, 112). Within Buddhism, the goal is to identify the specific causes of unhappiness and suffering (the major cause being attachment to false beliefs about reality), in order to develop techniques with which we find liberation from suffering and, in its stead, to bring happiness (Flanagan 2007, 116). Although Aristotelian philosophy and Buddhism share many similarities in their view of the good life and the virtuous person, Buddhists tend to place greater emphasis on the importance of loving kindness and compassion (Flanagan 2011, 167–68).

The moral science developed by Flanagan places ethical emphasis on flourishing, and in order to establish a sound account of human flourishing it must pass the test of a Rawlsian reflective equilibrium. Internally, a moral conception should be conceived as good within one's culture. Externally, a moral conception needs to pass intercultural comparison, and to stand up to competing moral visions of the good. This external test is called *wider* reflective equilibrium, with which we seek to bring our moral conceptions into dialogue with a host of spaces of meaning so as to ensure progress in the moral realm.

What can neuroscience, as a subset of the scientific space of meaning, tell us about flourishing and happiness? While not able to metaphysically ground particular values itself, neuroscience can aid in empirically evaluating particular conceptions of happiness, well-being, and overall life satisfaction. This has been done, argues Flanagan, in numerous scientific studies on the benefits of meditative practices for well-being, happiness,

and the reduction of stress, anxiety, and depression. Neuroscientifically, the subjective state of a positive mood is "reliably correlated with a high degree of leeward pre-frontal activity" (Flanagan 2007, 164). In this way, by combining the criteria of wider reflective equilibrium with the natural method and its focus on neuroscientific investigation into phenomenal consciousness, one can show with good precision what practices can produce the desired positive states required for living the good life. This is another crucial step toward a normative mind science. This "empirically inspired eudaimonics" can, according to Flanagan, "help to cure the disease of disenchantment by marking off reliable ways to flourish" (Flanagan 2007, 107–08). By bringing consciousness into the purview of scientific investigation (through the natural method) one can extract a normative, yet scientifically acceptable, account of human flourishing.

A place for ethical normative within a naturalistic ontology has positive implications, argues Flanagan, for the coherence of a naturalized spirituality. Normative mind science, thus, lends peace to the often assumed conflict between science and spirituality within the spaces of meaning. We find meaning in "spaces that are truthful, good, and beautiful" (Flanagan 2007, 187). There is no universal meaning, although "most everyone, no matter what space they are most absorbed in and by, will also typically want friends, companionship, family, and perhaps passionate love" (Flanagan 2007, 188). Nevertheless, it is not possible to talk about the meaning of life, as people should be given, within their particular context, the chance to construct meaning for their own lives. Flanagan concludes that, as long as naturalism is able to uphold the normativity of our meaning-finding enterprise, spiritual naturalism remains possible and plausible. Naturalism can "make room for robust conceptions of the sacred, the spiritual, the sublime, and of moral excellence" (Flanagan 2007, 189-90). It can also tolerate an *expressive theism*, which views theism not as a propositional statement that can either be true or false but as a *story* about the origin of the physical world. That is, "Whatever you wish that feels compelling, satisfying, rich and deep. We are only talking about stories" (Flanagan 2007, 191). There is no need to go in the direction of assertive or ontological theism. Naturalism, argues Flanagan, provides its own form of transcendence and framework against which to make sense of meaning, happiness, and flourishing.

Naturalism can, moreover, provide a corrective ontological lens through which to view existing religious traditions. Flanagan has recently argued for a naturalistic reinterpretation of Buddhist beliefs and practices. Buddhism is, generally speaking, a metaphysically modest tradition, but still involves beliefs in nirvana, nonphysical minds, realms of heaven and hell, karmic laws, and so on. Such beliefs have to be rejected, but remaining would be "an interesting and defensible philosophical theory with a metaphysics... a theory about how we come to know and what we can know, and an

ethics, a theory about virtue and vice, and how best to live" (Flanagan 2011, 3). Such an updated tradition would be rendered compatible with the neo-Darwinian theory of evolution and its commitment to scientific materialism (Flanagan 2011, 3). In this way, naturalism provides a bridging philosophical theory between the rich tradition of Buddhism and the theories and methodologies of the natural sciences.

To summarize, by locating the crucial steps for Flanagan's normative mind science it becomes possible to identify the criteria against which to judge the overall successfulness of this project, criteria which Flanagan himself sets up and acknowledges as crucial for the success of his project. First, Flanagan needs to separate his own naturalism from reductionist scientism. Second, consciousness needs to be fully naturalized. Third, Flanagan needs to establish the coherency of the idea that one can derive a positive account of human flourishing from the neurological workings of the brain. As I will argue below, Flanagan fails to (1) fully naturalize consciousness, and (2) derive a normative account of human flourishing from a neuroscientific portrayal of consciousness. Therefore, Flanagan's project fails the second and third criteria. Instead, it results in a metaphysical mysterianism which threatens the coherency of a normative mind science and undermines his own critique of the "the new mysterians."

## FLANAGAN'S CORRELATION ACCOUNT: MYSTERIANISM IN DISGUISE

The goal of Flanagan's neurophysicalism is to correlate mental events with physical ones through the natural method of bringing together neuroscience, psychology, and phenomenological accounts. This correlation account, which sees mental states as functions of physical interactions, seeks to map out the specific properties that underlie the instantiation of a first-person perspective. A scientific third-person perspective cannot fully capture the subjective point of view relevant for phenomenological experiencing. However, Flanagan argues that the coherency of constructive naturalism is secured as long as it is possible to correlate particular physical properties with the emergence of mental states. He further suggests that a robust naturalistic account of agency can be preserved, despite the tight causal link between the mental and the physical. That physical events underlie mental events is to be expected, and so Libet's experiment remains congruent with neurophysicalism. Flanagan's physicalist construal of consciousness resembles in many ways the biological account of consciousness that has been proposed by John Searle. Searle, compared to Flanagan, more keenly stresses the irreducibility of consciousness, but they both view consciousness as the totality of contingent mental and physical correlations, and the ambition is to explicate such correlations with as much detail as possible (Searle 2008, 69-80). Moreover, they both seem to suggest

that a physical account of such correlations is enough to establish the superiority of naturalism over non-naturalistic theories and explications of consciousness. Correlation is the best we can hope for.

However, to correlate particular physical and mental events does not seem to get us any closer to explicating the nature of consciousness. This point seems to be shared by Flanagan, who recognizes that "the thirdperson or impersonal point of view fails to capture the relevant first-person phenomenology" (Flanagan 1992, 118). There is an epistemic gap and the "nature of our access to what we are made of and to how we function as complex systems is different in kind and provides different information" compared to a first-person phenomenology (Flanagan 1992, 117). Yet, this mystery, or epistemic gap between the sciences and the subjectivity of persons, does not bother Flanagan. He thinks that it is enough for a naturalistic accommodation project to isolate "the specific properties that subserve first-person experience," and by explaining "why only you can capture what it is like to be you" (Flanagan 1992, 117, 94). However, this attempt at correlating physical and mental happenings is not sufficient for explaining the phenomenology of consciousness, especially not for a mental realist such as Flanagan. On Flanagan's natural method one seeks to explain consciousness by identifying the particular mechanism(s) causally responsible for the origination or instantiation of mental phenomena. The problem is that the appeal to mental ↔ physical correlation is not explanatorily sufficient: that is, it fails to identify the relevant mechanisms underlying mental states and experiencing. (For more on explanation in philosophy of mind, see Nagel 2012, 47–48, 51–52; Seager 1999, 33–59; Hasker 1999, 62-80; Flanagan 1991, 129-31; Campbell 2009, 34-51; Rowlands 2001, 51–74; Chalmers 2010, 59–100, 305–36; Horgan 1993, 555–86; Kim 1988, 225–39; Kim 1993, 237–43, 250–54; Kim, 2007, 93–114.)

A correlation account would be sufficient if you adopted an anti-realist stance on qualia and phenomenological properties, perhaps in a similar vein to Daniel Dennett who opts for the elimination of qualia. Flanagan's mental realism affirms the reality of subjective consciousness and so seeks to bring out its neural underpinnings. The problem for this account is that mere correlation does not explain why a particular set of brain processes would give rise to qualia at all (Chalmers 1996, 115). It is left as a brute fact. Indeed, Flanagan objects to further philosophical enquiry at this point, stating that some "patterns of neural activity result in phenomenological experience; others do not. The story bottoms out there" (Flanagan 1992, 58). Of course, in any theory of consciousness, we would like to know *how* experience is possible and *why* I have experience<sup>1</sup> rather than experience<sup>2</sup> (Seager 1999, 55). Here we can start to see the mysterianism of Flanagan's neurophysicalism through this attempt to bypass the origination issue of phenomenal properties.

Flanagan, as we saw earlier, complains about the high standards for intelligibility associated with Colin McGinn's agnosticism. Because of this standard, McGinn concludes that the riddle of mind is unsolvable. I think that Flanagan's critique of McGinn has some force. The problem, however, is that Flanagan flees in the direction of another extreme by significantly lowering the standard of intelligibility so as to render mind consistent with the physicalist framework (with its ontological emphasis on physical processes). He is not interested in explicating the origin or nature of the mental—or how mental phenomena appear to conscious creatures—but only the causes and mechanisms underlying mental states. Flanagan, it seems, employs a form of definitional reduction in his account of qualia; the concept of qualia becomes identical to neural activity or the sum of such activity. This is why Flanagan denies the private character of subjective consciousness (1992, 65). Qualia, according to Flanagan, do not exist if they are defined as "essentially private, ineffable, and not subject to thirdperson evaluation" (1992, 67). The problem is that anyone who wishes to further pursue the quest for understanding qualia or mental states in general is told by Flanagan that "the story bottoms out here" (Flanagan 1992, 58).

First of all, qualia seem to escape physical reduction. Take as an example Flanagan's strategy of reducing sensory qualia to neural vectors, the raw feels "of taste, touch, sight, and the like" (Flanagan 1991, 326). Now, if this was true then qualia and with them their phenomenological features should in some sense supervene on the physical and be derivable from physical processes. That is, if we can get a full description of the physical and neurological interactions of the brain, then we should also get an account of the mental. If it is the case, as Flanagan claims, that sensory qualia "are just the characteristic spiking frequencies of activation patterns... in the relevant sensory pathways" then a full description of particular activation patterns should be enough for explaining qualia (Flanagan 1991, 328). If Flanagan maintains that physicalism contains the whole set of microphysical truths, then it should be possible to derive all macrophysical truths from microphysical truths. The problem is that no physical property seems sufficient for capturing the qualitative feel associated with subjective experiencing. This conclusion is shared by Flanagan as well who, as I have described, claims that the third-person perspective of science fails to capture a first-person phenomenology. Moreover, Flanagan's qualia realism would be in jeopardy if we collapsed the distinction (or accomplished a definitional reduction) between qualia and particular cognitive mechanisms. Consequently it would blur the line between Flanagan's neurophysicalism and Dennett's eliminativism.

Flanagan has written an article on zombies and conscious inessentialism, an article which casts further doubts on his physicalist account of mind. In the article (co-written with Thomas Polger), Flanagan argues that the existence of zombies "who are behaviourally indistinguishable from us appears

to be metaphysically, logically, and nomically possible" (Flanagan and Polger 1995, 314). Zombies are identically physically constituted, "behave just like we do, but are completely 'mindless' in the conscious sense" (Flanagan and Polger 1995, 313). This scenario, according to Flanagan, seems consistent with our current knowledge of evolution. Mother Nature and the "laws of nature as we know them in our vicinity allow that very intelligent, informationally-sensitive, but non-conscious creatures, could evolve. To the best of our knowledge, this is true" (Flanagan and Polger 1995, 317). He also writes, "Consciousness did not have to evolve. It is conceivable that evolutionary processes could have worked to build creatures as efficient and intelligent as we are, even more efficient and intelligent, without those creatures being subjects of experience" (Flanagan 1992, 129). Flanagan's and Polger's article seeks to refute Todd Moody's critique of conscious inessentialism, but it turns out that Flanagan's defense of zombies undermines the idea that the mental supervenes upon the physical, as it is conceivable that a physical fact can obtain without some higher level phenomena being realized. This seems to speak against neurophysicalism as the logical and metaphysical possibility of zombies undermines the physicalist strategy of linking antecedent physical causes with the emergence of consciousness. These zombies chew up, so to speak, the desired causal chain between the physical and mental domain. Flanagan's exploration of philosophical zombies further testifies to the non-derivability of the mental from the physical and to an epistemic gap between our knowledge of the physical and the reality of human mentality. The plausibility of a zombie scenario and, as we shall see, Flanagan's failure in closing the epistemic gap between the mental and the physical undermine the plausibility of a full naturalization of consciousness.

As we have seen, Flanagan employs a deflationary account of qualia and phenomenal properties, despite describing his view of mental phenomena as realistic. I suggest that the overall strategy of Flanagan is to reduce the hard problem of consciousness to an easy problem in order to render it epistemically possible for physicalism to make progress in the mind-body debate. The problem is that the easy problem is not so easy. Let me briefly reiterate the hard versus easy problem distinction before showing how this view can be challenged. When David Chalmers originally introduced this distinction, the "easy" problem referred to the idea of subjects discriminating, categorizing, and reacting to the environment. It further involves the deliberate control of behavior, difference between wakefulness and sleep, the reportability of mental states, and the integration of information by a cognitive system (Chalmers 2010, 3–34). The reality of these cognitive capacities is seen as an easy problem, according to Chalmers, because if is often assumed that the problems associated with such capacities can be settled by the standard methods of cognitive science and the identification of the relevant neural mechanisms. Although

this infamous distinction serves an important pedagogical purpose it seems, nevertheless, philosophically unstable. This is most clearly seen in Flanagan's project of reducing sensory qualia to neural mechanisms.

The idea, as formulated by Flanagan, is that sensory qualia are simply the product of a vast number "of patterned distributions of neural activity" (Flanagan 1991, 326). Different forms of qualia, or phenomenal states, correspond to different forms of activity in different kinds of receptors. Hence, taking into account the variety of physical receptors within human creatures, one can explain emotions, tastes, colors, perceptual experiences, and so on. This reductionist attempt, however, seems inherently problematic. Some perceptual experiences, such as seeing a ginger cat, involve phenomenal features. And, more importantly, it involves representational content (Lowe 1995, 267). There is "something like" to enjoy the experience of a ginger cat, and also "such an experience represents – or better, *presents* – our immediate physical environment as *being in some way*" (Lowe 1995, 267). Perceptual experiences are not just a matter of seeing this, it is a matter if seeing-as. Thus, how an object appears to us affects our individual conceptualization of the world, and thus the phenomenal and qualitative aspect of our visual experience. This thought is, of course, raised in Immanuel Kant's epistemology: "Thoughts without [sensible] content are empty, [sensible] intuitions are blind" (Lowe 1995, 267-68). This idea is also expressed through John Hick's theological appropriation of the philosophical system of Wittgenstein. Hick seeks to explain the variety of cultures through our conceptual interaction with the world: "These conceptual creations are the inner skeletons structuring the various forms of life, or ways of being human, that constitute the different cultures of the earth. And it is at this level, at which experience is pervaded, moulded, and coloured by human meanings, that I wish to maintain that all experience embodies concept-laden forms of interpretation" (Hick 2004, 142, my italics). To leave out the phenomenal aspect of perceptual experiences—as in the case of neurophysicalism—is to neglect an intrinsic ontological aspect of human experiencing and ways of interacting with the world. Unless Flanagan joins Dennett and abandons his realistic stance regarding qualia, such phenomenal features have to be addressed and explained within a physicalist framework. Here too Flanagan's physicalism turns into a form of mysterianism by leaving the ontology of qualia as an unexplainable brute fact. He either needs to close the epistemic gap between the reality of qualia and the physical workings of the brain or go in Dennett's eliminativist direction. To simply leave qualia unexplained does not seem to be a live option.

I am assuming in my critique that a physicalist is *required* to explain how particular higher level properties can be accommodated within an ontology that places emphasis on the physical; how a higher level property can also be a natural or physical property; that is, the kind of property that can be empirically investigated by a scientist. Flanagan writes the

following regarding this requirement: "In the end, we need to beware of the temptation to think that for physicalism to be true, the basic physical sciences must be able to *capture* all truths. This is stronger than requiring that physicalism be true; that is, it is stronger than requiring that everything that happens in physical" (Flanagan 1992, 101). Physicalism can very well be true "without being able to explain everything" (Flanagan 1992, 101). Hence, physicalism, Flanagan claims, is not undermined by phenomena (including subjective consciousness) that resist a physicalist explication. Such epistemological limitations are unsurprising and do not undermine "the view that what there is, and all there is, is physical stuff and its relations" (Flanagan 1992, 101). Am I simply setting an unfairly high standard for Flanagan's proposal? I think that there are good reasons for a physicalist to take the issue of explanation seriously. It is generally assumed, even by committed naturalists and physicalists, that physicalism is in the business of explanations, that it is an accommodation project (Horgan and Timmons 1993, 180-204) that faces significant location (Jackson 2000) or placement problems (Price 2011, 187-89). A physicalist who adopts a realist view of mental properties, and who argues that such properties are purely natural, is required to show how they relate to other natural properties. Given that Flanagan rules out elimination as the way to proceed, he needs to place/locate mental properties within the natural order. Frank Jackson suggests that physicalism, and metaphysical systems in general, continually face the location problem by seeking "comprehension in terms of a more or less limited number of ingredients," or anyway a smaller list than we started with" (Jackson 2000, 5). Serious metaphysics, as Jackson calls it, claims to be complete with regard to some subject matter and this "means that there are inevitably a host of putative feature of our world which we must either eliminate or locate" (Jackson 2000, 5). Physicalism, thus, is a classic example of serious metaphysics because it "claims that a complete account of what our world is like... can in principle be told in terms of a relatively small set of favoured particulars, properties, and relations, the 'physical ones'" (Jackson 2000, 6). Consequently, if one maintains that the mental is dependent on the physical (some relation of supervenience) then a full inventory of physical properties should entail the instantiation and capturability of the mental. This follows from the ontological commitment of physicalism. Flanagan's mysterianism with regard to human mentality is problematic because it casts doubt on the ontological coherency of physicalism. Indeed, if some phenomena are left as unexplainable brute facts, then a critical question arises: "why should we favor physicalism over non-physicalist ontologies?" If mental phenomena escape a physicalist accommodation, then why should we take mental properties to be purely physical? We have seen that Flanagan seeks to justify physicalism by appealing to a causal correlation between mental and physical events. Yet, correlation

does not answer the origination issue and is insufficient for grounding and explicating mental properties. A correlation account does not exclusively support physicalism. Rather it is fully compatible with non-physicalist explanations, such as an interactionist form of dualism. Interactionist dualism, while it affirms the ontological difference between the mental and the physical, rejects a strict ontological gap between the two spheres and it takes seriously the embodied character of human creatures (see Taliaferro 1994). Hence, interactionist dualism can comfortably accommodate the close connection between physical and mental events, showing that correlation alone will not establish a physicalist/naturalist understanding of consciousness. Given that Flanagan (1) adopts a realist view of consciousness, (2) rejects eliminativism, and (3) denies the strict deducibility of the mental from the physical (through his defense of zombies), he needs to provide some additional reason in order to convince his fellow peers to adopt a physicalist outlook on reality. This, then, is not to set an unfairly high explanatory standard for physicalism. Rather, Flanagan needs to demonstrate the superiority of a physicalist accommodation of consciousness over that of non-physicalist ones. Flanagan's mysterianist account of consciousness leaves open the possibility for non-physicalist construals of consciousness. So far, Flanagan has not provided sufficient positive reasons for thinking that consciousness can be accommodated within his neurophysicalism, and accounted for through the natural method.

This means that Flanagan, as a self-identifying physicalist, needs to engage in some traditional accommodation/location/placement work. A full naturalization (or physicalization) of consciousness is not optional, but an intrinsic part in establishing the desired normative mind science. Yet, by settling for a mere correlation account, by not isolating the specific physical properties subvening mental phenomena, by not explaining the relationship between mental and physical events (and by leaving the origination issue unaddressed), this project entails a problematic mysterianism. As we shall see below, another form of mysterianism comes to surface in Flanagan's grounding of normativity.

# Grounding Moral Values in Physicalism: Another Form of Mysterianism

Physicalism is a restrictive ontology. We have seen of how Flanagan employs a variety of strategies for retaining a realist view of consciousness. He wants to wear the badge of physicalism without having to join the eliminativists in denying the reality of human mentality. This ambitious project of placing consciousness within scientific reductionism is unsuccessful and entails a form of philosophical mysterianism. Hence, a crucial feature of Flanagan's normative mind science is on shaky ground.

We will now turn to the second core aspect of Flanagan's normative mind science, namely meaning and ethical normativity. In this context Flanagan is clearer on the negative implications of physicalism, suggesting that an ontology that fundamentally relies on the methodology of the natural sciences leaves no room for objective moral values—if such values are metaphysically interpreted. Rather than opting for a realist construal of moral properties, Flanagan takes ethical naturalism in the direction of pragmatism. A pragmatic interpretation of the ethical domain, argues Flanagan, does not entail any explicit ontological commitment. Ethical naturalism, conjoined with a pragmatist outlook on human practices, "implies no position of whether there really are, or are not, moral properties in the universe in the sense debated by moral realists, anti realists, and quasi realists" (Flanagan 1996, 120). The pragmatist judges specific moral visions, ethical stances, and settles moral disagreement, on the basis of their practical efficacy and it what ways they improve human living. This is a form of *perspectivism* because all attempts at ethical critique will be from within human practices and not "from some neutral, transcultural, or transcendental perspective" (Flanagan 1996, 133). As ethical critique is intrinsically perspectival, this view amounts to "some form of moderate relativism," a consequence which Flanagan welcomes (Flanagan 1996, 133). What we deem as the good life and essential for human flourishing are practice-dependent, and "all critique is immanent" (Flanagan 1996, 133).

Flanagan's ethical naturalism, and the science of eudaimonia, seeks to avoid a metaphysical realist view of moral values. It also seeks to go beyond sheer moral subjectivism. Values are not reducible to psychological desires, nor are they universal. Flanagan, however, gives the impression that some rights should be universally agreed on. He states that "everyone has a right to flourish," that "every human ought to be given equal chance to develop her talents and interests so as to live in a fulfilling and meaningful way," and given that "each human life has intrinsic worth," "we should work to make the conditions of living meaningfully universally available" (Flanagan 2007, 214, 201, 58). Such ethical statements are admirable, but unfortunately left unsupported. They are merely assumed within his overall argument for a naturalization of the ethical domain. Yet, if Flanagan wishes these ethical assertions not to come across as wishful thinking, he needs to demonstrate how such a universalist language can be squared with ethical perspectivism and pragmatism. In a similar way to the case of qualia, these universalist claims are left as brute facts.

Moreover, Flanagan believes that "morality resists theoretical unification" under any single perspective or principle. Given that moral consideration arises within natural and social environments in relation to a multitude of complex processes, we face an abundance of moral ambiguity. There is a constant tension between societal demands and what we individually desire for our lives. On one side we have concerns about "social stability, coordination, prevention of harms, and so on" (Flanagan 1996, 131). On another side we have intrapersonal concerns regarding "individual

flourishing, personal goodness, and the like" (Flanagan 1996, 131). How does Flanagan proceed to resolve this seeming tension, and how can he render consistent the universalist aim of his ethical system with individual desires and preferences? Flanagan does little in this regard, saying that it "seems best to leave the tension as is" (Flanagan 1996, 131). There is no point, he suggests, in trying to constructively solve this issue, but "more good comes from having to confront it [the tension] again and again" (Flanagan 1996, 131). However, in leaving this issue unresolved, another form of mysterianism emerges within Flanagan's naturalized eudaimonia.

Flanagan insists that morality is a naturalistic phenomenon that can be studied by way of scientific methodologies. Moral reasoning is not a spooky, or ontologically odd, feature of human nature, somehow separate from the rest of the natural order. Similarly to other cognitive capacities, the ability to reason in moral terms has arisen naturally through the course of evolution. Humans are modeled by "connectionists systems" and we are encouraged to learn and adopt norms that are socially endorsed. A person who embodies socially supported values will gain in self-esteem, and she will flourish because she aligns herself with the values of a larger community. Flanagan utilizes Paul Churchland's neural-network theory for understanding moral learning and development, how a person acquires moral knowledge. On this model, "moral capacities are instantiated as skills... by a complexly configured matrix of synaptic connections" (Flanagan 2007, 134). The acquisition of new moral knowledge is realized in terms of the connections and relationships between neural units. This process is ongoing, and a subject has to learn to recognize the correct behavior and attitude for different moral contexts. Given that moral knowing is a skill, which we engage in from early childhood, this is fundamentally a process of learning how. Flanagan argues that according to "moral network theory, there is a straightforward analogy between the way a submarine sonar device that needs to learn to distinguish rocks from mines might acquire the competence to do so and the way a human might acquire moral sensitivities and sensibilities" (Flanagan 1996, 124).

This theory is proposed by Flanagan to demonstrate how a physicalist can talk about objective moral knowledge, in what ways there can be progress within the moral domain, and how physicalism does not entail full-blown moral subjectivism (Flanagan 1996, 124). Flanagan's construal of moral truths, however, makes it tricky to understand what he means when he says that a physicalist account of moral learning "yields objective knowledge," given that such knowledge is pragmatically construed and dependent on the beliefs of individuals and larger communities (Flanagan 1996, 124). Moral knowledge is a matter of practical success. Nevertheless, Flanagan is aware that what constitutes practical success, and moral progress, differs from individuals to individuals and from communities to communities. Consequently, "the tension between impartial moral

demands and what conduces to individual flourishing is ubiquitous" (Flanagan 1996, 131). Flanagan, as we saw earlier, does not seek to address this tension, but takes the mysterianist exit instead. The problem with this metaphysical quietism is that it remains unclear to what extent moral knowledge on a pragmatic view amounts to something truly objective or whether it is a matter of mere socialization.

Given that moral truth is defined solely within human practices, a deflationary account of moral discourse seems to follow. Flanagan denies W. V. O. Quine's remark that a naturalist is committed to ethical coherentism, that "a coherency theory is evidently the lot of ethics" (Flanagan 1982, 70). Given that Flanagan opts for a practice-centered procedure for ethical evaluation, moral knowledge seems reducible to "mere socialization" and agreement between speakers. The pragmatic criterion of "what works" is too weak to lend the appropriate normativity for physicalism. Indeed, Flanagan seems to agree on this point, saying that "there is no remotely reliable guarantee that these 'workable ways' will be morally workable" (Flanagan 1996, 133). This is not to say that it is impossible to understand the nature of ethical practices on the basis of pragmatism. My critique is aimed more specifically at Flanagan's proposal and what seems to be a failure in synthesizing ethical pragmatism with a universalist and at many times objectivist moral language; what we get instead is a form of mysterianism. In this way, Flanagan does little to actually naturalize the ethical domain, but leaves this core feature of a normative mind science ungrounded.

Flanagan's resistance to moral realism springs from a desire to uphold a purely physicalist conception of the world. If it is the case that "all phenomena are natural and subject to causal principles," "made of natural stuff," and "explicable in terms of natural law," then such an ontology seems to rule out moral realism, in the sense of there being irreducible moral properties (Flanagan 1996, 56, 57). As suggested, physicalism is a restrictive ontology with a limited supply of ingredients. Flanagan, contrary to other physicalists, does not seek to retain moral realism through a supervenience account (that normative facts hold in virtue of non-normative facts) (Zangwill 1997), nor does he seek to demonstrate the realness of moral properties as emergent features of physical systems (Murphy and Ellis 1996). Flanagan is to his credit an honest physicalist that is willing to face the ontological implications for the moral realm. Nevertheless, Flanagan's project of formulating a scientifically coherent normative mind science that is consistent with the tenets of physicalism relies heavily on the possibility of grounding ethical normativity within physicalism. This project envisions an intimate connection between morality, meaning, and spirituality; "living morally is a condition for a meaningful life" and meaning is a basic ingredient for a robust spirituality (Flanagan 2007, 199). The problem, then, is that Flanagan employs normative vocabulary

without appropriately grounding such vocabulary within his ontology, which consequently results in a problematic mysterianism.

#### CONCLUSIONS: OPTIMISTIC AND PESSIMISTIC MYSTERIANISM

This article has critically evaluated Flanagan's normative mind science, and the attempt within this framework to create conceptual space for a normative mind science. It was stated earlier in this article that Flanagan, in order to successfully formulate a scientifically acceptable normative mind science, needs to (1) naturalize human consciousness, and (2) derive a normative account of human flourishing from neuroscientific depictions of human mentality.

A physicalist understanding of consciousness and ethical practices seems, from the look of Flanagan's normative mind science, to contain several philosophical problems. It is generally conceived that to do philosophy in a physicalist manner is to explain how certain higher level domains can be accounted for within an ontological framework which stresses the primacy of the physical. Flanagan, despite opting for physicalism, leaves higher level phenomena as brute facts. This can most clearly been seen with regard to consciousness, but also to some extent with respect to ethics and the issue of normativity. Flanagan's physicalism, despite positing a strong critique against philosophical non-constructivism, amounts itself to a mysterianist approach. Therefore, it becomes difficult to make sense of Flanagan's strong displeasure with McGinn's agnostic position. On McGinn's view, there is an explanation for how consciousness emerged from physical happenings, although this explanation lies beyond human comprehension. It is because of our cognitive limitations that the mind-body problem is unsolvable. He writes, "Minds are biological products like bodies, and like bodies they come in different shapes and sizes, more or less capacious, more or less suited to certain cognitive tasks" (McGinn 1989, 350). Humans are not cognitively suited for solving the enigma of consciousness. This agnostic naturalism, however, does not exclude the possibility that a more intelligent nonhuman creature (perhaps a very intelligent alien civilization) or a supercomputer could explain the emergence and nature of mind. Ironically, McGinn's mysterianism comes across as more optimistic when compared to Flanagan's physicalism. Science can map out correlations between mental and physical events but is, according to Flanagan, unable to explain why it is that some particular patterns of neural activity give rise to phenomenal experience while other neurological events do not. The natural method can only take us so far, and the story bottoms out there. This is a pessimistic mysterianism which leaves the nature of consciousness, and ethical normativity, unexplained and not properly grounded in physicalism. Consequently, a reader of Flanagan is left doubtful that physicalism really is a suitable ontology for providing a philosophically coherent vision of the good life.

### References

- Campbell, Neil. 2009. "Why We Should Lower Our Expectations about the Explanatory Gap." Theoria 75:34–51.
- Chalmers, David. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. New York, NY: Oxford University Press.
- Flanagan, Owen. 1982. "Quinean Ethics." Ethics 93:56-74.
- . 1991. *The Science of Mind*. Cambridge, MA: MIT Press.
- ——. 1992. Consciousness Reconsidered. Cambridge, MA: MIT Press.
- ——. 1996. Self-Expressions: Mind, Morals, and the Meaning of Life. New York, NY: Oxford University Press.
- ——. 2007. The Really Hard Problem: Meaning in a Material World. Cambridge, MA: MIT Press.
- Flanagan, Owen, and Thomas Polger. 1995. "Zombies and the Function of Consciousness." Journal of Consciousness Studies 4:313–21.
- Hasker, William. 1999. The Emergent Self. Ithaca, NY: Cornell University Press.
- Hick, John. 2004. An Interpretation of Religion: Human Responses to the Transcendent. New Haven, CT: Yale University Press.
- Horgan, Terence. 1993. "From Supervenience to Superdupervenience; Meeting the Demands of a Material World." *Mind* 102:555–86.
- Horgan, Terence, and Mark Timmons. 1993. "Metaphysical Naturalism, Semantic Normativity, and Meta-Semantic Irrealism." *Philosophical Issues* 4:180–204.
- Jackson, Frank. 2000. From Metaphysics to Ethics: A Defence of Conceptual Analysis. New York, NY: Oxford University Press.
- Kim, Jaegwon. 1988. "Explanatory Realism, Causal Realism, and Explanatory Exclusion." Midwest Studies in Philosophy 12:225–39.
- . 1993. Supervenience and Mind: Selected Philosophical Essays. New York, NY: Cambridge University Press.
- 2007. "Reduction and Reductive Explanation: Is One Possible without the Other?" In Being Reduced: New Essays on Causation and Explanation in the Special Sciences, edited by Jakob Hohwy and Jesper Kallestrup, 93–114. Oxford, UK: Oxford University Press.
- Libet, Benjamin. 1985. "Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action." *Behavioural and Brain Sciences* 8:529–66.
- Lowe, E. J. 1995. "There Are No Easy Problems of Consciousness." *Journal of Consciousness Studies* 2:266–71.
- McGinn, Colin. 1989. "Can We Solve the Mind-Body Problem?" Mind 98:349-66.
- Murphy, Nancey, and George Ellis. 1996. On the Moral Nature of the Universe: Theology, Cosmology, and Ethics. Minneapolis, MN: Fortress Press.
- Nagel, Thomas. 1974. "What Is It Like to Be a Bat?" Philosophical Review 83:435-50.
- . 2012. Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature Is Almost Certainly False. Oxford, UK: Oxford University Press.
- Price, Huw. 2011. Naturalism without Mirrors. New York, NY: Oxford University Press.
- Rowlands, Mark. 2001. The Nature of Consciousness. New York, NY: Cambridge University Press. Seager, William. 1999. Theories of Consciousness: An Introduction and Assessment. New York, NY:
- Seager, William. 1999. Theories of Consciousness: An Introduction and Assessment. New York, NY:
  Routledge.
  Searle, John. 2008. "Reductionism and the Irreducibility of Consciousness." In Emergence:
- Searle, John. 2008. "Reductionism and the Irreducibility of Consciousness." In Emergence: Contemporary Readings in Philosophy and Science, edited by Mark A. Bedau and Paul Humphreys, 69–80. Cambridge, MA: MIT Press.
- Sinnott-Armstrong, Walter, and Lynn Nadel. 2010. Conscious Will and Responsibility: A Tribute to Benjamin Libet. New York, NY: Oxford University Press.
- Taliaferro, Charles. 1994. Consciousness and the Mind of God. New York, NY: Cambridge University Press.
- Zangwill, Nick. 1997. "Explaining Supervenience: Moral and Mental." *Journal of Philosophical Research* 22:509–18.