

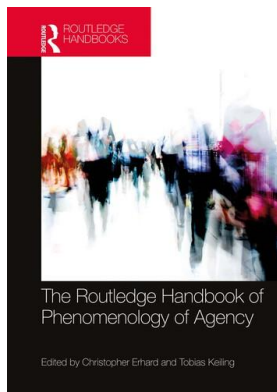
This article was downloaded by: 10.3.97.143

On: 20 Mar 2023

Access details: *subscription number*

Publisher: *Routledge*

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: 5 Howick Place, London SW1P 1WG, UK



The Routledge Handbook of Phenomenology of Agency

Christopher Erhard, Tobias Keiling

Will-power

Publication details

<https://www.routledgehandbooks.com/doi/10.4324/9781315104249-25>

Robert Hanna

Published online on: 30 Oct 2020

How to cite :- Robert Hanna. 30 Oct 2020, *Will-power from:* The Routledge Handbook of Phenomenology of Agency Routledge

Accessed on: 20 Mar 2023

<https://www.routledgehandbooks.com/doi/10.4324/9781315104249-25>

PLEASE SCROLL DOWN FOR DOCUMENT

Full terms and conditions of use: <https://www.routledgehandbooks.com/legal-notices/terms>

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

21

WILL-POWER

Essentially embodied agentive phenomenology, by way of O'Shaughnessy

Robert Hanna

In the beginning was the Act.

(Goethe 1990: 153)

Introduction

Goethe got it right. While I was a graduate student at Yale in the mid-to-late 1980s, during that heady, wonderful period between finishing coursework and starting actual PhD-dissertation-writing labors, I read three big, brilliant books that significantly influenced my later work on cognition, mental content, knowledge, logic, rationality, the mind-body relation, mental causation, intentional action, free will, practical agency, and persons:

Gareth Evans's *Varieties of Reference* (Evans 1982),
Brian O'Shaughnessy's *The Will* (O'Shaughnessy 1980), and
Richard Wollheim's *The Thread of Life* (Wollheim 1984).

At the same time, I was also steeped in Kant's Critical philosophy; in existential phenomenology (especially the later Husserl, Scheler, Heidegger, early Sartre, and Merleau-Ponty); in Wittgenstein's *Investigations* (Wittgenstein 1953); in Harry Frankfurt's essays, many of them collected in *The Importance of What We Care About* (Frankfurt 1988); and in Richard Rorty's radical metaphilosophy in *Philosophy and the Mirror of Nature* (Rorty 1979) and *Consequences of Pragmatism* (Rorty 1982). Those were the days of philosophical wine and roses.

A Recipe for Long-Term Philosophical Agency about Rational Human Minded Animal Agency: Throw all the above-listed philosophical ingredients, including the wine and roses, into a big pot, add water till it reaches the brim, and heat slowly to an intellectual boil, stirring occasionally, for twenty years. And then, as the novelist John Cheever remarked in a closely-related context, "you apply butt to chair." In the beginning, and also at the end, is the Act.

Following that recipe step by step, I eventually wrote *Rationality and Logic* (Hanna 2006), *Embodied Minds in Action* (co-authored with Michelle Maiese, Hanna and Maiese 2009), *Cognition, Content and the A Priori* (Hanna 2015), and *Deep Freedom and Real Persons* (Hanna 2018a). The present chapter/essay is a rational reconstruction of some of the results of all that long-term philosophical agency about rational human minded animal agency, concentrating especially on three issues. First, I briefly describe the causally efficacious power of the minded will (or what Kant calls *Willkür*, aka “the power of choice”) in basic intentional actions, that is, *intentional body movements*, against the metaphysical backdrops of my *essential embodiment* theory of the mind-body relation and my metaphysics of free agency, *Natural Libertarianism*. Second, I describe some relevant comparisons and contrasts between my views about essentially embodied intentional action and O’Shaughnessy’s *dual aspect* theory of intentional action, about which he very provocatively says:

Schopenhauer said: “I cannot imagine this will without this body.” If this wonderful remark is right – and . . . it is – then the psychological phenomena that occur when one engages in intentional physical action . . . depend in general upon the *actual existence* of the body. . . . Then it follows that [my position] cannot be an interiorist position, even though I believe that trying is a psychological event that causally explains the purely physical willed bodily event in the case of all bodily actions. For what here has come to light is, the essential mutual interdependence of bodily strivings and of the animal body upon each other.

(O’Shaughnessy 2008: vol. 2, 600)

And third and finally, I argue for the existence of a *self-evident, veridical phenomenology* of essentially embodied free agency, aka rational human minded animal agency.

Essentially Embodied Trying and Natural Libertarianism

Here is a thesis about the nature of basic intentional action:

A is a basic intentional act of an essentially embodied rational human minded animal, aka a real human person *P*, if and only if *A* is an intentional body movement of *P* that is structurally caused and actively guided and controlled by *P*’s simultaneously *trying* to perform *A*, which in turn is a physically irreducible conscious effective first-order desire to perform *A*, which in turn is *P*’s *will*.

For convenience, I will call this thesis *Essentially Embodied Trying*.

One crucial thing to notice about the Essentially Embodied Trying thesis is that what I am calling “intentional bodily movements” are *essentially embodied events in a minded animal’s life*, and not anything exclusively or fundamentally physical. This in turn directly implies what, in *Embodied Minds in Action* (Hanna and Maiese 2009), Michelle Maiese and I call *The Essential Embodiment Theory* of the mind-body relation and mental causation. Here are the six central theses of that theory:

- 1 *The Essential Embodiment Thesis*: Creatures with conscious, intentional minds are necessarily and completely neurobiologically embodied.

- 2 *The Essentially Embodied Agency Thesis*: Basic acts (e.g., raising one’s arm) are intentional body movements caused by an essentially embodied mind’s synchronous trying to make those very movements and its active guidance of them.
- 3 *The Emotive Causation Thesis*: Trying and its active guidance, as the cause of basic intentional actions, is primarily a pre-reflective, desire-based emotive mental activity and only derivatively a self-conscious or self-reflective, deliberative intellectual mental activity.
- 4 *The Mind-Body Animalism Thesis*: The fundamental mental properties of conscious, intentional minds are (a) non-logically or strongly metaphysically (namely, synthetically) a priori necessarily reciprocally intrinsically connected to corresponding fundamental physical properties in a living animal’s body (mental-physical property fusion), and (b) irreducible truly global or inherently dominating intrinsic structures of motile, suitably neurobiologically complex, egocentrically centered and spatially oriented, thermodynamically irreversible living organisms (neo-Aristotelian hylomorphism).
- 5 *The Dynamic Emergence Thesis*: The natural world itself is neither fundamentally physical nor fundamentally mental, but is instead essentially a causal-dynamic totality of forces, processes, and patterned movements and changes in real space and real time, all of which exemplify fundamental physical properties (e.g., molecular, atomic, and quantum properties). Some but not all of those physical events also exemplify irreducible biological properties (e.g., being a living organism), and some but not all of those biological events also exemplify irreducible fundamental mental properties (e.g., consciousness or intentionality). And both biological properties and fundamental mental properties are dynamically emergent properties of those events.
- 6 *The Intentional Causation Thesis*: A mental cause is an event or process involving both consciousness and intentionality, such that it is a necessary proper part of a nomologically jointly sufficient essentially mental-and-physical cause of intentional body movements. In so doing, it is a dynamically emergent structuring cause of those movements. Then, under the appropriate endogenous and exogenous conditions, by virtue of synchronous trying and its active guidance, conscious, intentional essentially embodied minds are mental causes of basic acts from their inception in neurobiological processes to their completion in overt intentional body movements.

Notice especially that according to this conception, a *structuring cause*

- i is *not a prior event-cause*, because it is simultaneous and synchronous with the effect, as a complex dynamic pattern spread out over time,
- ii is *equally formal and final in the Aristotelian senses* and *not (merely) efficient or material in the Aristotelian senses*, in that
 - ii a it actualizes potential energy, and it is inherently goal-directed or at the very least forward-directed in time,
 - ii b it thereby activates, controls, and guides causally efficacious energy flows, and
 - ii c it is itself genuinely causally *efficacious* (as opposed to merely causally *relevant*) in an independently necessary, independently insufficient, but also jointly sufficient way, *via* those causally efficacious energy flows, but
 - ii d it is *not* reductively identical with – in the “downward-identity” sense of being “nothing but,” or “nothing over and above” – causally efficacious energy flows themselves, and

- iii is a non-reducible proper part of a whole causally efficacious complex dynamic organismic process, that is, a causally efficacious non-reversible, non-equilibrium self-organizing organismic energy flow.

In this way, The Essential Embodiment Theory says that our dynamically emergent, irreducible, sentient, and sapient minds are also necessarily interdependent with our own living organismic animal bodies and not essentially distinct from them; that we are far-from-equilibrium, asymmetric, complex, self-organizing thermodynamic systems; that we act by intentionally moving our bodies by means of our desire-based emotions and trying; and that our conscious, intentional, caring, and rational necessarily and completely neurobiologically embodied minds are basically causally efficacious precisely because they are metaphysically continuous with our biological lives, and life is basically causally efficacious in physical nature.

The simple upshots of The Essential Embodiment Theory, then, are:

- i in thinking about the mind-body problem we should decisively replace the early modern Cartesian and Newtonian *ghost-in-the-machine* metaphysics with a post-Cartesian and post-Newtonian but also at the same time neo-Aristotelian hylomorphic *immanent-structure-in-the-non-equilibrium-thermodynamics* metaphysics, and
- ii the irreducible conscious, intentional, caring minds of cognizers and agents grow naturally in suitably complex living organisms, as irreducible, non-dualistic, non-supervenient, asymmetric thermodynamic structures of those organisms.

Correspondingly, the three crucial points in the Essentially Embodied Trying thesis are:

- i that the trying is itself essentially embodied,
- ii that the trying is simultaneous and synchronous with the action, and not a prior event-cause of it, and
- iii that the causation characteristic of trying is inherently structural-dynamic and goal-directed, that is, it is genuinely efficacious neo-Aristotelian hylomorphic structural causation, fusing together “formal” causation and “final” causation in the classical Aristotelian senses.

To put the Essentially Embodied Trying thesis in a name-dropping context, however, it is a constructive extension of Harry Frankfurt’s devastating *critique of classical causal theories of action*, his *hierarchical desire* conception of the will and personhood, and his *guidance-control* conception of intentional action (Frankfurt 1988a, 1988b, 1988c, 1988d, 1988e, 1988f), together with O’Shaughnessy’s action-theoretic notion of *trying* (O’Shaughnessy 1973), when framed against the dual backdrop of my theory of logically driven cognitive rationality in *Rationality and Logic* (Hanna 2006), and The Essential Embodiment Theory of the mind-body relation and mental causation in *Embodied Minds in Action*.

But in order to make all of this more clear and distinct, I will also have to provide a brief description of a background metaphysical theory I defend: *Natural Libertarianism*.

What is free will? What is practical agency? What is human personhood? And how are human free will, practical agency, and human personhood really possible in the natural world as it is correctly characterized by the modern natural sciences, especially physics, chemistry, biology, and cognitive neuroscience? Or more compactly put: given the truth of modern science, how is human free agency really possible? Let us call this *the freedom question*.

In *Deep Freedom and Real Persons* (Hanna 2018a), I provide what I think is a rationally decisive and true answer to the freedom question.

But in order to formulate this answer properly, I will need some precisely defined terminology. I will say that *free will* is a conscious subject's power to choose or do what she wants to, or to refrain from so choosing or so doing, without preventative constraints and without internal or external compulsion, with at least causal responsibility; that *practical agency* is a conscious subject's power to choose or do things freely in the light of principles or reasons, including *moral* principles or reasons, on the basis of self-conscious processes of deliberation and decision; and that a *real person* is an animal that is capable of free will, practical agency, and (moral or non-moral) responsibility. So the freedom question is asking how human free will, human practical agency, and real human personhood *really exist*, and how we *truly have them*, in the natural world as correctly described by modern science.

Or, on the contrary, given the truth of modern science, are we really nothing but “biochemical puppets” (Harris 2012) or “moist robots” (Daniel Dennett, as quoted in Schuessler 2013), that is, nothing but natural automata, or natural machines, whose evolutionary and neurobiological mechanisms continually generate the cognitive illusion that we are free agents? If so, then we would be in an even worse cognitive place than Pinocchio, a wooden puppet who longed to be a real boy. We would be nothing but “meat puppets” (Meat Puppets 1994) dreaming that we are real human persons. The issue being raised here, then, is how we should understand the implications of the modern natural sciences for our classical conception of ourselves as *rational and moral animals*, in the face of the possibility that we are under a serious cognitive illusion about this. Indeed, some contemporary philosophers even think that once we are liberated from this serious cognitive illusion, we will finally clearly see that we are *nothing but* highly complex “biochemical puppets” and that “physics makes us free” in a deterministic, block universe (Ismael 2016).

By sharp contrast to these philosophers, like Kant, I hold that we *directly experience* practical agency, or what he calls “practical freedom,” and also that practical agency requires metaphysically robust human free will or *deep freedom* – what he calls “transcendental freedom” – which really exists in the natural world:

Practical freedom can be proved through experience. . . . We thus cognize practical freedom through experience as one of the natural causes, namely a causality of reason in the determination of the will, whereas transcendental freedom requires an independence of this reason itself (with regard to its causality in initiating a series of appearances) from all determining causes of the world, and to this extent seems to be contrary to the law of nature, thus to all possible experience, and so remains a problem.

(Kant 1997: A 802–803/B 830–831, 675–676)

I also fully agree with Kant that this kind of freedom “seems to be contrary to the law of nature, thus to all possible experience, and so remains a problem.” In Kantian terminology then, in *Deep Freedom and Real Persons* (Hanna 2018a) I am trying to solve Kant's problem by developing a metaphysics *according to which* “transcendental freedom” and “practical freedom” or “autonomy” are themselves fully natural facts. But in my own terminology, I am trying to solve Kant's problem by developing a metaphysics *according to which deep freedom and practical agency are themselves fully natural facts*.

On the view I propose, our double capacity for free will and practical agency – which, for convenience, I will call *free agency* – is an irreducible fact. At the same time, our irreducible capacity for free agency does *not* exist over and above the rest of the physical world – it

is categorically *not* a mysterious dualistic, extra-physical fact. On the contrary, it is a fully natural, biological, and neurobiological fact – a natural fact of life. So the key to our free agency is not that we possess mysterious, non-natural, atemporal causal powers to choose or act in violation of the causal laws of nature. On the contrary, it is simply that, insofar as we are minded living organisms,

- i *all* of our intentional activities – by which I mean the things that we ourselves do, and do not merely happen to us – are inherently vital and non-mechanical,
- ii *some* of our intentional activities are naturally creative and authentically original, in just the way that a work created by a “human, all too human” artistic genius is authentically original, but not in any god-like or magical way, and
- iii *none* of the general causal laws of nature, especially including the Conservation Laws, is ever violated by us.

These three natural facts, namely,

- i* vital, non-mechanical sourcehood,
- ii* natural creativity, and
- iii* living in bounded natural open space (having non-equilibrium-thermodynamic “degrees of freedom”), consistently with *all* of the general causal laws of nature, especially including the Conservation Laws,

are all strongly supported by the self-evident, *veridical phenomenology* of essentially embodied free agency that I will spell out and defend in the final section of this chapter/essay.

O’Shaughnessy and Me

I am more than happy to disclose the elective affinities and relations of influence between O’Shaughnessy’s work and mine, because I do think that his work is both exceptionally brilliant and also, unjustly, much-neglected.

Moreover, it is clear from the following passage (and many others) in the second, 2008 edition of *The Will* that our theories of willing, trying, and intentional action are very much in philosophical harmony with each other:

[According to the dual aspect theory of physical action] the event of acting and/or willing is one and the same as the full activation of the motor system. It should be noted that this is not that the event of *willing* activates the motor system. Rather, it is that the immediate *act-progenitors*, act-desire and act-intention, activate that system, and that this activation-syndrome falls simultaneously under “try,” “strive,” “do,” “will,” and finally by the end of the motor-process also under “arm-raising” or “walking” or “talking” (as the case may be).

(O’Shaughnessy 2008: vol. 2, 511)

Nevertheless, without actually having painstakingly compared the first, 1980 edition of *The Will* with the second, 2008 edition, line by line, I am also strongly of the impression that many passages in the first edition, as well as certain passages in his 1973 paper “Trying (as the Mental ‘Pineal Gland’),” suggest that between 1970 and 1980, O’Shaughnessy at times

slipped back into a classical *causal theory of action*, of the kind (to my mind) so devastatingly criticized in by Frankfurt (1988c). As Frankfurt crisply puts it:

The problem of action is to explicate the contrast between what an agent does and what merely happens to him, or between the bodily movements that he makes and those that occur without his making them. According to causal theories of action, which currently represent the most widely followed approach to the understanding of this contrast, the essential difference between the two types is to be found in their prior causal histories: a bodily movement is an action if and only if results from antecedents of a certain kind.

...
 Despite its popularity, I believe that the causal approach is inherently implausible and that it cannot provide a satisfactory analysis of the nature of action. I do not mean to suggest that actions have no causes; they are as likely to have causes, I suppose, as other events are. My claim is rather that it is no part of the nature of an action to have a prior causal history of any particular kind. . . . It is integral to the causal approach to regard actions and mere happenings as being differentiated by nothing that exists or that is going on at the time those events occur, but by something quite extrinsic to them – a difference at an earlier time among another set of events entirely. . . . It is no wonder that [causal] theories characteristically run up against counterexamples of a well-known type. . . . No matter what kinds of causal antecedents are designated as necessary and sufficient for the occurrence of an action, it is easy to show that causal antecedents of that kind may have as their effect an event that is manifestly not an action but a mere bodily movement. . . . In my judgment causal theories are unavoidably vulnerable to such counterexamples, because they locate the distinctively essential features of an action exclusively in states of affairs which may be past by the time the action occurs. This makes it impossible from them to give any account whatever of the most salient differentiating characteristic of action: during the time a person is performing an action, he is necessarily in touch with the movements of his body in a certain way, whereas he is necessarily not in touch with them in that way when movements of his body are occurring without his making them.

(Frankfurt 1988c: 69–71)

Nevertheless, by 2008 or thereabouts, the later O’Shaughnessy and I are completely in agreement that

- i the spontaneous, desire-driven, telic act of the will that consists in *trying*, and
- ii the event that consists in *causally efficaciously bringing about intentional body movements* are not only *synchronous* events, but also – and here is where O’Shaughnessy and I both go one big metaphysical step beyond Frankfurt – *ontologically fused* events that are, necessarily, complementary proper parts of the rational human minded animal agent’s life over that stretch of time.

It is trying that efficaciously causes the intentional body movement, but not *as such* or *per se*, rather only as ontologically fused with a larger causally efficacious dynamic process, because the conscious event of trying constitutes only a proper structural part of the whole dynamic process. So the metaphysical picture is one of individually necessary and individually insufficient but *jointly sufficient synchronous and ontologically fused event-causation*. Similarly a driver

causally efficaciously drives a car, but nor *as such* or *per se*, rather only as a synchronous proper part of the whole dynamic process involving the causally efficacious car and the causally efficacious driver. The driver is causally efficacious in a different way from the car, insofar as he synchronously autonomously controls the car by means of his own intentional body movements, but he is not literally moving the car along the road by means of those movements – unless, of course, he’s Fred Flintstone.

This is not to say, however, that O’Shaughnessy ever *officially* held something equivalent to The Essential Embodiment Theory of the mind-body relation:

[The dual aspect theory advanced in *The Will*] is a dual aspect theory of bodily action, and nothing else. On the larger issues of materialism, dualism, and dual aspectism generally, I am intentionally silent. . . .To be sure, the specific conclusions reached concerning bodily action accord better with a general dual aspectism than with Cartesian dualism. But that is about all I have to say on these more general issues.

(O’Shaughnessy 2008: vol. 1, 44)

But at the same time, he also comes *really, really close* to The Essential Embodiment Theory, even lower down on the same page and onto the next:

I believe we shall find in [the question of the constitution of bodily action] a key to the mind-body problem: a key through the agency of bodily action. . . . In my opinion, it enables us to put to philosophical use the fact that mind and body come closest together, and in a wholly overt mode, in the phenomenon of physical action. . . . These two ontological domains are incomparably wedded in the phenomenon of physical action. Through this phenomenon we encounter a partial resolution of a major philosophical problem: a resolution through the agency of the very special character of physical action.

(O’Shaughnessy 2008: vol. 1, 44–45)

One important domestic *difference* between O’Shaughnessy’s theory of bodily action and mine, however, concerns what he calls “mere bodily events” like breathing, laughing, and blinking – as opposed to fidgeting, humming a tune to oneself, or moving one’s tongue in one’s mouth while driving one’s car (see, e.g., O’Shaughnessy 2008: vol. 1, 37, 40, 42; vol. 2, 356–360, 463–469). Contrary to most other action-theorists, O’Shaughnessy and I do agree that fidgeting, humming, or moving one’s tongue in one’s mouth while driving *are all intentional actions*. But at the same time O’Shaughnessy holds that breathing, laughing, and blinking are all passive and non-intentional; whereas I do not.

Why not? One reason is that I think that everything that is genuinely mental, no matter how “obscure,” and no matter how supposedly “sub-personal” or “unconscious,” is at *least minimally phenomenally and pre-reflectively occurrently conscious*. This is what I call *The Deep Consciousness Thesis*, aka The DCT (Hanna and Maiese 2009: chs. 1–2). In *Embodied Minds in Action* (Hanna and Maiese 2009), and again more recently in *Cognition, Content, and the A Priori* (Hanna 2015: 97–105), I have argued for The DCT in three ways:

- i via a critique of the very idea of the “unconscious,” as inherently confused between, on the one hand, different degrees and modes of *pre-reflective* consciousness, e.g., “absorbed” or “absent-minded”/spaced-out consciousness while driving one’s car, and, on the other hand, underlying neurobiological processes (this rejection of the very idea of

- the “unconscious” as confused and bogus is of course similar to Sartre’s critique of the Freudian unconscious [Sartre 1966: 90–96]),
- ii via the *phenomenology* of pre-reflective consciousness and other “obscure” modes of consciousness, and
 - iii by pointing out, by way of a *reductio* argument, that if there are “sub-personal” or “unconscious” mental acts, states, or processes that are truly *not* conscious, then much or even most of our mental life and mental activity is merely mechanical and robotic, hence nothing that has basic normative value, (moral or non-moral) responsibility, any sort of real guidance or control, etc.

The DCT, in turn, radically strengthens my action theory to comprehend anything that flows from the agent herself, as an *ultimate source* (*not* however in a classical *agent-causal* sense, but instead *only* in the natural-libertarian sense of what I call *deep freedom* – see the final section below), including absorbed or absent-minded/spaced-out actions like driving one’s car, but also sleeping (whether in dream or dreamlessly), forgetting, breathing, laughing, blinking, fidgeting, humming a tune to oneself, and moving one’s tongue in one’s mouth while driving one’s car. Of course, events that are mentally *caused* by the minded animal can also *fail* to flow from the agent herself as an ultimate source, as in mad-scientist-style manipulation cases, pathological mental causation while suffering from a mental illness or a pathological condition like Obsessive-Compulsive Disorder, aka OCD, and so on (again, what is missing here is *not* classical agent-causation, but instead *only* deep freedom in the natural-libertarian sense – again, see the final section below, and for its specific application to the case of OCD, see also the final section below). But I do also think that all of these too are at least minimally phenomenally and pre-reflectively occurrently conscious.

Moreover, I think of all mental states, acts, or processes, especially including *desires*, not only as at least minimally phenomenally and pre-reflectively occurrently conscious but also in *conative* terms: they are also all inherently active, intentionally directed, telic, and at least *poised for causing actions*. Where mentality is concerned, then, it is *through-and-through active or poised for action*; hence it is all *spontaneous* to some extent (Hanna and Thompson 2005). At the same time, there are also different modes of spontaneity, in that some mental acts, states, or processes are inherently somewhat less self-initiated and somewhat more “primed” or “triggered” than others, in the sense that they are the result of *vital responsiveness and sensible tracking* – e.g., the conscious experience of bodily pain, or what I call *nociperception* (see Hanna 2018b: ch. 4). Nevertheless, *none* of the so-called “passive” mental acts, states, or processes, not only including absorbed or absent-minded/spaced-out actions like driving one’s car, but also including nociperception, sleeping (whether in dream or dreamlessly), forgetting, breathing, laughing, blinking, fidgeting, humming a tune to oneself, and moving one’s tongue in one’s mouth while driving one’s car, as well as intellect-infused activities like cognitive insight or discovery, aesthetic experiences, respect, and reverence, are in fact “passive” at all, but on the contrary, instead, they all manifest one or another kind of *responsive, sensible spontaneity* (cf. Schopenhauer 1907: §42, 169).

But whether a *first-order desire* – say, wanting to drink a beer now – becomes *effective* and thus becomes the agent’s *will* at that time, or over time, thus moving the agent all the way to action, is a function of the desire’s position in the overall desire-hierarchy and the overall dynamic desire-structure constituting the agent’s will: as, e.g., when the agent also *has the second-order desire* of wanting to want to drink a beer now, and is *thereby* moved all the way to beer-drinking. The dynamic desire-structure, in turn, is the basic source of agency; and in this way, mental causation is *agentive will-structuring*. So my action theory is *neo-hylomorphic*,

deploying an updated *Aristotelian* metaphysics of form-and-matter, or structure-and-stuffing, as indeed is my theory of the mind-body relation: the mind is the activating form of the living animal body and its far-from-equilibrium, complex, self-organizing thermodynamic processes; and, in turn, agency is a certain configuration (really: a designated set of dynamic configurations/patterns, in orientable space and over elapsed time, making up a real person's life) of the minded animal's desires. Thus, for me, the will is an immanent structuring cause of the complex dynamic process constituting the minded animal's causally efficacious bodily life, via its organization of essentially embodied desires, and not, as in classical causal theories of action, a mental event temporally prior to and ontologically distinct from bodily movement.¹

Crucial to this overall account is what Maiese and I, in *Embodied Minds in Action* (Hanna and Maiese 2009: chs. 3–5 and section 8.3), riffing on the fundamental O'Shaughnessy-inspired notion of *trying*, call “effortless trying.” Effortless trying is pre-reflective and conscious, but *non-self-consciously* conscious, effective desiring. Hence, effortless trying is a mode of trying that synchronously guides and controls our causally efficaciously bodily life – say, various skillful body movements during dancing, playing sports, walking, or driving one's car, but also breathing, laughing, and so on – but which does *not* require self-conscious desires, or reflective self-initiation, rather only pre-reflectively conscious desires and desire-structuring. In turn, moreover, effortless trying is presupposed by all self-conscious or reflectively self-initiated *effortful* trying. So the (everyday, normal) essentially embodied basic act of, e.g., willing to type a word on my laptop computer is an episode of effortless trying, because it is a pre-reflectively conscious synchronous mental immanent structuring cause of the specific complex dynamic process constituting the minded animal's causally efficacious hand and arm movements, via its organization of essentially embodied desires; and the basic act itself in all its ontological glory is the complete dual aspect effortless-trying-event that synchronously structurally brings about the non-basic self-conscious, reflectively self-initiated act of creating a meaningful sequence of words on the laptop screen.

As I have indicated, I accept Frankfurt's basic picture of the hierarchical desire-structure of the will – but not his somewhat *intellectualist* picture of agentic phenomenology:

Now what conceptualization of [agentic phenomenology] fits its contours in the most authentic and perspicuous way? My own preference has been for a model that involves levels of reflexivity or self-consciousness.

(Frankfurt 1988f: 164, *underlining added*)

This seems somewhat intellectualist to me in its appeal to *self-consciousness*, which I construe in a broadly Kantian way as “apperception”: that is, a capacity for making higher-order psychological reports about lower-order conscious states. Now all such reports are judgments, and judgments are essentially *conceptual* activities. But for me, willing is fundamentally *essentially non-conceptual, pre-reflective* structural control of how my desires are integrated into my essentially embodied conscious life. Otherwise put, willing is not top-down mental impact from a separate highest-level executive self, but instead simply getting all of my minded bodily resources sufficiently structurally organized and together at any given time. Mental causation is *structuring causation carried out by the whole minded animal*. Or more idiomatically put: willing is simply *getting your essentially embodied, desiderative sh** together*, and thereby actually choosing or doing something. So a given second-order desire that fixes precisely which first-order desire is the effective one, at a certain time, is just a certain identifiable desiderative position (I want to want X, I want to want Y, etc.) in the overall will-structure

of the whole minded animal, indexed to a certain act-context, not some sort of little homuncular atomic willing self that initiates everything. In the final section, I will present a detailed picture of the full set of conditions that need to be satisfied in order for an agent to “get . . . all of [her] minded bodily resources sufficiently structurally organized and together at any given time.”

So much for the important domestic differences between O’Shaughnessy’s views and mine. Another way in which O’Shaughnessy’s view and mine importantly *converge* is with respect to what he calls “the long-term body image” (O’Shaughnessy 2008: vol. 1, 273–298). O’Shaughnessy’s long-term body image, in turn, is equivalent to what Maiese and I, following Shaun Gallagher, call *the body-schema* (Hanna and Maiese 2009: 69–73; see also Gallagher 2005: 37–38). According to us, the body-schema is the self-evidently veridical, essentially non-conceptual content of *primitive bodily awareness*; and according to me, it also captures the core of *the veridical phenomenology of essentially embodied free agency*, to which I now finally turn.

The Veridical Phenomenology of Essentially Embodied Free Agency

Many theories of free will – including Natural Libertarianism – hold that necessarily, I can freely choose or do *X* *only if* my choosing or doing *X*,

- i lacks an antecedent nomologically sufficient cause, and
- ii I myself am the causally sufficient ground, origin, or source of choosing or doing *X*.

The conjunction of these two necessary conditions is what Kant rather misleadingly calls “transcendental freedom,” and what Robert Kane more aptly calls “ultimate responsibility” (Kane 1996: 35). But as I see it, these two necessary conditions provide only *part* of a single necessary but individually insufficient condition – namely, what I call *the real causal spontaneity condition* – for deep freedom, ultimate sourcehood, or up-to-me-ness. More precisely, according to Natural Libertarianism, the real causal spontaneity condition says this:

A necessary but individually insufficient condition of deep freedom and (moral or non-moral) responsibility is that all my choosings and doings are causally efficacious in the physical world in a way that is fully consistent with, but also not entailed or otherwise necessitated by, the total set of deterministic or indeterministic causal natural laws, especially including the Conservation Laws, together with all the settled quantity-of-matter-and/or-energy facts about the past, especially The Big Bang.

According to Kant, transcendental freedom is how a human person can, “from itself” (*von selbst*) (Kant 1997: A533/B561, 533), be the spontaneous mental cause of certain natural events or processes. If I am that human person, then insofar as I am transcendently free, it follows that I am an ultimate source of my choices and intentional actions precisely because certain events or processes in physical nature are *up to me* – or to use Kant’s own phrase, *in meiner Gewalt* (literally: “in my control” or “in my power”; Kant 1996: Ak 5, 94–95, 215–216). So otherwise put, Kant’s misleadingly labeled transcendental freedom is in effect deep freedom of the will (aka ultimate sourcehood, up-to-me-ness), which in turn necessarily *includes* but also significantly *exceeds* the real causal spontaneity condition alone. More specifically, it significantly exceeds the real causal spontaneity condition by adding

- i a condition requiring what I call “the capacity for self-commitment to a live option,” or *the Kierkegaardian Either/Or*, and also
- ii an *ownership* condition.

Fully explicitly now, a rational animal or real person *P*'s choosing or doing *X* is deeply free if and only if

- i *P* is the far-from-equilibrium, spatiotemporally asymmetric, complex, self-organizing, organismic, finelygrainedly normatively attuned, thermodynamic, causal-dynamically unprecedented, causal-nomologically constrained-yet-also-necessarily-underdetermined, creative, and self-guiding efficacious cause of choosing or doing *X*, or refraining from so choosing or so doing (*that is, the real causal spontaneity condition is satisfied*),
- ii *P* can choose or do something *X*, as opposed to not choosing or doing *X*, and *X* would never actually happen (or: would not have happened) if *P* were not to choose (or: had not chosen) *X*, so that *P* has a live option for self-commitment in the Kierkegaardian sense of an Either/Or, even if, in context, there are no alternative possibilities in the classical sense of indeterministic branching futures, and the future is temporarily not open (*that is, the capacity for self-commitment to a live option condition is satisfied*), and
- iii *P*'s choosings and doings belong to her and her life, as a self-identical real person, and do not belong to some other agent or agency (*that is, the ownership condition is satisfied*).

This completes a general description of *the internal structure of deep freedom*.

According to Natural Libertarianism, however, even the complex core metaphysical fact of deep freedom (= real causal spontaneity + the capacity for self-commitment to a live option or Kierkegaardian Either/Or + ownership) does not exhaust the *complete* fact of free will, which as I have said is more properly characterized as *free agency* in order to accommodate the further fact of *practical agency*. More precisely put, the free agency of rational animals or real persons has not only a complex structure, but more specifically a *hierarchical* or ordered-and-leveled, structure. Even more precisely still, according to Natural Libertarianism, the capacity for free agency has three logically distinct ordered sub-levels within it, such that each lower-numbered level is presupposed by, but also ontologically nested in, all the higher-numbered levels:

- i the capacity for veridical psychological freedom,
- ii the capacity for deep freedom (aka ultimate sourcehood, up-to-me-ness), and
- iii the capacity for principled authenticity.

In *Deep Freedom and Real Persons* (Hanna 2018a: chs. 4–5), I fully spell out and defend the Natural Libertarian account of the internal structure of deep freedom and free agency, against the backdrop of a detailed critique of what I call “The Three Standard Options,” as well as Classical Compatibilism and Classical Incompatibilism. For the record and because it is relevant to what follows, The Three Standard Options are:

Hard Determinism: Free will and Universal Natural Determinism are mutually inconsistent, free will is impossible, and Universal Natural Determinism is true.

Soft Determinism: Free will and Universal Natural Determinism are mutually consistent, free will exists, and Universal Natural Determinism is true.

Classical Libertarianism: Free will and Universal Natural Determinism are mutually inconsistent, free will exists, and Universal Natural Determinism is impossible,

- either (i) because free will exists as an essential property of a special agent-substance, existing outside the natural causal order or inside that order (classical agent-causationism),
- or (ii) because some indeterministic processes exist in nature and free will is among them (event-causal indeterminism),
- or (iii) because some indeterministic processes exist in nature and free will is not among them because it exists over and above natural processes (non-causal indeterminism).

Correspondingly, the thesis that free will and Universal Natural Determinism are mutually logically or metaphysically consistent is Classical Compatibilism. So Soft Determinism is a form of Classical Compatibilism. And the thesis that, on the contrary, free will and Universal Natural Determinism are mutually inconsistent is classical Incompatibilism. So Hard Determinism and Classical Libertarianism are both forms of classical Incompatibilism.

Here is a brief follow-up comment about classical agent-causal theories. If an (otherwise) classical agent-causal theorist were to reject the metaphysics of substance in favor of a metaphysics of dynamic processes in minded animals of a certain level of cognitive complexity, then her view would be very close to mine. In fact, this is the case with Helen Steward's view in *A Metaphysics for Freedom*. For example, she writes:

I am convinced that to solve the problems surrounding animal agency simply it is not enough to cast off mistaken theories of causation. . . . It needs also to be shown that real, biological processes might enable us to sustain the idea that an *animal* may be truly in charge of what it does, so that its actions are more than merely the byproduct of its innards and parts. The task requires some reflection on the organizational principles of living creatures, for it is only through such reflection . . . that we can start to understand where the difference really lies between, on the one hand those things that are true agents, and, on the other, mere machines, entities that nothing will ever be up to, however impressive they may be. . . . I am exceedingly hopeful that the next few years will see the beginnings of a revolution in our conception of the human person, as philosophical and everyday conceptions of the scientific picture of the world are freed from outdated Newtonian ideas and begin to take more note, both of the complexities of science as it really is and of the undeniable fact of our animal nature.

(Steward 2012: 198–199)

So whether, on the one hand, like Steward, one starts out by officially calling oneself a classical agent causal theorist and then adopts a process metaphysics of agency and persons that in fact undermines and radically transforms the classical metaphysics of agent-substances, or whether, on the other hand, like me, one starts out by officially rejecting the background substance-dualist metaphysics of the classical agent-causal theory and developing a process metaphysics of free will, practical agency, and persons, against the backdrop of The Essential Embodiment Theory of the mind-body relation and mental causation, and then incorporates the core ultimate sourcehood/up-to-me-ness/source-incompatibilist idea of agent causal theories within that non-dualist, process-metaphysics framework, is probably just a nominal or stylistic difference at the end of the day. Thus, a neo-classical or post-classical agent-causal theory like Steward's, and a source-incompatibilist, Essentially Embodied-Agency theory like mine, can come out to pretty much the same thing.

In any case, the thrilling upshot of all the argumentation in *Deep Freedom and Real Persons* (Hanna 2018a: chs. 1–5) is that Natural Libertarianism is *neither* one of the Three Standard Options, *nor* a version of Classical Compatibilism, *nor* a version of classical Incompatibilism. Triply on the contrary, it is a non-classical, non-standard, original and unorthodox, slightly mind-bending, anti-mechanist Libertarian view that constitutes an *Incompatibilistic Compatibilism* and fuses Wilfrid Sellars’s dialectically opposed conceptions of the Manifest Image and the Scientific Image (Sellars 1963) by means of a metaphysical synthesis I call *liberal naturalism*. So for the purposes of the rest of this chapter/essay, I will now concentrate on my sub-account of *the capacity for veridical psychological freedom*, because it captures the nature of *essentially embodied agentive phenomenology*.

Strictly speaking, psychological freedom, when it is unspecified as to whether it is veridical or non-veridical and formulated as a weak disjunction – “veridical-or-non-veridical psychological freedom” – is *also* a necessary but not sufficient condition of deep freedom; and deep freedom, ultimate sourcehood, or up-to-me-ness, is a necessary but not sufficient condition of principled authenticity.

Psychological freedom, *per se*, without regard to its veridicality or non-veridicality, is my first-order consciousness of being both negatively and positively free. Or otherwise put, psychological freedom, *per se*, is my subjective experience of having an unfettered and really causally spontaneous will. This consciousness, in turn, can be

- either (i) a correct or true consciousness of being both negatively and positively free, such that it is an actual fact that I am both negatively and positively free, in which case, it is what I call *veridical psychological freedom*, so that at any time, the intentional subject is either in one kind of state or else in the other, never both,
- or (ii) an incorrect or false consciousness of being both negatively and positively free, that is, a mere seeming to be both negatively and positively free, such that I in fact am neither negatively nor positively free, in which case, it is what I call *non-veridical psychological freedom*.

Correspondingly, I am also committed to the thesis of *strong metaphysical disjunctivism* about the difference between veridical psychological freedom and non-veridical psychological freedom,² which says that

- iii veridical psychological freedom and non-veridical psychological freedom essentially share no intentional or phenomenological content whatsoever, even if they accidentally share some (or even many) other psychological or non-psychological properties, and
- iv the difference between veridical psychological freedom and non-veridical psychological freedom is, in principle, inherently *discriminable* for rational human agents, even if, in context, or in a specific range of contexts, it is actually *undiscriminated* because the agent’s capacity for discrimination is adversely affected or suppressed in that context or those contexts.

So obviously psychological freedom, *per se*, when it is non-veridical psychological freedom, is consistent with my *not* really being negatively or positively free. Furthermore, non-veridical psychological freedom is also inherently conceptually determined, that is, inherently open to “cognitive penetration.” By sharp contrast, *veridical* psychological freedom is essentially non-conceptual; hence, it is also pre-reflectively conscious, or non-self-conscious and non-reflective, and therefore impervious to “cognitive penetration.” The currently popular

“cognitive penetration thesis” just means that for any given kind of human cognition, its representational content is necessarily and sufficiently determined by propositional content and thereby also necessarily and sufficiently determined by conceptual content, on the standard assumption that all propositional content is, ultimately, reducible to conceptual content. So my central point here is that since veridical psychological freedom is essentially non-conceptual, then it is *impervious* to cognitive penetration, and thereby *immune* to propositional or conceptual errors.

Nevertheless, since rational human agents like us are *also* self-conscious and reflective, then the possession of psychological freedom for rational human agents like us, whether it is veridical psychological freedom or non-veridical psychological freedom, also entails a further capacity for having self-directed beliefs to the effect that we are negatively and/or positively free. In the case of veridical psychological freedom, obviously, those self-directed beliefs to the effect that we are negatively and/or positively free, by virtue of the essentially non-conceptual content of veridical psychological freedom, are *true* beliefs; whereas in the case of non-veridical psychological freedom, by virtue of its erroneous propositional and conceptual content, those self-directed beliefs are *false* beliefs. Thus, we can be deceived about our freedom and have what I call *sheer illusions* of deep freedom.

These sheer illusions can occur in drug-induced or pathological dreams,³ hallucinations, or, most poignantly, in causally overdetermined pathological waking psychological states, such as cases of sociopathic paranoid schizophrenics who believe they are choosing and acting with free will, but are actually delusional and insane, and acting under an irresistible compulsion when they commit crimes, hence are correctly legally judged to be “not guilty by reason of insanity.” One such case – or it seems to me from the available evidence – is the real-world waking nightmare of a former Yale Law School student, Ketema Ross:

Early one morning in 2007, Ross heard President George W. Bush [Yale] ‘68 telling him that his next-door neighbors were traitors who needed to be gotten rid of. Ross broke into the elderly couple’s apartment and beat them with a broom handle. (They both survived the attack.) Charged with assault, he pleaded not guilty by reason of insanity. Now Ross says he has recovered his sanity, and a court order says he is no longer “a substantial danger.” And, after seven years of confinement in a psychiatric hospital, he has regained his freedom, mostly: by court order, he was conditionally discharged on January 11 [2015].

(Bass 2015: 49)

Now psychological freedom, *per se*, is not a sufficient condition of deep freedom, precisely because, as some philosophers (including Spinoza, Leibniz, Hume, and more recently Frankfurt in 1988b, and many other Frankfurt-inspired philosophers of agency) have correctly pointed out, both the pre-reflective consciousness, and also the self-conscious or reflective belief, of having an unfettered and spontaneously really causal will, and of being both negatively and positively free, are perfectly consistent with Universal Natural Determinism. Formulated in my terminology, if either Universal Natural Determinism, Universal Natural Indeterminism, or the disjunctive combination of these that I have called *Natural Mechanism* were true, then all of these conscious or self-conscious states would be cases of *non-veridical* psychological freedom and/or false self-directed beliefs about my deeply free will.

Nevertheless, both veridical psychological freedom and also veridical psychological freedom’s discriminability from non-veridical psychological freedom collectively yield

a conjunctive necessary epistemic condition of deep freedom, by way of the capacity for self-commitment to a live option, or Kierkegaardian Either/Or. This condition says that no one could have a capacity for self-commitment to a live option *X* and at the same time

either (i) be in a state of non-veridical psychological freedom,
 or (ii) truly believe herself to be prevented from choosing *X* or doing *X*,
 or (iii) truly believe herself to be inwardly or outwardly compelled to choose *X* or do *X*,
 or (iv) truly believe herself to be otherwise unable to choose or do what she wants.

More precisely, by a *self-consciously or reflectively fettered, epiphenomenal, will* I will mean one's self-conscious or reflective awareness to the effect that, and correspondingly one's belief about oneself to the effect that, one is helplessly violated by inner or outer forces. But more briefly put, this is when someone vividly feels like a natural automaton (biochemical puppet, moist robot, "meat puppet," etc.), or like a tool in the hands of some other powerful manipulative agent or agency. This would be vividly true of, e.g., the Ketema Ross case, and also of pathological conditions like OCD. More specifically, mentally caused body-movements under OCD fall short of free agency in at least four ways:

- i they fail the condition of veridical psychological freedom, since in mentally causing body-movements brought about by her OCD, the subject self-consciously experiences herself as being helplessly overwhelmed by inner forces,
- ii they fail the Kierkegaardian Either/Or condition, since in mentally causing body-movements brought about by her OCD, the subject does not commit to these as a live option for her – on the contrary they are something ultimately forced on her by neurobiological processes internal to her body,
- iii they fail the ownership condition, since in mentally causing body-movements brought about by her OCD, the subject is not performing intentional body-movements that flow from that person *herself*, as an ultimate source – on the contrary, they are something that is *external to the self*, even despite their being mentally caused by the subject, and
- iv as a consequence of (i) to (iii), the bodily movements brought about by the subject's OCD are not something for which she is deeply (morally or non-morally) responsible.

So I am saying that in order to have deeply free will, then, we must *not* have a self-consciously or reflectively fettered, epiphenomenal, will: on the contrary, we must not only *have* veridical psychological freedom, but also *be at least fully disposed to believe, or actually believe, ourselves to have an unfettered, non-epiphenomenal, real causally spontaneous will.*

Ironically, this is as true of self-styled Hard Determinists as it is of everyone else. I am absolutely sure that when these philosophers choose and act, under normal conditions, they do not actually feel or believe in accordance with their own metaphysics of free will – that they do not either really feel like biochemical puppets, moist robots, "meat-puppets," etc., or really believe themselves to be natural automata of any kind. To be sure, the logical scope of their philosophical beliefs extends universally over all people, including themselves: but epistemically speaking, it is one thing to apply property *P* to everyone, including of course oneself, *third-personally*, and quite another thing to apply property *P* to oneself, *first-personally*. Universal instantiation is neither semantically nor epistemically equivalent to first-personal indexical predication. If *everyone* is *P* then it necessarily follows that *I* am *P*, because I am one of the people in the domain of discourse. But if everyone is *P*, even if I believe that everyone

is *P*, it does not necessarily follow that *I really believe that I am P*, because I still have to identify myself with one of the many people in the domain of discourse, and I might self-consciously or non-self-consciously refuse to do that.

Otherwise put, if Hard Determinists really *did* believe that *they themselves were natural automata* (biochemical puppets, moist robots, “meat puppets,” etc.), then obviously they would seek – or at least obviously they would at least *need* – psychiatric help for the treatment of well-attested symptoms of schizophrenia, like the unfortunate Ketema Ross (Bass 2015; see also Mellor 1970; Frith 1992; Spence 1996). Of course, they do *not* really believe it; hence, they have effectively compartmentalized, segregated, and causally isolated their purported philosophical beliefs from their basic agentic commitments or “beliefs” in the specific sense of essentially non-conceptual, hierarchical-desiderative, volitional *believing-in* or *faith* (*Glaube*) – as it were, their heart of hearts – and are, as a consequence, about as healthy and sane as professional academic philosophers usually are. Furthermore, and more generally, I submit that no healthy, sane real human person ever really and truly believes himself or herself, in the sense of essentially non-conceptual, hierarchical-desiderative, volitional *believing-in* or *faith*, their heart of hearts, to be a natural automaton (biochemical puppet, moist robot, etc.), *no matter what their free will metaphysics says*.

Consider, e.g., Classical Compatibilism/Soft Determinism. The very idea of “real” freedom of choice or action, when taken together with a self-consciously or reflectively fettered, epiphenomenal, will, would be as absurd and pointless as freedom of action together with either a naturally determined or a powerfully manipulated will. Indeed, as Frankfurt rediscovered and influentially pointed out, psychological freedom, *per se*, although it is consistent with Universal Natural Determinism (namely, when it is non-veridical psychological freedom), nevertheless remains essential to our rational human personhood, as veridical psychological freedom, at the level of effective first-order desires, together with the self-conscious or reflective awareness of having veridical psychological freedom, at the level of what Frankfurt calls *decisive* or *self-identifying* second-order volitions (Frankfurt 1988b, 1988d, 1988f). This is because a self-consciously or reflectively fettered, epiphenomenal, will utterly defeats and undermines our belief in our own intentional agency.

Now consider Hard Determinism. Virtually every theory that falls under these rubrics *also* has a “debunking strategy” or “error-theory,” which says that our brains mechanically create the cognitive illusion that we are really free, even though we are really natural automata. But then those theories have the following serious epistemic problem. Suppose that you hold the following view:

Free will *is* an illusion. Our wills are simply not of our own making. Thoughts and intentions emerge from background causes of which we are unaware and over which we exert no conscious control. We do not have the freedom we think we have.

(Harris 2012: 5)

If that is true, then we are all natural machines with an irresistibly strong tendency to create cognitive illusions for ourselves. Therefore, under the supposition that her theory is true, any holder of such a view *cannot rule out* the real possibility that she has created a cognitive illusion for herself by defending Natural Mechanism together with a debunking strategy or an error-theory. But if she cannot rule this out, *then she is not rationally justified in believing in her own theory*. So her belief in her own theory is cognitively self-stultifying. This conclusion, in turn, *debunks the would-be debunkers*.

Now suppose that someone were to object to that last argument by saying:

This argument, purporting to show that Hard Determinism is self-stultifying, goes by *very fast*. Would you embrace the claim that since one cannot rule out the possibility that one is in a lifelong dream, one is not rationally justified in any of one's external world beliefs?

This is how I would reply:

No, I wouldn't embrace *this* claim. Indeed, I've published two papers on Cartesian dream skepticism, in which I explicitly refute it (Hanna 2000; Hanna 1992). Nevertheless, the crucial point in the present context is whether, as you imply, there's a basic analogy between the Cartesian dream-skeptic's (as it turns out, unsound) argument and my self-refutation argument against the Hard Determinist. And what I'd say to *that* is also no, for there's a crucial *disanalogy* between what I am critically arguing *here* against the Hard Determinist, and what the Cartesian dream skeptic would (unsoundly, as it turns out) argue. In turn, that's because, since the Hard Determinist is explicitly committed to a *neurobiologically-based error theory* that applies to *everyone's* capacities for reasoning philosophically, then in an *easily-accessible, nomologically-identical, nearby possible world*, it also applies to the Hard Determinist's capacity for reasoning philosophically too. Now when there are *relevantly alternative* skeptical possibilities, those possibilities do indeed undermine rational justification. Hence, at least so far, my self-stultification argument holds up.

Interestingly and importantly, the closely related view that Derk Pereboom calls "Hard Incompatibilism" does *not* have the entailment that there is no genuine concept of free will (Pereboom 2001, 2014). According to the Hard Incompatibilist, real free will is possible, but actually does not exist. So while, according to Pereboom, there is no actual *fact* of real free will, there is nevertheless an intelligible *concept* and also a vivid *phenomenology* of real free will, and correspondingly, these can be taken to support the thesis that, in some non-trivial sense, we really are "deterministic agent-causes" (Pereboom 2015). In these respects, Hard Incompatibilism is a significantly more subtle and defensible view than Hard Determinism.

But where Hard Incompatibilism has serious internal difficulties, I think, is in reconciling the explicit rejection of the fact of real free will with an explicit acceptance of the genuine *concept* of real free will, the vivid *phenomenology* of free will, and the thesis about our "deterministic agent-causal" powers. This mixed attitude of rejection/acceptance seems to me to produce serious cognitive (and practical) dissonance. According to Pereboom, I really and truly am a biochemical puppet and a moist robot: I know this via mechanistic physics and biology, and "scientific philosophy." Yet at the same time, by virtue of the genuine concept and vivid phenomenology of real free will, I also cannot help believing outside my laboratory or study that I am (in some sense) deeply free and (morally or non-morally) responsible for my choices and actions, and therefore, somewhere *else* in my head and in my heart, I feel in my bones that I am *not* a biochemical puppet or moist robot. When I am wearing my white lab coat and doing experiments, or in my open-necked Brooks Brothers shirt and nice blue jeans teaching philosophy, I believe I am really a puppet and a robot; but when I get home, change into my comfortable T-shirt and baggy shorts, and open a beer, I believe I am really *not* a puppet or robot. This is philosophical *schizophrenia*, not a stable philosophical position.

More specifically, Pereboom claims that the “standard phenomenology of agency” is *evidentially neutral* as between Classical Libertarianism on the one hand, and Hard Determinism/Hard Incompatibilism/Soft Determinism on the other. And according to Pereboom, the “standard phenomenology of agency” is this:

The phenomenology appears to reveal that in paradigm cases, actions are caused not solely by events or states, but are rather actively caused by agents themselves.

(Pereboom 2015: 278)

Pereboom’s evidential-neutrality conclusion may follow *if* we construe the “standard phenomenology of agency” narrowly enough, but two things have gone seriously wrong here.

First, Pereboom has *not* considered the inherently evidentially richer, essentially embodied, agentive phenomenology of spontaneous, vital, non-mechanical sourcehood, natural creativity in “natural open space,” and more generally he has completely overlooked the phenomenology of agency as we find it in the work of O’Shaughnessy, Schopenhauer, Kant, Kierkegaard, and Frankfurt. More specifically, the five crucial additional elements in this inherently evidentially richer phenomenology of agency are:

- i the O’Shaughnessy-inspired and Schopenhauer-inspired identification of the agent’s sourcehood with *her inherently spontaneous and vital, hence non-mechanical, necessary, and complete minded animal embodiment,*
- ii the Kant-inspired *free-agent-as-creative-natural-artist-in-natural-open-space* analogy,
- iii the Existentialism-inspired Kierkegaardian “either/or” conception of *self-commitment-to-a-live-option,*
- iv the Frankfurt-inspired *hierarchical-desire* conception of the will and personhood, and
- v the equally Kant-inspired and Existentialism-inspired idea that free agency is inherently normatively motivated by *the pursuit of principled authenticity.*

Second, and correspondingly, Pereboom has not considered, or even looked-for, any view about the metaphysics of free will and practical agency that is even *remotely* like Natural Libertarianism; on the contrary, he has strategically boxed-in his metaphysical options by critically considering only the various versions of Classical Libertarianism as opposing doctrines.

So the plausibility and soundness of Pereboom’s argument depends entirely on his having started with a needlessly poverty-stricken and boxed-in domain of phenomenological materials and metaphysical options. Therefore, to conclude this critical line of argument by *modus tollens*-ing Pereboom’s *modus ponens*: *if* we do *not* start with these needlessly poverty-stricken and boxed-in materials, but instead on the contrary seriously consider the inherently evidentially richer agentive phenomenology of essentially embodied, spontaneous, vital, non-mechanical sourcehood, natural creativity in natural open space, self-commitment-to-a-live-option, willing-as-hierarchical-desiring, and the pursuit of principled authenticity, together with the metaphysics of Natural Libertarianism, *then* Hard Incompatibilism is clearly false.

It is also very important to point out in this connection that having a self-consciously or reflectively fettered, epiphenomenal, will is categorically *not* the same as the classical, Lord-Byron-style, Romantic self-consciousness of being in the grip of a grand passion that carries you away with it. The crucial contrast between these two is the inherent, categorical difference between

- i believing yourself to be a *mere natural machine*, which is wholly powerlessly caused or helplessly manipulated by something inside your own body or outside your own body, and
- ii believing yourself to be *fully alive*, driven, energized, and invigorated by some naturally purposive and naturally creative vital power that is immensely bigger than you are.

The first-order consciousness or subjective experience of life and vitality, insofar as it essentially non-conceptually and veridically picks out immanent structural properties of non-equilibrium thermodynamic systems, is inherently anti-mechanical and uncomputable. In other words, insofar as the Lord-Byron-style Romantic phenomenology is veridical, then it entails what I call *local incompatibilism with respect to Natural Mechanism* (Hanna 2018a: chs. 4–5). This point is especially telling because the Lord-Byron-style Romantic phenomenology is sometimes used as an intuitively evidential ground for believing in Compatibilism and Soft Determinism (see, e.g., Arpaly 2007). But this line of argument just confuses one kind of “carried away” phenomenology, with a categorically different kind of “carried away” phenomenology. A natural automaton can *never* have an essentially non-conceptual and veridical Lord-Byron-style phenomenology. At best, it could only be that *its psychic motor is racing*.

Therefore, and finally, the crucial difference between the self-evident agentive phenomenology I am talking about, and, e.g., either Terence Horgan’s agentive phenomenology (e.g., Horgan 2011), which presupposes the metaphysical backdrop of Classical Compatibilism/Soft Determinism, or Pereboom’s agentive phenomenology, which presupposes the metaphysical backdrop of Hard Incompatibilism, *is that my essentially embodied agentive phenomenology is veridical*, and presupposes the metaphysical backdrop of Natural Libertarianism. By a first sharp contrast, Pereboom’s agentive phenomenology is *non-veridical*, and nothing but a *cognitive illusion*. And by a second sharp contrast, even though Horgan’s view does not *explicitly* or *officially* present itself as involving either a non-veridical phenomenology or an illusion theory – indeed, it presents itself as involving a *veridical* phenomenology and a *preservationist* theory – I do also think that all versions of Classical Compatibilism/Soft Determinism ultimately collapse into Hard Determinism/Hard Incompatibilism, via their commitment to an essentially non-veridical and thereby *epiphenomenal* agentive phenomenology, just as, as Jaegwon Kim has (to my mind) convincingly argued in *Supervenience and Mind* (Kim 1993), *all* versions of Non-Reductive Physicalism about the mind-body relation ultimately collapse into Reductive Physicalism, via the Causal-Exclusion Argument, aka *the epiphenomenalism problem*. Moreover, at the philosophical end of the day, as Kim puts it in *Physicalism, Or Something Near Enough*:

Epiphenomenalism strikes most of us as obviously wrong, if not incoherent. . . . It is the kind of doctrine . . . that, even if we had to acknowledge it as true, could not serve as a guide to life; it cannot serve as a premise in our practical reasoning, and it is not possible for us to live as though it is true.

(Kim 2005: 70–71)

In other words, and by way of conclusion, Classical Compatibilism/Soft Determinism entails Non-Reductive Physicalism, which entails epiphenomenalism, which entails a non-veridical agentive phenomenology and implicit philosophical self-stultification about free agency – that is what Kim is saying. I heartily agree with Kim on this particular point. And O’Shaughnessy, I am quite sure, would also heartily agree with us.

Related topics

See Chapters 13 (on Merleau-Ponty), 19 (Horgan and Nida-Rümelin), 20 (Smith), 23 (Strawson), and 24 (De Monticelli).

Notes

- 1 Strictly speaking, Davidsonian theories of action do not require that the cause of action be temporally prior, since Davidson's theory is a third-person, interpretation-theoretic account, whereby a given belief-desire pair constitutes a reason that "rationalizes" an underlying physical event-cause, upon which the belief-desire pair are naturally or nomologically supervenient, and supervenience is a synchronous relation. Nevertheless Davidson's and Davidsonian theories have other serious problems, including epiphenomenalism and its own deviant causal chain worry. See Hanna and Maiese (2009: ch. 3).
- 2 Analogously, I am committed to strong metaphysical disjunctivism about sense perception. See Hanna (2015: ch. 3).
- 3 On the other hand, ordinary, *non*-drug-induced, *non*-pathological dreams exhibit real causal spontaneity, as well as satisfying the other conditions for deep freedom. So, strictly speaking, ordinary dreams are what I call *veridical illusions* of deep freedom, because they include
 - i essentially non-conceptual veridicality about our deep freedom, together with
 - ii propositional and conceptual errors about many details of the course of one's actual life and of the actual world in which ordinary dreaming deep freedom purports to be exercised,in a way that's highly analogous to the familiar bent-stick-in-water or bigger-moon-near-the-horizon cases of veridical illusion in the philosophy of perception. In turn, *this entails that our ordinary sleeping and dreaming life is as much a part of our free agential lives as ordinary waking life is*. But the philosophy of sleep is a long story for another day – or to be more precise, for another book.

References

- Arpaly, N. (2007) *Merit, Meaning, and Human Bondage*, Princeton, NJ: Princeton University Press.
- Bass, C. (2015) "By Reason of Insanity," *Yale Alumni Magazine*, pp. 48–53, <http://www.yalealumnimagazine.com/articles/4076/by-reason-of-insanity>.
- Evans, G. (1982) *Varieties of Reference*, Oxford: Clarendon/Oxford University Press.
- Frankfurt, H. (1988) *The Importance of What We Care about*, Cambridge, MA: Cambridge University Press.
- Frankfurt, H. (1988a) "Alternate Possibilities and Moral Responsibility," in H. Frankfurt (1988), *The Importance of What We Care about*, pp. 1–10.
- Frankfurt, H. (1988b) "Freedom of the Will and the Concept of a Person," in H. Frankfurt (1988), *The Importance of What We Care about*, pp. 11–25).
- Frankfurt, H. (1988c) "The Problem of Action," in H. Frankfurt (1988), *The Importance of What We Care about*, pp. 42–52.
- Frankfurt, H. (1988d) "Identification and Externality," in H. Frankfurt (1988), *The Importance of What We Care about*, pp. 58–68.
- Frankfurt, H. (1988e) "The Importance of What We Care About," in H. Frankfurt (1988), *The Importance of What We Care about*, pp. 80–94.
- Frankfurt, H. (1988f) "Identification and Wholeheartedness," in H. Frankfurt (1988), *The Importance of What We Care about*, pp. 159–176.
- Frith, C. (1992) *The Cognitive Neuropsychology of Schizophrenia*, Hillsdale, NJ: Erlbaum.
- Gallagher, S. (2005) *How the Body Shapes the Mind*, Oxford: Clarendon/Oxford University Press.
- Goethe, J.W. (1990) *Faust*, trans. W. Kaufmann, New York: Anchor Books.
- Hanna, R. (1992) "Descartes and Dream Skepticism Revisited," *Journal of the History of Philosophy* 30, 377–398.
- Hanna, R. (2000) "The Inner and the Outer: Kant's 'Refutation' Reconstructed," *Ratio* 13, 146–174.
- Hanna, R. (2006) *Rationality and Logic*, Cambridge, MA: MIT Press.

- Hanna, R. (2015) *Cognition, Content, and the A Priori: A Study in the Philosophy of Mind and Knowledge*, aka THE RATIONAL HUMAN CONDITION, Vol. 5, Oxford: Oxford University Press.
- Hanna, R. (2018a) *Deep Freedom and Real Persons: A Study in Metaphysics*, aka THE RATIONAL HUMAN CONDITION, Vol. 2, New York: Nova Science.
- Hanna, R. (2018b) *Kantian Ethics & Human Existence*, aka THE RATIONAL HUMAN CONDITION, Vol. 3, New York: Nova Science.
- Hanna, R. and Maiese, M. (2009) *Embodied Minds in Action*, Oxford: Oxford University Press.
- Hanna, R. and Thompson, E. (2005) "Neurophenomenology and the Spontaneity of Consciousness," in E. Thompson (ed.), *The Problem of Consciousness*, Calgary, AL: University of Alberta Press, pp. 133–162, https://www.academia.edu/9962718/Neurophenomenology_and_the_Spontaneity_of_Consciousness.
- Harris, S. (2012) *Free Will*, New York: Free Press.
- Horgan, T. (2011) "The Phenomenology of Agency and Freedom: Lessons from Introspection and Lessons from Its Limits," *Humana Mente Journal of Philosophical Studies* 15, 77–97, http://www.humanamente.eu/PDF/Issue15_Paper_Horgan.pdf.
- Ismael, J. (2016) *How Physics Makes Us Free*, Oxford: Oxford University Press.
- Kane, R. (1996) *The Significance of Free Will*, Oxford: Oxford University Press.
- Kant, I. (1996) *Immanuel Kant: Practical Philosophy*, trans. M. Gregor, Cambridge, MA: Cambridge University Press.
- Kant, I. (1997) *Critique of Pure Reason*, trans. P. Guyer and A. Wood, Cambridge, MA: Cambridge University Press. In-text citations include page-numbers from both the A edition (1781) and B (1787) editions of the CPR and also the Cambridge edition.
- Kim, J. (1993) *Supervenience and Mind*, Cambridge, MA: Cambridge University Press.
- Kim, J. (2005) *Physicalism, or Something Near Enough*, Princeton, NJ: Princeton University Press.
- Meat Puppets, The (1994) "We Don't Exist," <https://search.yahoo.com/yhs/search?p=the+meat+puppets+we+don%27t+exist+youtube&ei=UTF-8&hspar=mozilla&hsimp=yhs-003>.
- Mellor, C.S. (1970) "First Rank Symptoms of Schizophrenia," *British Journal of Psychiatry* 117, 15–23.
- O'Shaughnessy, B. (1973) "Trying (as the Mental 'Pineal Gland')," *Journal of Philosophy* 70, 365–386.
- O'Shaughnessy, B. (1980) *The Will: A Dual Aspect Theory*, 2 vols., Cambridge, MA: Cambridge University Press.
- O'Shaughnessy, B. (2008) *The Will*, 2 vols., second edition, Cambridge, MA: Cambridge University Press.
- Pereboom, D. (2001) *Living Without Free Will*, Cambridge, MA: Cambridge University Press.
- Pereboom, D. (2014) *Free Will, Agency, and Meaning in Life*, Oxford: Oxford University Press.
- Pereboom, D. (2015) "The Phenomenology of Agency and Deterministic Agent-Causation," in M. Altman and H. Gruenig (eds.), *For Horizons of Authenticity in Phenomenology, Existentialism, and Moral Psychology: Essays in Honor of Charles Guignon*, New York: Springer, pp. 277–294.
- Rorty, R. (1979) *Philosophy and the Mirror of Nature*, Princeton, NJ: Princeton University Press.
- Rorty, R. (1982) *Consequences of Pragmatism*, Minneapolis: University of Minnesota Press.
- Sartre, J.-P. (1966) *Being and Nothingness*, trans. H. Barnes, New York: Washington Square Press.
- Schopenhauer, A. (1907) *On the Fourfold Root of the Principle of Sufficient Reason*, trans. Mme. K. Hillebrand, London: George Bell and Sons, <http://www.gutenberg.org/files/50966/50966-h/50966-h.htm>.
- Schuessler, J. (2013) "Philosophy That Stirs the Waters," *New York Times*, 29 April 2013, http://www.nytimes.com/2013/04/30/books/daniel-dennett-author-of-intuition-pumps-and-other-tools-for-thinking.html?emc=eta1&_r=0.
- Sellars, W. (ed.) (1963) "Philosophy and the Scientific Image of Man," in *Science, Perception, and Reality*, New York: Humanities Press, pp. 1–40.
- Spence, S.A. (1996) "Free Will in the Light of Psychiatry," *Philosophy, Psychiatry, and Psychology* 3, 75–90.
- Steward, H. (2012) *A Metaphysics for Freedom*, Oxford: Oxford University Press.
- Wittgenstein, L. (1953) *Philosophical Investigations*, trans. G.E.M. Anscombe, New York: Macmillan.
- Wollheim, R. (1984) *The Thread of Life*, Cambridge, MA: Harvard University Press.