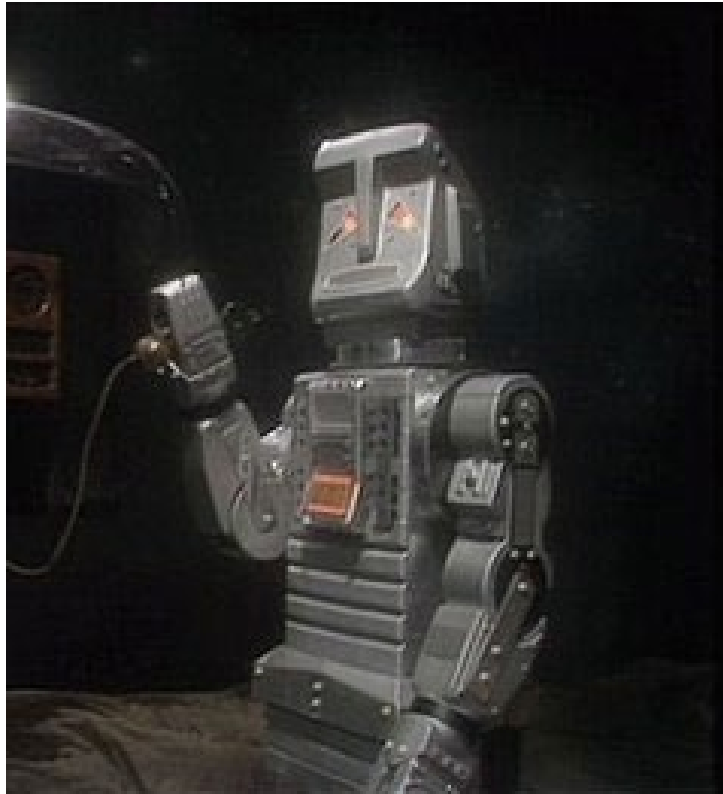


The Refutation of Compatibilism and Soft Determinism, and A New Proof That Incompatibilistic Real Free Agency Really Exists, and That You Really Have It

Robert Hanna



Marvin the Paranoid Android (Wikipedia, 2024)

When nature has unwrapped, from under this hard shell, the seed for which she cares most tenderly, namely the propensity and calling to *think* freely, the latter gradually works back upon the mentality of the people (which thereby gradually becomes capable of *freedom* in acting) and eventually even upon the principles of *government*, which finds it profitable to itself to treat the human being, *who is now more than a machine*, in keeping with his dignity. (Kant, 1784/1996: p. 22, Ak 8: 41-42)

So you think that you're both naturally determined and a free agent, and also that this philosophical position is conceptually coherent (see, e.g., Dennett, 1984)? Wrong, wrong, and wrong again. In this essay, I'll refute compatibilism and soft determinism, and then proceed to a new proof that incompatibilistic real free agency really exists and that you, the reader of this very essay, really have it.

I'll start by presenting my account of the mind-body relation and mental causation.

According to my view of the mind-body relation and mental causation, *consciousness is subjective experience* and, in turn, *consciousness is a form of life*. More precisely, however, in *Embodied Minds in Action* (Hanna and Maiese, 2009), Michelle Maiese and I claim that the mental-physical relation in minded living organisms like us is nothing more and nothing less than (i) a *synthetic a priori two-way necessary complementarity relation*, and also (ii) a *neo-Aristotelian hylomorphic relation*, that is, a mental-to-physical and also physical-to-mental entangled necessary equivalence of “fused” inherently activating irreducible *formal* or *morphic* mental properties on the one hand, and complex non-equilibrium thermodynamic *material* or *hyletic* biological physical properties on the other, such that, (iii) as *minded animals*, i.e., as conscious living organismic animal bodies, we're an indissoluble and physically irreducible form-matter composite, by virtue of which we're always “minding our bodies” (Hanna, 2011), that's (iv) *inherently poised for causally efficacious intentional action*, spontaneously initiated and creatively guided by our synchronous acts of desire-based *willing* (Hanna, 2020). In short, *our minds are physically irreducible forms of animal life* and *we're essentially embodied minds in action*; and this is what Maiese and I call *the essential embodiment theory* of the mind-body relation and mental causation.

Now I'll analyze *real free agency*.

By “free agency,” I mean “free will and practical agency,” which in turn means (i) that you really can consciously and self-consciously choose and do what you want to, or refrain from so choosing or doing, thus without being in any way compelled or prevented by irresistible inner or outer forces (free will), and (ii) that you really can consciously and self-consciously choose and do what you want to, for reasons, and with moral responsibility (free agency). And by “moral responsibility” for X, I mean (i) that X is something you really chose or did, whose objective moral value flows from and directly attaches to your free choice or action, and (ii) that moral responsibility requires freedom—if you weren't free to choose or do X, you couldn't be responsible for it.

Real free agency in this sense is *an objective fact about the natural universe*, and not merely something *psychological*—for example, a *belief in free agency* or a *consciousness of free agency*. In principle, those could exist even if real free agency didn't. Nevertheless, real free agency does *also* include a rich psychological aspect: namely, capacities for consciousness, self-consciousness, desires, and choices. And if the capacity for rationality is added to those capacities, then it's also possible to believe in your own free agency.

Fully explicitly now, something X has real free agency if and only if (i) X is a certain kind of complex living organism, namely an animal, and *not* a machine, whether this machine is a *deterministic* automaton, whose behaviors are necessitated by all the settled facts about the past, together with the laws of nature, or an *indeterministic* automaton, whose behaviors occur according to statistical laws, and are always more or less random, (ii) X is conscious, i.e., X has a capacity for subjective experiences, that is also at least sometimes actualized, (iii) X is self-conscious, i.e., X has a capacity for being conscious of its own subjectively experiential acts, processes, or states, that is also at least sometimes actualized, (iv) X really can consciously &/or self-consciously choose and do what X desires to choose and do, or consciously &/or self-consciously refrain from so choosing and so doing, without being in any way compelled or prevented by irresistible inner or outer forces, and (v) X has a *live option* LO, which means that (v.1) X can commit themselves to choosing or doing LO, or not, (or: X could have committed themselves to choosing or doing LO, or not), (v.2) LO would never actually happen unless X were to choose it or do it, (or: LO would never have actually happened unless X had chosen it or done it), and (v.3) X actually chooses and does LO, or not (or: X actually chose and did LO, or not).

In short, X is a *free agent*, and according to my view, free agency is equally non-deterministic and non-*indeterministic*, precisely because it's *purposive and self-organizing*, or internally and spontaneously goal-oriented and organismic or non-mechanical, and also includes a live option. In the jargon of contemporary metaphysics of free will, this collectively yields a *neo-organicist* version of *source incompatibilism without alternative possibilities* (Hanna, 2018a: ch. 5). Moreover, if X *also* has a capacity for rationality—namely, a complex capacity for logical inference, for practical decision-making, and for formulating, recognizing, and being guided by rules, principles, and instrumental or non-instrumental ideals, standards, and values (see, e.g., Hanna, 2006, 2015)—then by virtue of having rationality + free agency, not only is X a free agent, non-deterministic, non-*indeterministic*, purposive, organismic or non-mechanical, and the possessor of a live option, but also X is a *rational free agent*, who is *the ultimate source*—against the metaphysical backdrop of source incompatibilism without alternative possibilities—of their really free choices and actions, for which they are *causally and morally responsible*. For a detailed and extended argument that the preceding analysis of real free agency is not only conceptually coherent but also rationally defensible and correct, see (Hanna and Maiese, 2009; Hanna, 2018a: chs. 1-5, 2020, 2024a: esp. chs. 1, 3, 5 and 16).

Most contemporary professional academic philosophers and natural scientists hold that you are *not* really a free agent, because they also believe that the truth of contemporary natural science entails a thesis I'll call "natural mechanism." The thesis of *natural determinism* says that everything that happens now and in the future is

strictly fixed by the laws of nature together with all the actual facts about the past. And the thesis of *natural indeterminism* says that at least some things and perhaps all things that happen are *not* strictly fixed by the laws of nature together with all the actual facts about the past, but also happen more or less randomly, according to mathematical laws of probability. Conjoining natural determinism and natural indeterminism, the thesis of *natural mechanism* says that everything that happens is either naturally deterministic, naturally indeterministic, or some mixture of both, and that all its causal and quantitative characteristics are not only fixed by the general causal laws of nature, especially those laws governing the conservation of quantities of matter or energy, together with all the settled facts about the past, especially including The Big Bang, but also calculable from those laws and facts on an ideal digital computer. If natural mechanism is true, then you're really free, because, instead, no matter what you may believe about your freedom, you are really a deterministic or indeterministic natural automaton, ultimately caused by The Big Bang.

Now let's focus on natural determinism. The widely-held thesis of *compatibilism* says that *it's logically and metaphysically possible* to be both a free agent and also naturally determined. And the widely-held thesis of *soft determinism* says that we're *actually* both free agents and also naturally determined.

As per the above, if we're naturally determined, then we're machines. But if we're machines, then we're not free agents, since all free agents are conscious and self-conscious living organisms.

If you really believe that you're a machine, then you can't also consistently believe that you're a free agent, since the belief that you're a machine means that you also believe that you're not the ultimate source of your choices and actions, which is a necessary condition of free agency. So if you believe that you're a machine, then you believe that you're not a free agent.¹

Moreover, if you actually *were* a machine, then you couldn't actually *know* it, since knowledge, as sufficiently justified true belief, requires free agency in your choice of beliefs.

Therefore, self-evidently, both compatibilism and soft determinism are false. On the contrary, the view I call *natural libertarianism* is true (Hanna, 2018a). Moreover, natural

¹ For an interesting dialogical debate about this crucial point, see (Horgan and Nida-Rümelin, 2020). Horgan takes the compatibilist/soft determinist position, and Nida-Rümelin takes the—in effect—natural libertarian position. Obviously, I strongly disagree with Horgan and strongly agree with Nida-Rümelin.

libertarianism is not only consistent with but entailed by a comprehensive neo-organicist approach to the philosophy of science and nature that I call *science for humans* (Hanna, 2024a).

The only remaining important philosophical question in this connection then is, why do so many contemporary professional academic philosophers and scientists believe the theses of compatibilism and soft determinism? My own view is that this is brought about by their dogmatic, rationally unjustified commitment to *materialism* or *physicalism* and *scientism*. So in their professional academic lives, they're committed to beliefs that the act of living their everyday lives directly contradicts. They're *professional academic* compatibilists and soft determinists, and *existential* natural libertarians, living an essentially inauthentic double life. The late Daniel Dennett was a paradigmatic example (Hanna, 2024b).

Now I'm going to present a new proof that incompatibilistic real free agency really exists, and that you, the reader of this very essay, really have it.

I'll start by defining *pretending*. Someone X pretends to be something Y or some kind of thing F if and only if (i) X is conscious, (ii) X is self-conscious, (iii) X is in fact *not* either Y or an F, (iv) X consciously or self-consciously either (va) overtly acts like or imitates Y or an F for the purpose of deceiving others or themselves (hence self-deception) into falsely believing that X is Y or an F, (vb) overtly acts like or imitates Y or an F for the purpose of entertaining others (e.g., film actors or stage actors) or themselves, or (vc) merely imagines acting like or imitating Y or an F for the purpose of either (vc1) simply amusing themselves or (vc2) as an interesting thought-experiment. This definition fairly closely follows the definitions of "pretence" and "pretend" in the *Oxford Encyclopedic Dictionary* (Hawkins and Allen, 1991: p. 1146), with a few philosophical elaborations for the sake of clarity, distinctness, and completeness.

From this definition, it follows necessarily that if someone, X, *can* pretend to be something Y or some kind of thing F, then in fact X is *not* Y and *not* an F. Someone can pretend to be *only what in fact they are not*. Someone X can of course *portray* themselves in a film or on stage, but that's not *pretending* to be themselves. No one can pretend to be *what they in fact are*.

Now, granting me that analysis of *real free will* and that definition of *pretending* for the purposes of argument, then *please pretend to be a machine—say, a robot or a super-computer* for a minute or two. Yes, it's OK if you do this by means of getting up and walking around the room in a very stiff-legged and stiff-armed way and repeatedly say "Beep beep, does not compute!" in a loud artificial-sounding voice, like Robby the Robot

in the 1956 movie, *Forbidden Planet*, or in the 1960s TV program, *Lost in Space*, or else by means of muttering in an artificial-sounding depressed voice,

I didn't ask to be made: no one consulted me or considered my feelings in the matter. I don't think it even occurred to them that I might have feelings. After I was made, I was left in a dark room for six months... and me with this terrible pain in all the diodes down my left side (Wikipedia, 2024),

like Marvin the Paranoid Android in the 1981 British TV series *Hitchhiker's Guide to the Galaxy*—based of course on the novels by Douglas Adams—as per the image at the top of this essay. And it's also perfectly OK if you do this just by means of sitting in your comfortable chair or at your desk and singing the song "A Bicycle Built For Two" like HAL did in Stanley Kubrick's 1968 classic science fiction movie, *2001: A Space Odyssey*—actually the voice of the Canadian actor Douglas Rain—or something of that sort. Or you could even pretend to be a Large Language Model like ChatGPT, by answering written questions according to an algorithm that tracks the probability of one bit of text being followed by another bit of text, over an existing corpus of texts (Hanna, forthcoming). Or if you prefer to do any of this by means of merely *imagining* acting like or imitating a machine, then that's perfectly OK too.

Then, let's suppose (i) that you did indeed pretend to be a machine when I asked you to do so, and also (ii) that counterfactually, you could have easily refrained from pretending to be a machine when I asked you to do so, if you had desired to refrain—either because you just didn't feel like pretending to be a machine at that moment, or because you thought it was extremely silly to pretend to be a machine, or whatever.

From what I've already argued, it follows necessarily (i) that you're conscious and self-conscious, hence you're *minded*, (ii) that you really could consciously &/or self-consciously choose and do what I asked you to choose and do, or else really could have consciously &/or self-consciously refrained from so choosing or so doing, without being in any way compelled or prevented by irresistible inner or outer forces, (iii) that you really did have a live option, namely, pretending to be a machine, or not, (iv) that you're *not* a machine, because, by the definition of *pretending*, one can pretend to be only what one in fact is not, therefore since you pretended to be a machine, then necessarily you're *not* a machine, (v) that because you're minded but not a machine, then you're a certain kind of complex living organism, namely a *minded animal*, and finally (vi) that incompatibilistic real free agency exists and you, the reader of this very essay, really have it. **QED**

The key to the soundness of this proof is the fact that pretending is a mode of what I've called *authentic human creativity*, which brings about a categorical improvement or upgrade in the intrinsic specific character or quality of its creative inputs or materials (Hanna, forthcoming). In the case of pretending, the creative inputs or materials are the everyday or quotidian life of the pretender, so that the pretender becomes, temporarily, an icon or image of the thing or kind of thing that's pretended—namely, something or some kind of thing that the pretender is *not*. So pretending is a mode of *self-transcendence* and *self-transformation*: it's authentic human creativity that's self-applied to the pretender's own life.

This self-transcendence and self-transformation is vividly manifest in the kind of pretending that's characteristic of film acting or stage acting, for example, the actors who performed Robby the Robot, Marvin the Paranoid Android, and the voice of HAL, but also especially of *great* film actors or stage actors—see, example, the novelist Frances Burney's marvellous description of a morning spent in the company of the great 18th century actor, David Garrick, in 1775 (Burney, 2001: pp. 42-44). Nevertheless, authentic human creativity is also present in *any* kind of pretending, no matter how humble (as in the pretending of children in make-believe play), immoral (as in the pretending of con artists and scammers), existentially inauthentic or phony (as in the pretending of French waiters so brilliantly described by Jean-Paul Sartre in *Being and Nothingness* [Sartre, 1943/1956: pp. 101-103]), or tragic (as in the case of the Samurai warlord's political double or "shadow warrior" in Akira Kurosawa's brilliant 1980 film *Kagemusha*).

Of course, pretenders can *use* machines, including computing machinery or digital technology more generally; but in all such cases, the machines aren't *themselves* pretending anything. For no machine—and therefore no computing machinery or digital technology more generally—can ever exemplify authentic human creativity (Hanna, forthcoming), and all free will necessarily includes and expresses authentic human creativity (Hanna, 2018a: chs. 1-5). So when we pretend that we're machines, this vividly exemplifies the metaphysically, morally, and politically profound fact that, as Kant pointed out in 1784, we are *not* machines, and therefore we should never be treated *like* machines by the State. Instead we *are* authentically creative free agents innately possessing human dignity, and should always and only be treated as such by each other and by the State alike. And if the State is inherently incapable of doing this, then the State must be devolved out of existence, permanently exited, and replaced by an essentially better and inherently dignitarian kind of social institution (Hanna, 2017, 2018b, 2023a, 2023b).

But because the false and conceptually incoherent belief in compatibilism and soft determinism, and in our machinehood, undermines not only our belief in our own free

agency but also our own free agency itself, then it also undermines our moral and sociopolitical agency, and we turn ourselves into passive, quietistic slaves. How very convenient it would be for US President-elect Donald Trump and his Republican myrmidons, then, if we were mind-fucked into believing in that compatibilism and soft determinism are true, and that we're essentially machines (Hanna, 2024c).²

² I'm grateful to Scott Heftler and James Schofield for thought-provoking conversations on and around the main topics of this essay, and also to James Schofield for drawing my attention again to (Horgan and Nida-Rümelin, 2020).

REFERENCES

(Burney, 2001). Burney, F. *Frances Burney: Journals and Letters*. London: Penguin.

(Dennett, 1984). Dennett, D. *Elbow Room: The Varieties of Free Will Worth Wanting*. Cambridge MA: MIT Press.

(Erhard and Keiling, 2020). Erhard, C. and Keiling, T. (eds.) *Routledge Handbook: The Phenomenology of Agency*. London: Routledge.

(Hanna, 2006). Hanna, R. *Rationality and Logic*. Cambridge: MIT Press. Available online in preview at URL = <[https://www.academia.edu/21202624/Rationality and Logic](https://www.academia.edu/21202624/Rationality_and_Logic)>.

(Hanna, 2011). Hanna, R. "Minding the Body." *Philosophical Topics* 39: 15-40. Available online in preview at URL = <[https://www.academia.edu/4458670/Minding the Body](https://www.academia.edu/4458670/Minding_the_Body)>.

(Hanna, 2015). Hanna, R. *Cognition, Content, and the A Priori: A Study in the Philosophy of Mind and Knowledge*. THE RATIONAL HUMAN CONDITION, Vol. 5. Oxford: Oxford Univ. Press. Available online in preview at URL = <[https://www.academia.edu/35801833/The Rational Human Condition 5 Cognition Content and the A Priori A Study in the Philosophy of Mind and Knowledge OUP 2015](https://www.academia.edu/35801833/The_Rational_Human_Condition_5_Cognition_Content_and_the_A_Priori_A_Study_in_the_Philosophy_of_Mind_and_Knowledge_OUP_2015)>.

(Hanna, 2017). Hanna, R. "Exiting the State and Debunking the State of Nature." *Con-Textos Kantianos* 5: 167-189. Available online at URL = <<https://revistas.ucm.es/index.php/KANT/article/view/89975>>.

(Hanna, 2018a). Hanna, R. *Deep Freedom and Real Persons: A Study in Metaphysics*. THE RATIONAL HUMAN CONDITION, Vol. 2. New York: Nova Science. Available online in preview at URL = <[https://www.academia.edu/35801857/The Rational Human Condition 2 Deep Freedom and Real Persons A Study in Metaphysics Nova Science 2018](https://www.academia.edu/35801857/The_Rational_Human_Condition_2_Deep_Freedom_and_Real_Persons_A_Study_in_Metaphysics_Nova_Science_2018)>.

(Hanna, 2018b). Hanna, R. *Kant, Agnosticism, and Anarchism: A Theological-Political Treatise*. THE RATIONAL HUMAN CONDITION, Vol. 4. New York: Nova Science. Available online in preview at URL = <[https://www.academia.edu/36359665/The Rational Human Condition 4 Kant Agnosticism and Anarchism A Theological-Political Treatise Nova Science 2018](https://www.academia.edu/36359665/The_Rational_Human_Condition_4_Kant_Agnosticism_and_Anarchism_A_Theological-Political_Treatise_Nova_Science_2018)>.

(Hanna, 2020). Hanna, R. "Will-Power: Essentially Embodied Agentive Phenomenology, By Way of O'Shaughnessy." In (Erhard and Keiling, 2020: ch. 21, pp. 312-333). Available online in preview at URL = [https://www.academia.edu/36883518/Will-Power Essentially Embodied Agentive Phenomenology By Way of OShaughnessy](https://www.academia.edu/36883518/Will-Power_Essentially_Embodied_Agentive_Phenomenology_By_Way_of_OShaughnessy)>.

(Hanna, 2023a). Hanna, R. "In Defence of Dignity." *Borderless Philosophy* 6: 77-98. Available online at URL = <https://www.cckp.space/single-post/bp6-2023-robert-hanna-in-defence-of-dignity-77-98>>.

(Hanna, 2023b). Hanna, R. "Dignitarian Post-Capitalism." *Borderless Philosophy* 6: 99-129. Available online at URL = <https://www.cckp.space/single-post/bp6-2023-robert-hanna-dignitarian-post-capitalism-99-129>>.

(Hanna, 2024a). Hanna, R. *Science for Humans: Mind, Life, The Formal-&-Natural Sciences, and A New Concept of Nature*. Berlin: Springer Nature. Available online in preview at URL = [https://www.academia.edu/118055113/Science for Humans Mind Life The Formal and Natural Sciences and A New Concept of Nature Springer Nature 2024](https://www.academia.edu/118055113/Science_for_Humans_Mind_Life_The_Formal_and_Natural_Sciences_and_A_New_Concept_of_Nature_Springer_Nature_2024) >.

(Hanna, 2024b). Hanna, R. "The Death of Dennett and The End of Materialism." Unpublished MS. Available online at URL = [https://www.academia.edu/117893526/The Death of Dennett and The End of Materialism April 2024 version](https://www.academia.edu/117893526/The_Death_of_Dennett_and_The_End_of_Materialism_April_2024_version) >.

(Hanna, 2024c). Hanna, R. "Trump, Truth, and Logic." Unpublished MS. Available online at URL = [https://www.academia.edu/123816347/Trump Truth and Logic November 2024 version](https://www.academia.edu/123816347/Trump_Truth_and_Logic_November_2024_version) >.

(Hanna, forthcoming). Hanna, R. *Digital Technology for Humans: The Myth of AI, Human Dignity, and Neo-Luddism*. Berlin: De Gruyter-Brill.

(Hanna and Maiese, 2009). Hanna, R. and Maiese, M. *Embodied Minds in Action*. Oxford: Oxford Univ. Press. Available online in preview at URL = [https://www.academia.edu/21620839/Embodied Minds in Action](https://www.academia.edu/21620839/Embodied_Minds_in_Action)>.

(Hawkins and Allen, 1991). Hawkins, J.M. and Allen, R. (eds.), *The Oxford Encyclopedic English Dictionary*. Oxford: Clarendon/Oxford Univ. Press.

(Horgan and Nida-Rümelin, 2020). Horgan, T. and Nida-Rümelin, M. "On the Satisfaction Conditions of Agentive Phenomenology: A Dialogue." In (Erhard and Keiling, 2020: ch. 19, pp. 264-299).

(Kant, 1784/1996). Kant, I. "An Answer to the Question: What is Enlightenment?" Trans. M. Gregor. In I. Kant, *Practical Philosophy*. Cambridge: Cambridge Univ. Press. Pp. 17-22 (Ak 8: 35-42).

(Sartre, 1943/1956). Sartre, J.-P. *Being and Nothingness*. Trans. H. Barnes. New York: Washington Square Press.

(Wikipedia, 2024). Wikipedia. "Marvin the Paranoid Android." Available online at URL= https://en.wikipedia.org/wiki/Marvin_the_Paranoid_Android.