JNphi

Freedom under Naturalistic Dualism

Arturo Macías

Abstract

This article explores the issue of freedom in the framework of naturalistic dualism. Given that for physicalism the Universe is no more than the execution of the mechanical (either deterministic or stochastic) laws of Physics, reality is chance and necessity. On the other hand, consciousness is immediately real. By evaluating the set of possible futures conditional on their own actions a subject builds a legitimate mental object that is the scope of their "freedom". The existence of free will depends on the fact that the future cannot be remembered. The philosophical relevance of a better scientific understanding of time asymmetry is underscored and it is conjectured that it is related to the existence of fundamental (ontic) randomness in the fundamental Laws of Physics. The article closes with a short discussion on the moral consequences of this vision: the abhorrence that evil deeds produce is justified in absence of causative agency because they signal an unworthy conscious perpetrator.

Key Words: free will, naturalistic dualism, time asymmetry

DOI: 10.5281/zenodo.14272517

Introduction

This article discusses the classic issue of free will under naturalistic dualism. The literature on free will (O'Connor and Franklin, 2022; Timpe, 2016; Stewart, 2012) revolves around the tension between agency and either physical or metaphysical necessity. Those who consider that free will is a legitimate concept in a physicalist Universe are named "compatibilists", while those who don't are named incompatibilists.

The position here defended is that the free will problem is inexistent under naturalistic dualism. Agency is a part of the subjective realm, and is real as such. Moreover, given that consciousness is the most ontologically dense part of reality,

Corresponding author: Arturo Macías, Ph.D.
Address: c/Alcala, 522-Madrid (28.027)
e-mail ✉ arturo.macias@bde.es

subjective realities are more unproblematically real than the physical world itself. By evaluating (however imperfectly) the set of possible futures conditional on its own actions, the subject builds a mental object that is what it is defined as the scope of their "freedom". That subjective map is a legitimate mental object. On the other hand, consciousness is epiphenomenal, and arises from the autonomous physical reality, so (subjective) choice is real, but as the rest of the Universe, mechanistically determined.

While the philosophical problem is inexistent, free will depends on a very definite characteristic of reality: the fact that the future cannot be remembered. Time asymmetry is not properly understood in modern physics: the thermodynamic arrow of time is not clearly related to the information asymmetry that exists between past and future. It is conjectured that the future cannot be remembered because some of the fundamental laws of Physics imply fundamental (ontic) randomness.

The article closes with a short discussion on moral responsibility, where it is argued that it does not require causative agency. The abhorrence that evil deeds produce is granted because they signal an unworthy (or at least less worthy) conscious perpetrator.

In Section 1 of this article naturalistic dualism is summarized. In section 2 the definition of freedom in that framework is presented. Section 3 discusses the relation between time asymmetry and agency. Section 4 sketches how the proposed definition of free will deals with moral responsibility and Section 5 concludes.

Naturalistic dualism

This section describes naturalistic dualism (Chalmers, 1996; Gertler, 2020). Naturalistic dualism postulates the existence of two ontological domains of reality: first it is the conscious subject, or to be more exact the flow of consciousness over time, whose existence is immediate (it is the philosopher himself, and hopefully also the gentle reader). Since Descartes, the stream of consciousness has full ontological legitimacy (thinking means being immediately), and although memory is fallible and personal identity somewhat volatile (Minsky, 1986), even lies and errors, when they occur as a part of the conscious experience, have the legitimacy of "being" in the full sense of the word.

On the other hand, Modern Science since Newton postulates (with immense explanatory success) the existence of an irrational, automatic, and objective matter, whose behavior can be described in mathematical terms. Within this framework, Science is divided into two tasks: the discovery of the fundamental laws of Physics that can be mathematically expressed, but not explained (hypotheses non fingo), and the "scientific explanation" that consists of reducing observable phenomena to an application of the fundamental laws of

Physics. The scientific explanation thus understood is recursive, with the most complex phenomena explained in terms of the simpler in a hierarchy (the so-called reductionist hierarchy) where Biology[2] is based on Chemistry and Chemistry is determined by Physics (the laws of Chemistry are mainly a consequence of the quantum properties of the electron orbitals). In fact, the "reduction" begins in Physics itself, where heat and its properties are explained as an emerging phenomenon in the so-called Statistical Mechanics that provides the micro-foundation of Classical Thermodynamics.

Laplace's synthetic summarization of this worldview remains unsurpassed: "*We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which at any given moment knew all of the forces that animate nature and the mutual positions of the beings that compose it, if this intellect were vast enough to submit the data to analysis, could condense into a single formula the movement of the greatest bodies of the universe and that of the lightest atom; for such an intellect nothing could be uncertain and the future just like the past would be present before its eyes.*"

Laplace describes a deterministic Universe where that postulated intellect (generally known as "Laplace's demon") can perfectly know the future, but in Quantum Mechanics, physical prediction becomes probabilistic (and "positions and velocities are substituted by wave functions). By their very nature, the laws of Physics are "mechanical": physicalism is precisely the idea that the evolution of the Universe at all scales is fully determined by the execution of the mechanical laws that rule elementary particles and fields. When those mechanical laws allow for complete prediction, and no stochastic variables are needed we are in the particular case of deterministic mechanism.

For Descartes the mechanistic *res extensa* ended in the pineal gland, where the conscious supernatural soul (res cogitans) joined the automaton body. By contrast, Laplace's demon owes his omniscience to the autonomy of matter. Laplace suggests that there is no soul connected to the pineal gland and that Descartes is as much an automaton as the rest of the Universe. This position is anti-intuitive, but its denial implies that the laws of Physics, with their mechanical concatenation of causes and effects are not and cannot be complete, and that animism is justified in the last citadel of the human brain. That the atoms of the brain know where they are and have the courtesy to alter their general behavior when located in the cranial cavity, or that they are subjected there to the influence of a supernatural substance.

In any case, whether the soul descends from heaven or emerges from the mud, consciousness remains entirely real. On the other hand, if matter is autonomous, the mind is super-impressed on

[2] From Biochemistry to Ecology (through Cell Biology, Histology, Physiology...) this science encompasses multiple layers of reduction itself.

reality, perfectly synchronized, and parallel to it. This position on the matter-consciousness relationship is called "emergentist epiphenomenalism".

Emergentist epiphenomenalism is not a purely metaphysical proposition, because it implies that the assignment of conscience is not a necessary intermediate computation for the Laplace demon in order to predict the position and velocity of the particles (more exactly its wave functions) in the future. On the other hand, in the unlikely case that conscience plays an active role in the Fundamental Laws of Physics (and the Laplace demon needs to assess conscience to properly forecast the future of some Physical systems), the Laws of Physics involving sentience attribution could be reversed engineered to assess conscience of physical systems, allowing for the falsification of epiphenomenalism (De Barros & Oas, 2017; Yu & Nikolic, 2011). In the remaining of this paper it is supposed that conscience does not alter subatomic interactions and *consequently* it is epiphenomenal.

David Chalmers devotes many pages and several thought experiments to argue that consciousness cannot be "reduced" to a physicalist explanation like the one that brought down vitalism, but it is more pertinent to point out that there is nothing to reduce. Explaining the functioning of physical systems (for example, of living beings) as a consequence of the laws of Physics is part of Science, but if the emergentist epiphenomenal hypothesis is true, the most minute mechanical description of any phenomenon is entirely divorced from any attribution of sentience.

Understanding why some physical systems make an emergent consciousness appear (the so called “hard problem of consciousness”) or finding a procedure that quantify the intensity of consciousness emerging from a physical system (the so called “pretty hard” problem of consciousness) is impossible under naturalistic dualism: the Laplace's demon could know the exact future evolution of each neuron in Descartes's brain without this giving him the slightest information about whether Descartes thought and existed as a conscious subject.

In fact, the demon himself would not even know whether computing Descartes's future evolution would generate the realization of his conscious experience. If the answer were affirmative, only the “simulated Descartes” himself would know it (!), and if it were negative then nobody would. In naturalistic dualism consciousness is the ultimate noumenon.

In a memorable statement Epstein (2006) describes the scientific reductionist enterprise as “growing” the macrostructure from a micro-specification. Exact replication from mechanical rules is the maximum attainable goal for Science. The fact that even having a complete generative model does not contribute much to the discussion about the consciousness of artificial neural networks shows in practice the

limits of even perfect phenomenal knowledge to assess the existence of (noumenal) consciousness.

In fact, the “neural correlates of consciousness” research agenda (Koch, Massimini, Boly and Tononi, 2016) substantially depends on our trust on human subjective reporting. The extension of this methodology to animals is undermined for their lack of language, which impedes reporting. On the other hand, for computers, even superhuman cognitive and linguistic skills would be not enough to guarantee conscience, because the specific physical implementation of a neural network (and not only her outputs) is likely important for the emergence of conscience (see Marshall, Albantakis, Mayner, Koch, and Tononi, 2019).

The autonomy of matter and the fact (observable by the subject himself) that the flow of consciousness is synchronized with the evolution of the material system from which it emerges, reverses the Cartesian worldview and makes the mind a passive automaton (an epiphenomenon) of the brain. This article defends that the concept of freedom of a conscious being is well defined even within the framework of naturalistic dualism.

Definition of freedom

Suppose a conscious physical system (eg. a human brain). The dynamics of that physical system is as determined by the laws of Physics as the rest of the universe, but having an attached consciousness flow creates a well-defined boundary between that subset of the universe and the rest. Indeed, one essential characteristic of human consciousness is that it is a unitary experience, which emerges from a determined physical system (Massimini and Tononi, 2013).

The physical system from which consciousness arises may be connected to a device (named "body") that serves it as an interface to affect external reality. The conscious (epiphenomenal) subject is aware of their own ability to affect “external” reality and the idea of freedom is based on that perception. By evaluating (more or less imperfectly) the set of possible futures conditional on their own actions, the subject builds a mental object that is the scope of their "freedom".

Let’s focus on the scenario of a deterministic universe to simplify the exposition. We divide reality into two distinct and arbitrary sets that encompass the entire cosmos: “Brain” and “Rest.” Let us assume that Laplace’s demon possesses complete information about the set “Rest,” while remaining entirely ignorant of the details within “Brain.” Within this framework, the demon can compute all conceivable futures of the universe for every conceivable configuration of “Brain.” This ensemble of potential futures reveals what we term the “degrees of freedom of Brain”—essentially, how this specific subset of reality can

exert influence over the entire Universe. Although this process may appear cumbersome, it remains a legitimate and mechanistic exercise.

If “Brain” is not merely an arbitrary component of the Universe but rather a locus of unified conscious experience, the set of possible futures conditional on the configurations of “Brain” is the real set of “degrees of freedom” of the conscious being emerging from “Brain”. The self-assessed scope of possible futures among which a conscious being choses approximates those “degrees of freedom of Brain”, a well-defined materialist object defined in the previous paragraph with the help of the Laplace’s demon. The “degrees of freedom of Brain” describes the scope of “Brain’s” freedom (and the self-assessed scope of freedom is a subjective object approximating it), but the use that “Brain” will make of its freedom is determined by the Laws of Physics.

If the conscious physical system is volitional, it also has a subjective assessment about the desirability or utility of future states of the world. Laplace's demon (phenomenologically) observes the action of the conscious physical system as the mere ordinary application of the laws of Physics, but from the perspective of the epiphenomenal subject that action is an optimization of the universe for their own ends.

The reader may object that what has been argued so far does not refute that freedom is an illusion (the scope freedom is placed in the subjective domain), and indeed, we believe it is. Freedom is as much an illusion as the color red, the set of natural numbers or the statement that in chess the bishop stays on squares of a given color.

All the objects of our mind are illusions, and at the same time they are the only thing that fully and immediately exists. Among those objects, however, freedom is well defined, and linked to physical reality in a direct way. It is much more like a Hamiltonian than like a unicorn. It is true that, as opposed to a Hamiltonian, the self-assessed domain of free action is not a purely physical description of reality: the conscious subject does not attempt to physically predict their own action, but considers it as a "free variable". But consciousness is real and the subject makes not mistake when regards their self as a volitional unit.

On the other hand, the ability of the conscious being to think rationally to improve the universe is seen by the Laplace's demon as a mere mechanical consequence of a given state of the matter and of the laws of Physics, and of course, he is right too.

Time and consciousness

Chalmers explores in detail the matter-consciousness relationship through a mental object called a "philosophical zombie ", a human being identical to those that exist, but without consciousness. On the other hand, the experiment on what the absence of freedom means

does not involve any hypothetical exercise. Freedom only exists towards the future, so regarding the past we are all philosophical "free will" zombies: if freedom is the set of possible "states of the world" conditional on the subject's actions, and the subject's actions cannot affect the past, freedom with respect to the past is inexistent.

Our intuition about time is clear: there is a unique past that we remember over which human action is impossible; there is a present instant of consciousness in the Cartesian sense, and there is the future that is unknown and that can be affected by the present action. The experience of these three times is essential in our perception of the world and the structure of language.

Contemporary Physics and the Philosophy of Physics lack consensus regarding the reason why remembering the future is impossible (Hemmo and Shenker, 2022). In our view, this stands as the most philosophically significant unresolved scientific question. The thermodynamic time arrow is not related in any obvious manner to the massive information asymmetry between future and past that characterizes our experience. Additionally, it is extremely unlikely that such a pervasive and structural feature of reality comes from a mere biological limitation.

We risk in this matter a (dangerous) scientific hypothesis, whose exploration is left to others more qualified: the reason why the future cannot be remembered is because there is fundamental (ontic) randomness in the Universe. Any random experiment generates irreversibility: towards the past the random process has been carried out (we know all the realizations of the involved random variables), and towards the future we only have probability distributions of the (ontically) random variables.

In the architecture of modern physics, Born's rule turns quantum mechanical observation into a random experiment, but whether quantum measurement is truly ontic randomness is still an open issue (see Nath Bera, Acín et al., 2017 on quantum randomness and more generally Maudlin, 2019 for a philosophical inquiry on Quantum Mechanics).

From the perspective of distinguishing between the past and the future it is irrelevant whether the realization of a random experiment selects one event among several possibilities, or if it splits reality into several branches (as Everett's interpretation of Quantum Mechanics suggests). In both cases ontic randomness generates the possibility of several futures, and with that possibility the reversibility of the fundamental laws of the Universe is broken and remembering the future becomes impossible. The void left open by the impossibility of remembering the future is where the scope of freedom appears.

Responsibility in a mechanist Universe

Readers well-versed in the philosophical literature on freedom might notice the absence of an analysis regarding the implications of our definition of freedom for the freedom-responsibility dilemma. Up to this point, this article has primarily delved into descriptive philosophy, and the freedom-responsibility interplay can only be adequately addressed within a normative context. Consequently, while this issue will not be explored in detail, entirely avoiding it could imply a lack of confidence in the practical applicability of our proposed definition of freedom.

The folk intuition on moral responsibility is that the determination of behavior has a "strong part" that encompasses genetic inheritance, extreme nurture conditions, or neurological damage, and to the extent that behavior can be explained based on those strong determinants it is considered outside of the scope of moral responsibility. The rest of the behavior not explained by the strong determinants is what is considered susceptible of moral responsibility.

In this article, we have argued that freedom is not "absence of determination"; rather, it is defined as behavior that falls under the control of a conscious entity. Using Helen Stewart definition (Stewart, 2012), agency is the "power to act". A conscious being that can choose among several options has the power to act, no matter how determined is the way it effectively uses that power (vg. if you offer poison or food to one hundred people, and all of them chose the food, their unanimity detracts nothing from their "power to act"). This idea was advanced by Schopenhauer (1969 [1839]) in a memorable statement: "*Man can do what he wills but he cannot will what he wills",* which is vindicated under naturalistic dualism.

Beyond innocuous choice, on the moral realm, when a conscious being performs immoral acts, our condemnation does not come merely from the social desirability of the punishment, but mainly from the horror that a conscious being performs the considered acts. That intentionality comes down from heaven or emerges from the mud does not change that it is intention anyway. The fact that evil intention comes from a configuration of matter does not reduce the horror that the existence of that kind of corrupted conscience produces on those that are not so corrupted.

To clarify this position, let's consider three cases: i) A runs over his wife in a traffic accident, ii) B is a person with a normal and non-violent behavior until a brain tumor is found; then his behavior turns violent and he kills his wife, iii) despite being born into a middle-class family, and having no known problematic relatives, C develops increasingly violent behavior from puberty that culminates in him killing his wife. Folk theories of responsibility consider A and B fundamentally not responsible and C strongly responsible.

Now let's apply the proposed definition of freedom to the case considered before. While folk moral responsibility acquits B because he is not guilty of developing a tumor, in our view B and C are equally responsible because the path that turns you into a monster does not change the monster you have become. A's unintentional act is innocent, since it tells us nothing about the moral value of his current conscious self. The case of B is totally different. There was a time when B was a good person, but the tumor has already killed that good person, bringing about murderous post-tumor B. The intuition that post-tumor B must be absolved of his responsibility because he is not the true B is only an animistic fallacy.

The existence of consciousness is what gives moral relevance to physical systems. Any attribution of moral value can only be made to a conscious being (only the conscious being can be a moral object), and at the same time moral responsibility is only required from the conscious being (only the conscious being can be a moral subject). To a certain degree this allows for the metaphysical grounding of moral reciprocity: the immoral conscious subject excludes himself (to the extent of its immorality) from the moral circle to which his own conscience gives him access to. Of course, other considerations are relevant for exclusion from the moral circle, as the degree of hazard that the considered agent can pose to others (reciprocity arguments for moral exclusion are weakened when the excluded can instead be easily neutralized).

Conclusion

In this article the issue of freedom in the framework of naturalistic dualism is explored. Given that for physicalism the Universe is no more than the execution of the mechanical (either deterministic or stochastic) laws of Physics, reality is not more than chance and necessity.

On the other hand, under naturalistic dualism consciousness is still real and legitimate (more ontologically legitimate than anything else). By evaluating (more or less imperfectly) the set of possible futures conditional on their own actions, a conscious subject builds a mental object that is what we define as their scope of "freedom". If agency is "power to act", a conscious being that can choose among several options has the "power to act" no matter how determined is the way it effectively uses that power.

An important observation regarding freedom is its relation with time: in our Universe, where past is remembered and future unknown, freedom is always related to the future. It has been argued that the physical basis of the time asymmetry is the most important open issue in Physics from a philosophical point of view, and it is conjectured that it is related to the existence of ontic randomness in the fundamental laws of Physics.

The proposed definition of freedom is used to deal with the liberty-responsibility dyad. On the moral realm, when a conscious being performs an immoral act, our condemnation does not stem merely from the social desirability of the punishment, but mainly from the horror that a conscious being is capable of performing the considered acts.

The existence of consciousness is what gives moral relevance to physical systems. To a certain degree this allows for the metaphysical grounding of moral reciprocity: the immoral conscious subject excludes himself (to the extent of its immorality) from the moral circle to which his own conscience gives him access to.

References

Acacio de Barros J, Oas G. Can we falsify the consciousness-causes-collapse hypothesis in quantum mechanics? Foundations of Physics 2017; 47(10): 1294–1308.

Chalmers D. The Conscious Mind: In Search of a Fundamental Theory. Oxford University Press; 1996.

Epstein JM. Generative Social Science: Studies in Agent-Based Computational Modeling. Princeton University Press; 2006.

Gertler B. Dualism: How Epistemic Issues Drive Debates About the Ontology of Consciousness. In: The Oxford Handbook of the Philosophy of Consciousness. Oxford University Press; 2020.

Hemmo M, Shenker O. The Second Law of Thermodynamics and the Psychological Arrow of Time. The British Journal for the Philosophy of Science 2022.

Koch C, Massimini M, Boly M, Tononi G. Neural correlates of consciousness: progress and problems. Nature Reviews Neuroscience 2016.

Marshall W, Albantakis L, Mayner W, Koch C, Tononi G. Dissociating Intelligence from Consciousness in Artificial Systems - Implications of Integrated Information Theory. Proceedings of the 2019 Towards Conscious AI Systems Symposium, AAAI SSS19; 2019.

Massimini M, Tononi G. Sizing Up Consciousness: Towards an Objective Measure of the Capacity for Experience. Oxford University Press; 2013.

Maudlin T. Philosophy of Physics: Quantum Theory. Princeton University Press; 2019.

Minsky M. The Society of Mind. Pocket Books; 1988.

Nath Bera M, Acín A, Kuś M, Mitchell MW, Lewenstein M. Randomness in Quantum Mechanics: philosophy, physics and technology. Reports on Progress in Physics 2017.

Steward H. The Metaphysical Presuppositions of Moral Responsibility. The Journal of Ethics 2012.

Schopenhauer A. Essays on Freedom of the Will. Wiley-Blackwell; 1969.

Timpe K. Leeway vs. Sourcehood Conceptions of Free Will. In: The Routledge Companion to Free Will; 2016.

O'Connor T, Franklin C. Free Will. In: The Stanford Encyclopedia of Philosophy; 2022.

Yu S, Nikolić D. Quantum mechanics needs no consciousness. Annalen der Physik 2011; 523(11).