

# THE MIND–BODY PROBLEM AND LIBERTARIAN FREEDOM OF THE WILL

**Jonas Dagys**

Vilniaus universiteto  
Filsofijos istorijos ir logikos katedra  
Universiteto g. 9/1, LT-01513 Vilnius  
El. paštas: jonas.dagys@fsf.vu.lt

*The paper analyzes the relation between the mind body problem and the problem of free will within the physicalist framework. It is argued that the problem of free will, as far as metaphysics is concerned, is a specific case of the mind–body problem, namely the problem of how can mental events have physical effects. One need not be eliminativist to be a physicalist about the mind – reductionism or epiphenomenalism is perfectly compatible with the causal closure of the physical world, while preserving certain reality of the mental, but neither epiphenomenalist nor reductionist strategy can save libertarian freedom. If this is in fact so, then somewhat more general conclusion follows: physical determinism and fatalism do not exhaust the opposition to libertarians, and thus the falsity of the two would not in any case suffice to establish libertarian position as true.*

**Key words:** *mind–body problem, freedom of the will, physicalism, mental causation.*

In this paper, I am going to relate two often unrelated and independently pursued metaphysical problems – the mind–body problem and the problem of free will. Even though both problems are genuine metaphysical problems, their development has gone in two slightly different directions – discussion of the former more often invokes neurophysiology and psychology, while the latter is usually discussed in ethical and legal contexts. I am going to argue that the problem of free will, as far as metaphysics is concerned, in fact relates to mind-body problem as its more specific case – namely the problem of how can mental events have physical effects, and that the answer to the problem of free will is in a way

dependent on the way metaphysical relation of the mental and the physical is depicted. Having drawn this parallel I will try to transfer the problem of finding a space for mind in the natural world to the discourse about free will and attempt to sort out the theoretical possibilities open for someone of libertarian inclinations.

Most often the problem about the mind is that of finding a slot for the mind in the world of events that is causally closed under physics. The empirical thesis stating causal closure of the physical is not uncontroversial and not without its problems, but we seem to have “good reason to believe the empirical thesis that all physical effects are due to physical

causes” (Papineau 2000). I will argue that one can accept this thesis and still be a realist about the mental (and even some sort of mental causation). However, this path is blocked for the defender of the libertarian free will. One need not be eliminativist to be a naturalist about the mind – reductionism or epiphenomenalism is perfectly compatible with the causal closure of the physical world, while preserving certain reality of the mental. But neither epiphenomenalist nor reductionist strategy can save libertarian freedom. If this is in fact so, then somewhat more general conclusion follows: physical determinism and fatalism do not exhaust the opposition to libertarians, and thus the falsity of the two would not in any case suffice to establish libertarian position as true.

The problem under discussion could be investigated from the perspective of compatibility of the idea of free will and free action with our knowledge of natural world as supplied by physical sciences. Put it briefly, it could be expressed by the following argument:

1. If the natural world is such as suggested by physical sciences, it is not hospitable to free will; or, alternatively, if some world is closed under physics, then there is no such thing as free will in that world.
2. It is very likely that the natural world we live in is closed under physics.

Therefore

3. It is very likely to believe that there is no free will.

The logical form of the argument is nothing more than classical *modus ponens*, and it should not be questioned by anyone who allows for the application of basic laws of logic in this context. What is needed for establishing the conclusion of the argument is assigning

truth-values to the premises. In what follows I will supply my reasons for believing that the premises are true and thus that the argument is sound. I do not aim at convincing or converting anyone, but rather at deriving *prima facie* counterintuitive position.

In his book in progress Steven Horst advances the familiar thesis that “there is a kind of gulf between our discourses about the world of nature and our discourses about ourselves” and that “trying to bridge this gap is the most fundamental problem of modern philosophy” (Horst: forthcoming). Horst himself seems to be sceptical about the success of naturalistic reductionist programme in this field, as he is convinced that “there are mentalistic properties [call them] M, that cannot be naturalized” (ibid.). Similar claims, namely, that the explanatory resources of our natural sciences are in principle less than sufficient for explaining the phenomena of (or related to) consciousness can be found in works of philosophers like David Chalmers, Frank Jackson, Hubert Dreyfus, physicists like Brian Pippard and others.

It seems to be fairly uncontroversial that the mental domain is also the seat of what is usually referred to as free will, and that free will is one of the ways consciousness manifests itself. As Roger Penrose puts it, “if you use your consciousness in some way to influence what you do – that’s what we call free will” (Penrose 1999). But the problem with free will is not only that it cannot be suitably explained by, or reduced to, some natural phenomena, but that the very idea of it contradicts the present picture of the physical world.

This stems from two sources. On the one hand, physical sciences supply us with effective and reliable descriptions and explanations of the events in the world around us. In others

words, scientific theories predict and explain occurrences of physical events, and, unless one would go as far as denying that at least some of the events taking place are physical, she is committed to accepting physical theories as true (at least in the pragmatic explanatory sense). On the other hand, there seems to be some very strong evidence, derived from first person phenomenological experience, that sometimes we actually exercise our ability to choose freely in virtue of our capacity of free will, and that at least some of our actions are free and undetermined, or that we ourselves (whatever that means) are the sole origins of these actions. The problem arises when we want to put the two beliefs in conjunction, and express the idea of acting freely within the physical world.

It is usually agreed that there are what some call the objective correlates to consciousness: there are physical and chemical changes in one's brain and body that can be observed to be correlated (as causes or effects) or identified with changes in her conscious thoughts and experiences. These objective correlates to consciousness can be studied by the methods of science and (hopefully) will eventually be explained in terms of physics and chemistry of the brain and body. Given this, to be a realist about free will, one would need to find a way to assign real causal properties to this phenomenon (or else it would melt down to epiphenomenal appearance). A realist about free will is committed to showing that it actually influences events and has a role in their causal sequences. After all, to be taken as real free will must have effects! The problem of free will thus comes extremely close to the problem of mental causation, since possibility of specific mental causation is required by (if not identical to) what we take to be free will.

To proceed further a more explicit definition of what is understood by free will is needed. Usually freedom of the will is discussed in moral and legal contexts, where the concept appears to be central, and is understood in terms of responsibility for one's actions:

'Mental responsibility' pointed to a consideration of the extent to which the accused's mind is answerable for his physical acts; this must include a consideration of the extent of his ability to exercise will-power to control his physical acts. Inability to exercise will-power to control such acts entitles the accused to the benefit of the action' difficulty in controlling such acts may do so, if great enough to amount to substantial impairment (Gregory 1987: 194).

So free will can be defined literally, as the ability to choose freely any one of several more or less different actions, based on the conscious experience of a single given situation. For the purposes of this paper I will rely on the illuminating and precise formulation given by Kane (even though it is essentially equivalent to those provided by Galen Strawson or Martha Klein). Kane defines freedom of the will in terms of ultimate responsibility – the choice is free only if it is up to the agent, and that is the case only if the agent is ultimately responsible for that choice. (It is worth noting that responsibility here is devoid of any subjective connotations and has little to do with duty, guilt and the like, and therefore it is used with the term 'agent' rather than 'subject'). So, we have two requirements – alternative possibilities and ultimacy of that responsibility: As Kane puts it

An agent is *ultimately* responsible for some (event or state) E's occurring only if (R) the agent is personally responsible for E's occurring

in a sense which entails that something the agent voluntarily (or willingly) did or omitted, and for which an agent could have voluntarily done otherwise, either was, or causally contributed to, E's occurrence and made a difference to whether or not E occurred; and (U) for every X and Y (where X and Y represent the occurrences of events and/or states) if the agent is personally responsible for X, and if Y is and *arche* (or sufficient ground or cause, or explanation) for X, then the agent must also be personally responsible for Y. (Kane 1996: 35)

This is what is called libertarian free will, and it draws upon our strong intuition that human beings are free to decide their own actions. During the break at the opera one is free to choose whether can have a glass of wine, smoke a cigarette or chat with her companions. Some particular choice is always made, but whatever the choice one is free in the strong sense since she could have acted otherwise and the history of the universe would have gone different. This intuition suggests that at least some part of this history is up to us, and we have a significant role as it rolls through time.

Libertarian free will is incompatible with physical determinism, which asserts that the future of the universe (at least its physical core) is completely determined by its physical state at a given time. Ever since the development of Newtonian physics in XVII century, the problem of free will was precisely the one of locating this freedom within the universe governed by deterministic laws – by deciding the possibility of choices satisfying the requirement of ultimate responsibility. Since all the laws in Newtonian physics were deterministic in character, it suggested that the universe itself was deterministic.

It takes a simple argument to prove the impossibility of free action within deterministic universe (the standard version is due to

van Inwagen). Put briefly, it states that if no agent is free to influence the past and the laws of nature, and if both of the latter imply the future, then no agent is free to influence the future, and hence – no free will:

If determinism is true, then our acts are the consequences of the laws of nature and events in the remote past. But it is not up to us what went before we were born; and neither is it up to us what the laws of nature are. Therefore, the consequences of these things (including our present acts) are not up to us. (van Inwagen 1983: 16)

It was not until the rise of quantum mechanics with its probabilistic laws in the first half of the XX century that some hope was given to the proponents of libertarian free will: it appeared that it might be possible to accommodate consciousness and freedom within the physical picture of the world. As Chalmers observes:

Indeed, it can seem that quantum mechanics provides about as perfect a causal role for consciousness as one could imagine in a physical theory. Any indeterminism in quantum mechanics comes in at the point of 'collapse', which on the most common interpretation is triggered by a 'measurement', and it can seem that consciousness is the only non-arbitrary way to distinguish a measurement from other physical events. If so, then consciousness may be present in quantum mechanics' very foundations. (Chalmers 1997: 403)

The idea of fitting freedom into the 'gaps' opened by indeterministic laws of quantum mechanics was to establish an analogy or equivocation between making a free decision and an act of quantum mechanical measuring or occurrence of the quantum superposition break – the events that do not seem to be deterministically caused under physical laws,

since these laws are indeterministic, could be claimed to be sufficiently caused by conscious decision or mental act. But the difference between the world governed by deterministic laws and the one closed under indeterministic laws makes a rather apparent than genuine difference when the problem of mental causation or causal efficacy of free decision is considered. According to Kane:

one often hears the argument in contemporary free will debates that if quantum jumps or other undetermined events did sometimes have non-negligible effects on the brain or behavior, this would be of no help to defenders of an incompatibilist free will. Such undetermined effects would be unpredictable and uncontrollable by the agents ... just the opposite of the way we envision free and responsible actions. (Kane 2002: 8)

However, there is yet another problem that could be raised for grounding free choice in quantum indeterminacy – the problem of the relation between physical and mental causes. The problem becomes apparent, when one observes that quantum indeterminacy is governed by laws: probabilistic laws, but still laws in the strict sense. So choices that are rooted in quantum indeterminacy are free, but since the scope of freedom corresponds to the scope of probabilities assigned, these choices are not free in the way libertarians would have them.

First, the solution does not work because in quantum mechanics the principle of the physical closure is in force (except for some improbable dualistic interpretations, e.g. according von Neumann's and Wigner's interpretation it is human consciousness itself that causes the superposition break during the measurement). So deterministic causation does not give way to anarchy, it is replaced by

probabilistic causation, and all events within the quantum mechanical system are assigned objective probabilities. If we suppose that the agent's choice is grounded in indeterministic process (say superposition break of an electron within *tubulin dimer*, a complex protein which is the basic unit of which microtubules of the brain are built, would result in a certain choice to go the opera or to the pub), then the outcome of the choice will be subject to the same probabilities. And if we consider the long run of trials on the same choice, the statistical result should come close to these objective probabilities. So the indeterministic choices would be determined in a certain sense – determined by the objective probabilities and corresponding statistical distributions.

As a consequence, the objective probability of the occurrence of event can only be changed by changing the probability of its cause, and the objective probability of that can only be changed by altering the probability of its cause, and so on: slightly augmented version of van Inwagen's argument yields the same result as the one obtained for deterministic systems – no agent is free to affect the future, since no agent is free to modify the laws of nature or the states of universe before her birth. Thus the problem of free will can be seen as the problem of how to fit it into a world governed by natural laws, regardless of whether these laws are deterministic or not:

Since all of the surface features of the world are entirely caused by and realised in systems of microelements, the behavior of microelements is sufficient to determine everything that happens. Such a 'bottom up' picture of the world allows for top-down causation (our minds, for example, can affect our bodies). But top-down causation only works because the top level is already caused by and realized in the bottom levels. (Searle 1984: 94)

Yet another problem is that even if quantum indeterminacy would open the required space for freedom in the natural world, one could deny that quantum indeterminism is metaphysically rather than merely epistemically probabilistic. There are famous difficulties with the explanations based on statistical laws explored by Alberto Coffa. The problem here is that in contraposition to deductive-nomological explanations based on deterministic laws where “premises identify certain features of the world that are nomically responsible for the explanandum event <...> there are no things [that could figure in statistical explanations] which go in non-epistemic world of facts that can explain the event” as probabilistic (Coffa 1974: 71). This allows physicists like Steven Weinberg to claim that “in quantum mechanics there is still a sense in which the behavior of any physical system is completely determined by its initial conditions and the laws of nature” (Weinberg 1992: 37).

So the story goes. The proponents of free will seem to be left with the obscure ideas of agent causation, that invoke some mysterious way of ‘causing’ events without affecting their objective probabilities. According to Timothy O’Connor, “the agency theory affirms the completely general claim that objects have causal powers in virtue of their properties, so that objects sharing the same properties share the same causal capacities, but it denies that all such causal powers may be thought of as simple ‘functions from circumstances to effects’” (O’Connor 1995: 177). The tenet of the agency theory thus is the claim that there are two basic and fundamentally different sorts of causal properties, one of which applies exclusively to purposive agents.

This doctrine was thoroughly criticized by Jaegwon Kim (1998 and elsewhere) on the

basis that it fails to provide an appropriate account of causal interaction between the two types of properties, e.g. mental and the physical:

Suppose that a certain [mental] event ... causes a physical event. The causal closure of the physical domain says that this physical event must also have a physical cause. ... *What is the relationship between the two causes, one mental and one physical?* ... Given that any physical event has a physical cause, how is a mental cause *also* possible? (Kim 1989: 280-1)

If the physical closure of the world is to be preserved, agent causation breaks down to either physicalist monism or non-interactionist dualism, and hence non-realism about free will. Mental properties have neither causal, nor explanatory significance, and it is hard to see how they could have any effect upon the physical goings-on.

Let me turn to the second premise of the initial argument – the plausibility of the belief that the world is closed under physics. It has to be acknowledged that the closure principle is empirically underdetermined – it is not a regular empirical hypothesis, since it cannot be verified nor falsified in principle. It is rather a fundamental methodological assumption underlying the research in fundamental natural sciences and enforcing the physical explanations, and as long as these sciences can be supplied (at least pragmatic) justification of their theories, the causal closure principle is justified. Whatever physicists on holiday or in retirement might say, any physical theory (and practice) is inconceivable without the assumption that the studied system is closed under universal laws. This is so, since gaps in nomological network would put the need and validity of explanations in jeopardy. And so, since it is rational (in pragmatic sense) to

accept physical explanations, it is also rational to accept the complementary principle of physical closure of the explained system. It might be claimed such justification for the principle of causal closure is circular – science assumes it as true, while the principle supports scientific theories. Deeper analysis of the discussed principle leads to the analysis of the conception and applicability of natural laws. But in any case, we have good empirical evidence to believe that the physical world is in fact closed under physics, and no substantial independent reasons against it.

So we are left with a dilemma – either mental states are physical states, or they aren't. If they in fact are, then they are causally efficacious in virtue of their physical properties, but by the same token they are closed under physical laws, and thus there is no space left for free decisions that could satisfy the requirements of ultimate responsibility. On the other hand, if mental states are not physi-

cal states, it is difficult to see how they could have real effects within the physical universe, and thus how they could be anything more than epiphenomena. If that is the case, then conscious decisions would have no physical effects either, and since normally it is in virtue of its effects that certain choice is judged to be good or bad, choices without specifically physical effects would not amount to a relevant type of choices that could support moral responsibility. Thus in either case, as long as one is sensitive to causal closure of the physical, one cannot consistently believe in freedom of the will – there is no account of libertarian free will that could be made compatible with the natural sciences. But this need not imply that closure of the physical, and thus eliminativism about free will, entails any form of antirealism of moral norms. One just has to find the appropriate way to reconceptualize responsibility, praise and punishment without recourse to free choice.

## REFERENCES

- Chalmers, D. 1997. "Moving Forward on the Problem of Consciousness", in Shear, J. (ed.) *Explaining Consciousness: The Hard Problem*. Cambridge: Cambridge University Press, 1997.
- Coffa, J. A. 1979. "Hempel's Ambiguity". Reprinted in *Explanation*, ed. D.-H. Ruben. Oxford: Oxford University Press, 1993.
- Gregory, R. (ed.) 1987. *The Oxford Companion to the Mind*. Oxford: Oxford University Press.
- Inwagen, P. van. 1983. *An Essay on Free Will*. Oxford: Oxford University Press.
- Horst, S. (forthcoming) *Mind and the World of Nature*. Manuscript in progress. Accessed on Nov. 20, 2007 at: <<http://shorst.web.wesleyan.edu/mwn/mwn.frames.html>>. Quoted with the kind permission of the author.
- Kane, R. 1996. *The Significance of Free Will*. Oxford: Oxford University Press.
- Kane, R. 2002. *The Oxford Handbook to Free Will*, Oxford, Oxford University Press.
- Kim, J. 1989. "The Myth of Nonreductive Materialism" in his *Supervenience and Mind: Selected Philosophical Essays*, Cambridge: Cambridge University Press, 1993: 265-284.
- Kim, J. 1998. *Mind in a Physical World*. Cambridge: MIT Press.
- O'Connor, T. 1995. "Agent Causation", in T. O'Connor (ed.) *Agents, Causes, and Events: Essays on Indeterminism and Free Will*, New York: Oxford University Press, 1995: 173-200.
- Papineau, D. 2000 "The Rise of Physicism". Accessed on Dec 04, 2007 at: <<http://www.kcl.ac.uk/ip/davidpapineau/Staff/Papineau/OnlinePapers/Risephys.html>>.

Penrose, R. 1999. "Science and the Mind". *Lecture delivered at Institute for Theoretical Physics, University of California*. Accessed on Jan 20, 2002 at: <<http://doug-pc.itp.ucsb.edu/online/lecture/penrose/>>.

Searle, J. 1984. *Minds, Brains and Science*. Cambridge, MA: Harvard University Press.

Weinberg, S. 1992. *Dreams of a Final Theory*. New York: Pantheon Books.

## ŠAMONĖS IR KŪNO SANTYKIO PROBLEMA IR LIBERTARINĖ VALIOS LAISVĖ

**Jonas Dagys**

Santrauka

Straipsnyje tiriamas dviejų metafizinių problemų – sąmonės ir kūno santykio bei valios laisvės – santykis fizikalizmo požiūriu. Teigiama, kad valios laisvė yra specifinis sąmonės ir kūno santykio problemos atvejis, o būtent klausimas, kaip psichiniai įvykiai gali turėti fizinių padarinių. Fizikalizmas nereikalauja būti eliminatyvistu visų mentalinių reiškinių atžvilgiu, nes reduktyvizmas arba epifenomenalizmas neprieštaruoja kauzaliniam fizinio pasaulio uždaramui. Ta-

čiau nei epifenomenalistinė, nei reduktivistinė strategija nėra suderinama su valios laisvės realumu. Jei taip, tai galima daryti ir bendresnę išvadą – fizinis determinizmas ir fatalizmas neišsemia libertarizmui priešingų pozicijų ir todėl pirmųjų paneigimas nepakankamas pastarosios teisingumui įrodyti.

**Pagrindiniai žodžiai:** sąmonės ir kūno santykis, valios laisvė, fizikalizmas, psichinis kauzalumas.

*Įteikta 2007 12 17*