

Compatibilism and Conscious Will

Kompatibilismus a vědomá vůle

Michaela Košová

Katedra filosofie a dějin přírodních věd

Přírodovědecká fakulta

Univerzita Karlova v Praze

Viničná 7, 128 44 Praha 2

michaela.kosa@gmail.com

Abstract/Abstrakt

Daniel Dennett's compatibilism based on redefining free will via broadening the concept of self to include unconscious processes seems to disappoint certain intuitions. As Sam Harris points out, it changes the subject from the free will we seem to intuitively care about – conscious free will. This compatibilism is untenable since conscious will seems to be an illusion. However, if we take Dennett's idea of "atmosphere of free will" and view conscious will as an important concept or "user illusion" which is one of the atmosphere's building blocks, we can see how a new compatibilism could be reached. Although from the point of view of scientific thinking conscious will seems illusory, inspired by Wilfrid Sellars's conception of manifest and scientific images we can start to understand free will as existing on its own conceptual level. The confusion stems from mixing the two frameworks.

Kompatibilismus Daniela Dennetta, založený na redefinování svobodné vůle skrze rozšíření konceptu „já“ o nevědomé procesy, se zdá být v nesouladu s jistými intuicemi. Jak upozorňuje Sam Harris, vyhýbá se té svobodné vůli, o kterou nám zřejmě intuitivně jde – vědomé svobodné vůli. Tento kompatibilismus je neudržitelný, protože vědomá vůle se zdá být iluzí. Když ale přijmeme Dennettovu myšlenku „atmosféry svobodné vůle“ a nahlédneme vědomou vůli jako důležitý koncept nebo „uživatelskou iluzi“, která je součástí této atmosféry, můžeme najít cestu k novému kompatibilismu. Ačkoliv se z vědeckého pohledu zdá být vědomá vůle iluzí, inspirování koncepcí zjevného a vědeckého obrazu Wilfrida Sellarse můžeme porozumět svobodné vůli jako existující na své vlastní konceptuální úrovni. Zmatení přichází právě s mícháním zmiňovaných dvou rámců.

Introduction

In this paper I would like to introduce and discuss two approaches towards the problem of free will: Daniel Dennett's compatibilism and Sam Harris's determinism. My aim is to both problematize and find inspiration in ideas of these authors in order to better understand the confusion which comes together with the ever controversial question about the possibility of compatibility of free will and strict physical laws and the role that conscious awareness plays

in it. The main issue will be to characterize and confront two levels of thinking about free will – the unreflective intuitive level and the scientific, more fine-grained level.

In the first section I will introduce Daniel Dennett's conception, in the second Sam Harris's criticism of Dennett and his own conclusions. In the third section I will try to point to the strengths and weaknesses of both approaches and use the strengths to reveal such approach toward the problem which would enable us to understand the confusion concerning free will and find a possible way to clarify it. This should possibly lead to a new formulation of compatibilism and an answer to the question how we can keep both intuitive and scientific understanding of free will while avoiding paradoxes and tensions.

Compatibilism?

One of the crucial aspects of Daniel Dennett's conception is his attempt to demonstrate the illusoriness of our unreflective common-sense concept of free will which is closely connected to the concept of self. In order to understand the mechanism of how our concept of free will arises, it is important to turn to the concept of causation. Dennett mentions experiments conducted by Daniel M. Wegner which demonstrate the propensity of subjects to "misattribute decisions to themselves that are in fact being made by somebody else."¹ Causation is very problematic issue in the history of philosophy. We keep realising that the way we view causes and effects is largely dependent on our natural tendency to be

"overeager to interpret, to "notice" things causing other things when, in fact, both "cause" and "effect" are effects of complex machinery that is hidden from us – backstage, in effect."²

It is all about *our* interpretation of what we observe. We are not getting the "real" causes served on a platter and according to Wegner, our experience of conscious will arises from the process which interprets the connections between our thoughts and actions, not from the connections themselves.³ When we feel that we do something consciously and voluntarily, it is only our interpretation of what is actually going on.

In fact, as Dennett points out, we only observe our decisions *arriving*; we never see the whole process of them *being made*. "We have to see how we are going to decide something, and when we do decide, our decision bubbles up to consciousness from we know not where."⁴ From this perspective it seems that we are strangely bereft of the real responsibility for our actions. We are only observers of the results coming from the impenetrable depths of our unconscious minds. This has a lot to do with "the idea of the self as a unitary and cohering point of view on the world", an illusion which arises when we are trying to come up with an

¹ Dennett (2003, p. 243).

² Ibidem, p. 244.

³ Ibidem.

⁴ Dennett (1984, p. 78).

interpretation of a certain action and when we try to answer the question “Did *I do that?*”⁵ This “illusion” is very useful since it helps us interpret in an effective way certain events we encounter. Under deeper observation, however, it doesn’t really make much sense.

By intuitively accepting the idea that our conscious self is really “the entity” which actually makes all the voluntary decisions, we implicitly accept some sort of “supernatural free will”. If we don’t assume that the decisions arrived to consciousness from intricate webs of unconscious processes, the only possible explanation is that they had to come from “nowhere”; they simply appear to conscious self somehow miraculously. Thus, conscious self is the only entity available to take the responsibility: “...we exploit the cognitive vacuum, the gaps in our self-knowledge, by filling it with a rather magical and mysterious entity, the unmoved mover, the active self.”⁶ Dennett explicitly claims that free will thus understood does not exist. His position is succinctly captured in the following two sentences:

“If you are one of those who think that free will is only *really* free will if it springs from an immaterial soul that hovers happily in your brain, shooting arrows of decision into your motor cortex, then, given what *you* mean by free will, my view is that there is no free will at all. If, on the other hand, you think free will might be morally important without being supernatural, then my view is that free will is indeed real, but just not quite what you probably thought it was.”⁷

He suggests that free will is compatible with unconscious processes doing most of the work. We are not only the conscious tip of an iceberg; we are much more than that, all the intricate unconscious processes included. As Dennett often emphasises, there is a danger in excluding too much from our concept of self:

“As I never tire of pointing out, all the work done by the imagined homunculus in Cartesian Theater has to be broken up and distributed in space *and time* in the brain. It is once again time to repeat my ironic motto: If you make yourself really small, you can externalize virtually everything.”⁸

In other words, according to Dennett, free will is real, but it needs to be redefined in the face of deeper insight into its underlying mechanisms, especially into the basis of the concept of self. On the one hand, he uncovers the illusoriness of the conscious self understood as the ultimate source of decisions and, on the other, he tries to save free will by widening the concept of self: we are not only the conscious observers but also the unconscious (and possibly deterministic) processes lying behind.

What is interesting, however, is that in Dennett’s positive account the unreflective concept of “supernatural” conscious free will doesn’t disappear altogether. This idea, as I see

⁵ Ibidem.

⁶ Ibidem, p. 79.

⁷ Dennett (2003, p. 223).

⁸ Ibidem, p. 238.

it, is very important for the constitution of what Dennett calls “the atmosphere of free will”. This is an important conceptual atmosphere which enables us to think about the world in a certain way: we ascribe to people intentions, plans, hopes, etc., and we can honour them or blame them – all this because we perceive them as agents possessing free will:

“The idea that we have free will is another background condition for our whole way of thinking about our lives. We count on it; we count on people “having free will” the same way we count on them falling when pushed off cliffs and needing food and water to live...”⁹

It seems to me that the kind of free will that we ascribe to ourselves and to others has to be, in fact, the “supernatural” conscious kind. We can be aware of its illusoriness, but we still employ it, intuitively and unreflectively, on the everyday practical basis. Dennett never explicitly states it like this but I think that no other concept of free will would really do the job. I will return to this problem in the last section.

The mentioned “atmosphere of free will” is something that had to come into being gradually. Dennett attempts to convey an exhaustive analysis of mechanisms which play crucial role in the process of constituting the atmosphere of free will. He starts by uncovering continual emergence of reasons, intentions and interests out of complex set of conditions whose elementary roots are bereft of such attributions.

“In the beginning, there were no reasons; there were only causes. Nothing had a purpose, nothing had so much as a function; there was no teleology in the world at all. The explanation for this is simple: there was nothing that had interests.”¹⁰

The first “interests” evolved much later as a result of gradual process of complexity accumulation. They were no full-fledged interests which are characteristic of human beings today, however. We only call them interests because we project concepts of our present perspective onto much simpler things which exhibit familiar patterns. In the spirit of this reflection we say that simple replicator’s interest is self-replication.¹¹ The interests became better defined with the development of the replicators’ abilities to “defend their own interests” and “preserve this and that (their varieties of *homeostasis*)”.¹² This means that certain concepts which we consider to be capturing some irreducible aspects of reality are emergent – they are our cognitive reactions to the results of the initial conditions gradually changing in virtue of being formed under the influence of simple, arbitrary principles, namely the evolutionary principle as described by Darwin (his “strange inversion of reasoning”).

As I understand it, the fact that we tend to think in terms of reasons, intentions, etc., is due to *our* inability to view the whole complicated interplay of various factors. When we say

⁹ Ibidem, p. 10.

¹⁰ Dennett (1984, p. 21).

¹¹ Ibidem.

¹² Ibidem, p. 22.

that they “came to be”, we mean that *we* started to see the world this way. This could be nicely illustrated also by Dennett’s example describing two chess programs playing a match. Provided that the programs are complex enough, we view the combat as very suspenseful – from what we are actually able to consciously observe we can never predict which computer will win. However, the programs are in fact determined:

“What from one vantage point appear to us to be two chess programs in suspenseful combat can be seen through the “microscope” (as we watch instructions and data streaming through the computer’s CPU) to be a single deterministic automaton unfolding in the only way it can, its jumps already predictable by examining the precise state of pseudo-random number generator. There are no real “forks” or branches in its future; all the “choices” made by A and B are already determined.”¹³

Unable to “see” the actual determinism of the situation, we are led to think of the computers as making choices and having various possibilities. Even though they are much simpler than human beings, we can still find it useful to treat them as agents and ascribe intentions to them.¹⁴

According to Dennett, free will is not to be understood as some pre-existing feature of our existence; in fact, it evolves: “It is an evolved creation of human activity and beliefs, and it is just as real as such other human creations as music and money. And even more valuable.”¹⁵ Concept of free will emerges as a result of our reaction to certain complex conditions. As we could see, intentions were not there at the very beginning. They came with complexity and so did the idea of responsible agents.¹⁶ The idea of conscious free will is an emergent concept: even though it seems to capture something which escapes physical laws, it arose from completely natural conditions. It enables us to see the world in a specific way and helps us to act more effectively. Even though our perception of ourselves and other people is illusory on closer inspection, it is useful. This thought comes forward for example when

¹³ Dennett (2003, p. 80).

¹⁴ Ibidem, p. 81.

¹⁵ Ibidem, p. 13.

¹⁶ According to Dennett, the atmosphere of free will consists of concepts such as “*intentional action, planning and hoping and promising – and blaming, resenting, punishing and honoring*”. Ibidem, p. 10. It seems that this conceptual complex is a basis for our understanding of us and other human beings as moral and responsible persons. Thus, as I understand Dennett, free will, personhood and moral responsibility go hand in hand, being connected through intentions, planning of acts and subsequent praise or blame for these acts: as soon as beings evolve their own intentions and ability to plan acts following these intentions, they become responsible for their conduct because interaction between such beings naturally implies concurrent evolution of the atmosphere of free will and belief in free will (a kind of “bootstrapping”). We like or dislike certain actions and it is only logical to trace them back to those “bundles of intentions” that are their source. Ascribing responsibility is a tool which helps us influence behaviour of others as well as our own – praise encourages, blame discourages and the “air” of free will filled with responsibility makes us certainly think twice about our actions since they build our “moral image”.

Dennett discusses the shift from *design stance* to *intentional stance*. Imagine a designer who creates simple entities operating according to some simple principles. As he manages to develop more complex systems, there will come a point when he can start thinking of them as rational agents with intentions, beliefs, etc. This helps the designer to think of them on a higher level without being flooded by unnecessary details of the complicated mechanisms underlying the observable behaviour of the entities in question:

“It makes life blessedly easier for the high-level designer, just the way it makes life easier for us all to conceptualize our friends and neighbours (and enemies) as intentional systems.”¹⁷

The simplification brought about by the above described conceptualisation is similar to a familiar case of computer users. Dennett mentions the way software designers simplify and even distort the truth about the real workings of the computer so that it can be manipulated by the users intuitively. One can click and drag, hear various sound effects and orientate according to icons on the desktop – all this draws on usual and natural ways we perceive the world around us and act in it. Similarly, communication between people with a “self” provides access to such features of agents which are much easier to grasp and operate with than otherwise very intricate nets of brain processes. The concept of self certainly makes it easier to communicate what is going on in our brains and to influence other agents.¹⁸ It is all about making things more effective on a higher level of complexity. Reality may get distorted but the most important thing is that the “trick” works in the end.

What we encounter here is the problem of confrontation of two different levels of our view of the world. As Dennett points out (referring to Sellars), on the one hand, there is an unreflective every-day view – all the phenomena we see with naked eye, middle-sized objects and rates of change, etc. (“manifest image”). On the other hand, we also have “scientific image”. For example,

“we understand that while “water” is a mass noun for us, water is also a swarm of countable molecules, whose trajectories are trackable in principle, and sometimes even in practice (with the aid of prosthetic extensions of our senses).”¹⁹

We are capable of abandoning our manifest image and start looking at the world differently, using “more fine-grained level of description”. While adopting this new outlook we start realising how problematic our everyday commonsense thinking really is. However, the truth is that we need our intuitive understanding and it is the only way to think on a day-to-day basis. We can clearly see it in the case of free will problem: we are deliberators and if we want to deliberate effectively, we must be faithful to certain unreflective concepts (e.g. “open future”, even if determinism is true).²⁰ I believe that effectiveness is really the key issue here:

¹⁷ Ibidem, p. 45.

¹⁸ Ibidem, p. 248f.

¹⁹ Dennett (1984, p. 114).

²⁰ Ibidem, p. 114f.

we have to work with what we have and we have finite and bounded epistemic equipment. If we were able to process all the information concerning the intricate causal interactions pertaining to micro-level in short time, we wouldn't need "user illusions". But we do need the concept of self precisely because it provides us with the ability to predict behaviour and orientate in the world of complex beings while we only have to process relatively small amount of information. We project "intentions" and "beliefs" into others and ourselves because it works in the end – we are capable of accurate predictions and effective interactions. Concepts concerning personhood serve us well and thus influence dramatically the way we intuitively view reality – even though in this case we are not "knowers", as we tend to think, but "constructors" of useful conceptual tools. As by using computer, we don't learn about the mechanisms which make all the applications possible; we simply use them and learn on this user-level.

In my opinion, the main advantage of Dennett's conception of free will is that it brings forward a certain idea of "concept-emergence". A level of reality which seems utterly unique and irreducible (as the above discussed level of the atmosphere of free will) can arise from much simpler elements, provided that these elements are arranged into sufficiently complex structures (beings like us) entering mutual interaction. The important thing to realise is that the feeling of novelty and irreducibility results from the way the complex situation in question is perceived by us. Various levels of thinking about the world are to be considered. On practical or day-to-day level it would be highly ineffective to try to keep track of all the details and that's why appropriate simplification or even distortion comes in handy. In other words, intricate complexity yields novel properties *for us*, because *we*, finite and cognitively bounded beings, need to effectively handle this complexity.

From the everyday perspective, the scientific image is unnatural. Our basic understanding of the world is based on various intuitions and unreflective concepts which work like useful shortcuts or user illusions. Once we look deeper into the mechanisms that ground the familiar phenomena we suddenly encounter completely different world which so often contradicts our everyday perception. Thus, in order to avoid confusion, it is important to realise that the everyday view and the scientific view have to be carefully distinguished. Dennett goes in this direction but I don't think that he fully articulates the role of the intuitive level. He redefines free will in terms of the scientific view and by this introduces certain confusion into the problem – the "redefined free will" belongs to the scientific framework, but the original "intuitive free will" is something quite different. Dennett confuses the two possible meanings and bases his compatibilism on this confusion. I will discuss this further in the following sections.

Changing the subject

Sam Harris attacks Dennett's conception as problematic in his own account of free will whose style is perhaps more popular than philosophically deep but it rightly points to some important issues. This has to do with the role that our intuition plays in the problem. Harris emphasizes

the subjective strength of our everyday common-sense concept of free will. He directly opposes Dennett's compatibilism, especially his claim that we are not only the conscious tips of an iceberg, but also the intricate unconscious neural processes. Harris stresses again and again that the free will problem is based on psychological fact – the feeling most people have about their free will:

“Compatibilists generally claim that a person is free as long as he is free from any outer or inner compulsions that would prevent him from acting on his actual desires and intentions. (...) The truth, however, is that people claim greater autonomy than this. Our moral intuitions and sense of personal agency are anchored to a felt sense that we are the conscious source of our thoughts and actions.”²¹

According to Harris, compatibilists simply “change the subject”: they ignore the subjective feeling people have about their status as conscious agents and serve us with a specific concept of person instead. What makes the problem of free will so acute is the *feeling of agency and moral responsibility*, and to ignore this is to miss the whole point.²² According to Harris, there is no place for compatibilism because it ignores the only kind of free will worth talking about – and this kind of free will simply doesn't exist.

To support his claim that we simply don't have free will, Harris too turns to the problem of the attribution of agency and mentions the scientifically described cases showing the unreliability of our interpretation skills. According to him, our interpretation of the cause-effect relation between our thoughts and actions is even more erroneous than we tend to think: “There is no question that our attribution of agency can be gravely in error. I am arguing that it always is.”²³ It is not only actions but also intentions whose origins we interpret incorrectly. The problem is that we are consciously aware of our intentions and we intuitively believe these intentions to originate from our conscious selves. Their true source is, however, hidden from us. It belongs to the realm of unconscious brain events that we don't intend.²⁴

So for Harris it is all about the subjective feeling of agency and our intuitions. The concept of free will he is interested in is the unreflective concept which, according to him, most people share. This is what Dennett would call “supernatural” free will, since under closer observation it doesn't make sense. It implies decisions coming from nowhere and suddenly appearing in our consciousness. If the conscious self was their ultimate author, we would stand face to face with a strange idea that we create ourselves *ex nihilo* – that our decisions-building blocks of our moral character, pop up suddenly without any prior warning thanks to the god-like power of our conscious self and *nothing else*. However, when we give this problem a deeper thought it seems only natural that every decision has to be based on

²¹ Harris (2012, p. 27).

²² Ibidem, p. 31.

²³ Ibidem, p. 31.

²⁴ Ibidem, p. 32.

something. We need some prior background of knowledge and experience in order to deliberate and it should not be very surprising that *unconscious* brain processes play the main role. When we really think about it we realise that there is no reason to presuppose an agent independent of all the possible influences. On the contrary, both outer influences and those coming from our own brain are in fact necessary for a process of deliberation and the resulting decision to take place. Harris agrees with all this, but for him it doesn't imply that we are more than the conscious self. Rather, it implies that we, conscious selves, live in an illusion.

Harris, unlike Dennett, doesn't try to "save" free will; he simply states that we don't have it. He "analyses free will away" and doesn't seem to find any particularly positive role for the unreflective concept in our everyday lives. He acknowledges that thinking about free will in terms of its illusoriness might have some bad impact on certain moral tendencies (he mentions example of students who cheated more after being confronted with an argument against the existence of free will). On the other hand, he claims that "the truth" could possibly increase one's "feelings of compassion and forgiveness".²⁵ Anyway, his ultimate claim is that "The illusion of free will is itself an illusion." We might feel that we are the conscious authors of our decisions but as soon as we try harder and explore our experience more thoroughly, we realise that we don't even feel supernaturally free anymore:

"It is not that free will is simply an illusion – our experience is not merely delivering a distorted view of reality. Rather, we are mistaken about our experience. (...) Our sense of our own freedom results from our not paying attention to what it is like to be us. The moment we pay attention, it is possible to see that free will is nowhere to be found, and our experience is perfectly compatible with this truth. Thoughts and intentions simply arise in the mind. What else could they do?"²⁶

There simply seems to be no place left for conscious free will.

Two conceptual worlds

In my opinion, both Dennett's and Harris's conceptions share a common problem: neither Dennett nor Harris distinguishes properly between the two levels of conceptualization of free will.

Dennett refuses the unreflective concept (the one emphasized by Harris) as "supernatural" and illusory and redefines free will in terms of broader understanding of its mechanisms. On the other hand, he seems to acknowledge the importance of our thinking

²⁵ Ibidem, p. 44.

²⁶ Ibidem, p. 56. This might seem paradoxical and lead us to think that Harris contradicts himself. I believe, however, that he simply tries to show precisely that the paradox is a crucial feature of our understanding of free will. Conscious free will is an unreflective intuition - it is like a fuzzy picture: it makes perfect sense until we look at it too closely. It is simply "not meant" to be looked at too closely.

about ourselves and others as possessing free will. However, my problem with his conception is that when he talks about the atmosphere of free will he seems to be talking about the unreflective view of free will (the conscious free will), but he doesn't explicitly acknowledge its positive role. The question is whether the atmosphere of free will would be preserved provided that we forgot about the conscious "homunculus agent" altogether and thought about our own free will scientifically. I just don't think so. If we didn't use the "shortcut" or "user illusion", we would not be able to act effectively, "under the idea of freedom". It is quite possible that we have to view ourselves and others as conscious agents in order to preserve the effective functioning of our moral interactions.²⁷ This does not mean that we cannot also be aware of the true mechanisms, but we simply don't embrace this scientific view when we are acting in our social world on everyday basis. I appreciate Dennett's attempt to redefine free will so that we incorporate also unconscious processes into our concept of self but I agree with Harris that this is simply changing the subject. Dennett claims that we have free will, but *this* free will is not *the* free will whose concept actually enables us to enter the "atmosphere of free will".

I believe that also Harris misses an important point by not paying enough attention to the role of our intuitive concept of free will: it really seems that people normally act "under the idea of freedom" which means that they think of themselves and of others as free conscious agents. The fact that this idea doesn't make sense under closer inspection has nothing to do with the way we function in the world on a day-to-day basis. There is a level of conceptualisation which plays crucial role for us as complex beings interacting with other complex beings. We need shortcuts and "user illusions" to get by in this environment. We can also realise, by using microscope, that water isn't really what we normally see it to be. But this doesn't mean that we *should* deny its liquidity and transparency without further qualifications.

I have to agree with Harris's criticism of Dennett's "changing the subject" - it seems to me that he confuses the issue by redefining free will and putting the unreflective concept of conscious free will aside. However, Dennett does a good job by introducing different possible ways of understanding complex phenomena (e.g. our switching to "intentional stance"). If we follow Harris, it may lead us on our way to acknowledge the subjective strength of the unreflective concept. Following Dennett, in turn, can enable us to find a proper place for *this* free will. Combining the two, we do justice to our common-sense concept and at the same time we save its validity by looking at free will while distinguishing different possible stances.

In order to illustrate the point I would like to return to the parallel between our concept of free will and that of water. Once again, unreflective understanding of free will implies the idea that there is some ultimate unit – the conscious self who is the sole author of all the free decisions. This is the homunculus, the explanatory entity of the first order – something which

²⁷ See e.g. Baumeister, Masicampo, DeWall (2009). The authors suggest that their experiments point to a possibility that "disbelief in free will increases aggression and reduces helpfulness".

makes the mechanism of free will seem “supernatural” because we simply have to take it for granted and cannot ask about further, more fine-grained explanations which would be more akin to scientific inquiry. When we analyze the phenomenon closer, however, we discover immense net of unconscious processes and failures in our ability to reveal the actual causal relationship between our thoughts and actions. From scientific perspective there is no conscious self which would be the ultimate source of free decisions. In my opinion, something similar happens when we think about water. On a day-to-day basis water is something liquid, transparent, something we can drink and something we can drown in, etc. From scientific perspective water is neither liquid nor transparent – it is a collection of molecules of H₂O. This is supposed to be the true nature of water, viewed as a result of more fine-grained analysis. On this level of understanding it doesn’t make sense anymore to talk about manifest properties of water as they appear to our senses since this would be just a coarse approximation or even distortion.

We can, however, use the scientific view to explain the effects which appear in our unreflective and “crude” level of reality. We explain liquidity by referring to behaviour of the H₂O molecules, for example. In the case of free will it is not so much different. We discover the mechanisms which elicit in us the specific subjective feelings of conscious agency. However, this doesn’t mean that we should deny the validity of the unreflective intuitive thinking. Both the concept of liquidity and the concept of conscious free will are emergent: they represent the phenomena in question in such a way that many of the “scientific details” can be abandoned so that we can orientate more effectively in our world. They are clever shortcuts which may distort the actual mechanisms standing behind the phenomena but which are very useful in day-to-day practice nonetheless.²⁸

The point I am trying to reach is well illustrated by what Sellars says about the conflict between the everyday and the scientific framework:

“...the claim that physical objects do not really have perceptible qualities is not analogous to the claim that something generally believed to be true about a certain kind of thing is actually false. It is not denial *of* a belief *within a framework*, but a challenge to the framework. It is a claim that although the framework of

²⁸ There is, of course, a danger of taking the analogy too far. My thanks to anonymous reviewer who pointed out the disanalogy. As I understand the problem, liquidity shouldn’t be considered to be an illusion in the same way as free will. Science doesn’t deny liquidity in the same way it seems to deny free will – we can see how liquidity arises on a basis of molecular structure, etc. We can also come to see how our intricate brain processes give rise to the illusion of conscious free will. However, the difference is that we don’t feel it necessary to deny our subjective perception of how water feels to us - liquidity, transparency, etc. even though there is a certain distortion: we don’t see the molecules moving around (the same would go for perceived solidity of things around us despite the fact that they mostly consist of empty space between particular atoms). In the case of free will we don’t see the whole causal chain of brain processes leading to our decisions. Our finite epistemic capacities leave us “blind” to another possible, differently parcelled and more fine-grained view of the workings of the world. The point is that in the case of free will the “blind spot” seems so much more prominent because it has an immense impact on how we understand ourselves as moral beings. Whether we see molecules with the naked eye doesn’t really matter so much and we feel free to let it pass.

perceptible objects, the manifest framework of everyday life, is adequate for the everyday purposes of life, it is ultimately inadequate and should not be accepted as an account of what there is *all things considered*.”²⁹

Only when we approach free will problem from this perspective can we come to understand why “the truth” seems so surprising, unintuitive and controversial.

The core of the problem seems to be this: confusion happens when we mix the frameworks. It appears to me that it is simply impossible to build a conceptual bridge which would smoothly connect scientific and intuitive account of free will. We encounter an unavoidable abyss here. Scientific thinking is based on carefully articulated explanations and fine-grained analyses. The common-sense concept of free will ends up being revealed as unsatisfactory because of the weakness of its explanatory power. Wegner points out that the unreflective idea of free will is “homunculus-based”. We postulate a homunculus who decides things without any prior causes which would have some impact on the decisions. We come to

“an explanatory entity of the first order. Such an explanatory entity may explain lots of things, but nothing explains it. (...) A first-order explanation is a stopper that trumps any other explanation, but that still may not explain anything in a predictive sense. (...) There cannot be a science of this.”³⁰

We may intuitively feel that there is such a homunculus but as soon as we assume scientific approach and analyze the deliberation process more precisely, we uncover the pitfalls. This is the moment when we, together with Harris, might want to say that there simply is no such thing as free will.

Harris makes the above mentioned unwarranted step: he mixes the two frameworks. He adopts scientific framework and tries to implant the fine-grained concept of free will to the framework of our common-sense intuitions. This results in confusion and tension because, scientifically speaking, conscious free will is a nonsense, and, “naturally” or “intuitively” speaking, we don’t like the idea of determinism applied to ourselves as beings living in the world of moral interactions. Dennett does distinguish between the frameworks – he is aware of the fact that on a certain level of complexity it is useful to conceptualise things differently (a very good example is his “intentional stance”). But he doesn’t state it clearly that it is the “supernatural” conscious free will which serves us so effectively on the level of interactions between moral agents. He changes the concept of the self to save free will and by this he smuggles in the scientific framework. His “atmosphere of free will”, I believe, has to belong to the manifest, common-sense framework. This means that he shouldn’t say that supernatural free will doesn’t exist. He should say rather that it exists within its own manifest framework and is illusory within the other – the scientific framework.

Yet another thing which requires clarifying is the comparative adequacy of the frameworks. Even though we tend to perceive the scientific framework as the “true” one, or

²⁹ Sellars (1963, p. 27).

³⁰ Wegner (2005, p. 20).

the one we should prefer (Sellars's "scientia mensura"), this could be actually misleading. Science can indeed provide us with fine-grained detailed explanations by uncovering the intricate mechanisms behind things, but there is a functioning way of understanding the world which *ignores* these details. And it is this very fact of simplification which gives rise to new and wonderful level of existence – a world of incredibly complex and yet epistemologically bounded creatures. This paradoxical combination of complexity and epistemic boundedness is the true soil where free will can flourish. To handle their own complexity the creatures need to learn to understand³¹ themselves in a specific way: being finite and unable to process all the possible information available, they have to think in shortcuts and clever simplifications. Their world is a world where the intuitive unreflective concepts are its real building blocks. The world of conscious agents is real on its own level and cannot be torn down by scientific analysis. Science can only describe how various building blocks came into being but the bricks and pillars of manifest image still work together to give rise to a coherent building.

I believe that it is also very important to stress how the two frameworks differ in the way we employ them. Scientific stance can be adopted only temporarily and under specific conditions: it can take us considerable amount of time to analyze a certain phenomenon and this analysis takes form of such complicated information that we have to exert considerable amount of effort to grasp it. This all, of course, is in contrast with prompt interpretation of our environment facilitated by our intuitions. Scientific thinking can thus in many cases³² never become our day-to-day mode of orientation in the world – it is unnatural and ineffective. Manifest framework, by having its own building blocks (unreflective concepts which work as useful simplifications) and its own rules, has its own indisputable validity. It is a world on its own.

By keeping the two frameworks apart we can better understand the specific status of conscious free will. We can see its illusoriness from one perspective and acknowledge its validity from the other. We cannot escape the fact that the manifest image is natural for us. When we try to smuggle in the scientific concepts, our intuitions get a hard time. If we adopt the scientific perspective and try to suppress our intuitions for a moment, we can gain an understanding of how our worldview works on various levels. We can get insight into the process of gradual coming to be of intentions, purposes and even free will. We must remember, however, that we are still those same finite and epistemologically bounded creatures who need to function effectively in an immensely intricate reality. We simply cannot get by without useful simplifying concepts, because they make us understand the

³¹ This kind of understanding is not really meant to simply help us explain behaviour of complex beings; rather, it is meant to provide us with a tool for *interacting* with them.

³² I would like to thank anonymous reviewer for pointing out that there are, of course, cases, in which the scientific view helps us to overcome certain harmful intuitions and unreflective tendencies (e.g. racism) but here we are concerned with something much more basic – at least I believe that the notion of free will, however vague, is crucial for our understanding of human as not only an animal, but primarily a responsible person. Every other moral question builds on this picture of a man as responsible agent and is ultimately dependent on the "atmosphere of free will".

world on one, very important, level – a level we cannot escape because it is always with us, wired in our brains. The only way to avoid confusion is to fully realise that we are capable of adopting two different perspectives and two modes of understanding reality whose particular roles have to be carefully distinguished.

Conclusion

I have tried to show that both Dennett and Harris approach the problem of free will in a way which is in conflict with our intuitive understanding of us and other people. They simply state that conscious self is not the ultimate author of our decisions. Harris sticks to determinism while Dennett tries to save free will and comes up with a version of compatibilism. However, I believe that this is not the real compatibilism. Harris correctly criticises Dennett for his “trick” – redefining free will so that it fits scientific framework and then claiming that *this* “Dennettian” free will exists and thus compatibilism is true. The real compatibilism would have to acknowledge, just as Harris points out, our subjective feeling of conscious agency as the free will we in fact care about.

I proposed a way how to get to such compatibilism. It is inspired by Dennett’s theory of different possible “stances” and Sellars’s theory of different frameworks or images. Certain way of understanding the world is natural for us – we are wired this way and this fact enables us to effectively handle environment where complexity reaches new levels. On the other hand, we are also capable of more fine-grained approach – we became sophisticated enough to perform deeper analyses of the phenomena we encounter in our everyday lives. This led us to many unintuitive conclusions. The problem is that we tend to mix the frameworks. The only way we can clarify the situation is to acknowledge the abyss between the two frameworks. To say that free will doesn’t exist is like to say that physical objects don’t have perceptible qualities. They indeed do have perceptible qualities – for *us* on the everyday level. And we indeed do have free will – it is real *subjectively*, for *us* as finite beings who try to make their way through the world which is too complicated to be perceived in detail. Manifest image has its own validity; it is a world on its own which works in its specific way. The fact that there are complicated mechanisms which give rise to this world cannot change anything about it.

Bibliography

- Baumeister, R. F., Masicampo, E. J. & DeWall, N. C. (2009): „Prosocial Benefits of Feeling Free: Disbelief in Free Will Increases Aggression and Reduces Helpfulness.“ In *Personality and Social Psychology Bulletin* 35 (2): 260–268. Available from: <<http://psp.sagepub.com/content/35/2/260.full.pdf+html>>.
- Dennett, D. C. (1984): *Elbow Room: The Varieties of Free Will Worth Wanting*. Clarendon Press, Oxford.
- Dennett, D. C. (2003): *Freedom Evolves*. Viking, New York.

Harris, S. (2012): *Free Will*. Free Press, New York.

Sellars, W. (1963): „Philosophy and the Scientific Image of Man.“ In *Empiricism and the Philosophy of Mind*, W. Sellars, Routledge & Kegan Paul Ltd, London, 1963, pp. 1–40. Available from: <<http://selfpace.uconn.edu/class/percep/SellarsPhilSciImage.pdf>>.

Wegner, D. M. (2005): „Who Is the Controller of Controlled Processes?“ In *The New Unconscious: Social Cognition and Social Neuroscience*, ed. R. R. Hassin, J. S. Uleman & J. A. Bargh, Oxford University Press, New York, 2005, pp. 19–36. Available from: <http://scholar.harvard.edu/files/dwegner/files/who_is_the_controller_of_controlled_processes.pdf?m=1360040444>.