Why phenomenal zombies are conceivable whereas anti-zombies are not

Dmytro Sepetyi

Abstract. In this article, I discuss Keith Frankish's attempt to neutralize the zombie argument against materialism with a closely parallel argument for physicalism, the anti-zombie argument, and develop David Chalmers' reply to this species of arguments. I support Chalmers' claim that the conceivability of situations like the existence of an anti-zombie is problematic with an analysis that makes it plausible that the idea of an anti-zombie is incoherent, and argue that to counter this, a materialist should deny the absence of a priori entailment from the physical to the phenomenal; however, this would involve the denial of the conceivability of zombies and so make the anti-zombie argument superfluous.

Keywords: anti-zombie, coherence, conceivability, phenomenal zombie, physicalism.

Чому феноменальні зомбі можливі, а антизомбі — ні

Анотація. У цій статті обговорюється спроба Кейта Френкіша нейтралізувати аргумент зомбі проти матеріалізму, протиставивши йому аналогічний аргумент на користь фізикалізму – аргумент антизомбі; розвивається відповідь, яку дав аргументам цього роду Девід Чалмерс. На підтримку думки Чалмерса, що збагненність ситуацій, подібних до існування антизомбі, є проблематичною, запропоновано аналіз, який показує, чому ідея антизомбі є, ймовірно, некогерентною. Доведено, що для відхилення цього аналізу, матеріаліст мусить заперечувати відсутність апріорного випливання феноменального з фізичного; проте це означає заперечення збагненності зомбі, що робить аргумент антизомбі непотрібним.

Ключові слова: антизомбі, когерентність, збагненність, феноменальний зомбі, фізикалізм.

Актуальні проблеми духовності: зб. наук. праць / Ред.: Я.В. Шрамко. — Кривий Ріг, 2019. — Вип. 20. — С. 18–28.

Keith Frankish's anti-zombie argument is one of the most interesting attempts at a rebuttal of the zombie argument, which «in recent years ... has come to occupy a central role in the case against physicalist views of consciousness, in large part because of the powerful advocacy it has received from David Chalmers»[1]. It is representative of the series of similar arguments advanced by Marton[2], Sturgeon[3], Brown[4], and others, so that the brunt of the following discussion bears on these arguments too.

The zombie argument, as construed by Chalmers, proceeds from the conceivability of exact physical duplicates of human beings without phenomenal mental states (phenomenal zombies) to their metaphysical possibility, and from this to the falsity of physicalism. Frankish counters this with an apparently closely parallel argument to the opposite, and argues that this efficiently neutralizes the zombie argument, so that the result is «a stand-off». Chalmers replied that this is not the case, because the decisive premise of the zombie argument is obviously true, whereas the «parallel» premise of the anti-zombie argument is not[5]. It is obvious that phenomenal zombies are conceivable; it is not obvious that anti-zombies are. However, this reply is likely to need more defense than Chalmers provides. Anti-zombies are likely to *seem conceivable* for materialists (I suppose that Frankish, when advancing the argument, did really believe that anti-zombies are conceivable). If so, something more is required on the side of defenders of the zombie argument. In this article, I propose the explanation and defense of Chalmers' point that serves this purpose.

To begin with, let us rehash and compare the two opposing arguments.

The zombie argument is an argument about imaginary creatures called «phenomenal zombies», which are arguably metaphysically possible, and their metaphysical possibility means that physicalism is false. Phenomenal zombies are exact — atom-to-atom (or quark-to-quark, if you like), with all the structure and dynamics — physical copies of human beings but with no mental (phenomenal, subjective) states. There are all the same physical interactions and movements of cells, molecules, atoms, etc. inside

[1] *Frankish K.* The anti-zombie argument // Philosophical Quarterly. — 2007. — Vol. 57. — P. 651.

[2] *Marton P.* Zombies versus materialists: The battle for conceivability // Southwest Philosophy Review. — 1998. — Vol. 14. — P. 131-138.

[3] *Sturgeon S.* Matters of Mind: Conscious, Reason, and Nature. — London : Routledge, 2000.

[4] *Brown R.* Deprioritizing the A Priori Arguments Against Physicalism // Journal of Consciousness Studies. — 2010. — Vol. 17. — P. 47-69.

[5] *Chalmers D.* The Two-Dimensional Argument Against Materialism // Chalmers D. The Character of Consciousness. — New York : Oxford University Press, 2010. — P. 180.

and all the same movements of legs, arms, fingers, lips, tongues, eyelids, etc. outside; however, there are no subjective experiences and no mental subjects-experiencers; there is nothing it is like to be a zombie. The zombie argument, as construed by Chalmers[6], proceeds from the thesis that zombies are conceivable to the thesis that they are metaphysically possible and from this to the conclusion that physicalism is false.

Frankish's anti-zombie argument is an argument about imaginary creatures called «anti-zombies», such that their metaphysical possibility means that physicalism is true. Anti-zombies are exact — atom-to-atom (or quark-to-quark, if you like), with all the structure and dynamics — physical copies of human beings, with all human experiences but without any nonphysical states. There are all the same physical interactions and movements and all the same subjective experiences, and there is something it is like to be an anti-zombie; however, there is nothing nonphysical. The anti-zombie argument proceeds from the thesis that anti-zombies are conceivable to the thesis that they are metaphysically possible and from this to the conclusion that physicalism is true.

It seems that Chalmers' zombie argument and Frankish's anti-zombie argument are exactly parallel. Now, if the zombie argument, taken separately, has a considerable force, then the anti-zombie argument should have the same force in the opposite direction. Thus, the resultant force of the two arguments taken together is zero; the zombie argument is efficiently neutralized by the anti-zombie argument.

To understand the situation better, it is helpful to consider the common logical structure of the arguments. The structure involves two subarguments — let us call them «modal» (conceivability-to-possibility) and «substantial».

The modal (conceivability-to-possibility) subargument has a form:

(**M1**) X is conceivable.

(**M2**) Conceivability entails metaphysical possibility.

Therefore,

(**S1**) X is metaphysically possible.

[6]See: *Chalmers D.* Does Conceivability Entail Possibility? // Conceivability and Possibility / ed. T. Gendler and J. Hawthorne. — New York : Oxford University Press, 2002. — P. 145-200; *Chalmers D.* The Two-Dimensional Argument Against Materialism // Chalmers D. The Character of Consciousness. — New York : Oxford University Press, 2010. — P. 141-205.

(In the case of the zombie argument, X is a phenomenal zombie; in the case of the anti-zombie argument, X is an anti-zombie.)

The substantial subargument has a form:

(**S2**) If X is metaphysically possible, then physicalism is false/true.

(**S1**) X is metaphysically possible.

Therefore,

(**C**) Physicalism is false/true.

(In the case of the zombie argument, (S2) and (C) attribute to physicalism the value «false»; in the case of the anti-zombie argument, «true».)

The two subarguments constitute one argument, because the conclusion of the modal subargument (S1) is used as a premise in the substantial subargument.

Now, the initial premises of the zombie and anti-zombie arguments are (M1), (M2), and (S2). Of these, (M2) is shared and thus uncontroversial. (S2) is uncontroversial, too: both sides of the discussion agree that the metaphysical possibility of a phenomenal zombie entails the falsity of physicalism, whereas the metaphysical possibility of an anti-zombie entails its truth. The sides disagree only about which one — a zombie or an anti-zombie — is metaphysically possible. And this, as far as Chalmers' and Frankish's arguments go, is taken to be derived from the conceivability premise (M1). Thus, only (M1) is a controversial point of difference between the two arguments. We have two its opposing versions:

(**M1**$_z$) A phenomenal zombie is conceivable.

(**M1**$_{az}$) An anti-zombie is conceivable.

For the anti-zombie argument to neutralize the zombie argument, these two claims should be on equal footing? Are they? Chalmers suggests that they are not, because *it is not obvious that the truth of materialism is conceivable*, in the sense stronger than *prima facie* conceivability (= it is not obvious that anti-zombies are conceivable): many people find it inconceivable that consciousness is a physical process; although this inconceivability is not «so obvious that it should be used as a premise in an argument against materialism», the conceivability is not either[7].

[7] *Chalmers D.* The Two-Dimensional Argument Against Materialism. — P. 180.

Of course, this saves the zombie argument only if zombies are better off, that is, only if *it is obvious that zombies are conceivable*. So, Chalmers' response in a nutshell is that it is obvious that a zombie is conceivable, while it is not obvious that consciousness being a physical process (and so, an anti-zombie) is conceivable.

To estimate this response, we need to elucidate what is involved in the claim that something is conceivable, in the sense relevant to the zombie argument. There are some difficulties in this, because Chalmers' treatment of conceivability and possibility in the zombie argument had undergone a revision in the period between 1996 and 2002, without (as far as I know) his making an explicit statement of this revision. Let us pay attention to this revision-involving development.

First, consider the initial Chalmers' exposition of the zombie argument, in *The Conscious Mind*. Four important points should be noted about that exposition.

(**1**) Chalmers argued that the kind of possibility relevant for his argument is *logical possibility* — there being no incoherence in the description of zombies, as distinct from natural (nomological, causal) possibility, what is possible given the laws of nature of the actual world. Natural possibility is irrelevant because it may be that zombies are naturally impossible and materialism is false (if some laws of nature make it the case that some physical events in the brain cause, or produce, non-physical mental states). However, if the notion of zombies were incoherent (a world of zombies is logically impossible, whatever the laws of nature), this would mean that materialism is true; and if it is coherent, it seems that materialism should be false. Accordingly, Chalmers talks of «logical supervenience», which is a matter of logical necessity[8].

(**2**) Chalmers also discussed the proposition that the relevant kind of possibility can be what philosophers usually call «metaphysical possibility», and argued for *modal monism* — the claim that metaphysical possibility should be identified with logical possibility (perhaps, described in terms that accommodate Kripkean «*a posteriori* semantic twist»[9] by referring according to their *secondary intensions*).

(**3**) In the section «The possibility of zombies» Chalmers defended this possibility by straightforward appeal to obviousness rather than by

[8] *Chalmers D.* The Conscious Mind. — New York : Oxford University Press, 1996. — P. 34-38.

[9] *Ibid.* — P. 38.

inference from conceivability. First, he suggested that it is obvious that zombies are logically possible, that is, that there is no contradiction in the description of a zombie. Second, he proposed some «indirect arguments ... in order to establish that the obvious logical possibility really is a logical possibility, and really is obvious»[10].

(**4**) However, in an earlier fragment, Chalmers made an attempt to «spell out the relationships among logical necessity, conceptual truth, and conceivability»[11]. In particular, he suggested that «there seems to be no reason to deny that conceivability of a world implies possibility»[12], and supported this suggestion by considering traditional (Kripkean) purported counter-examples and showing that they can be declined by means of the two-dimensional analysis[13]. On this ground, Chalmers «henceforth take this for granted as a claim about logical possibility; any variety of possibility for which conceivability does not imply possibility will then be a narrower class»[14]. A few pages later, he proposed «using conceivability as a test for logical possibility»[15].

(4) can be seen as a precursor of his later construal that explicitly makes the conceivability-to-possibility entailment part of the foundation of the zombie argument. However, it is important that on this later construal, the kind of possibility that is (supposedly) entailed by conceivability is *not logical* but *metaphysical* possibility.

One can wonder if there is a difference between the claim that conceivability entails logical possibility and the claim that conceivability entails metaphysical possibility, given modal monism. However, there is a difference, because in Chalmers' later expositions of the zombie argument, modal monism is not «given» or argued for *before* the argument for the conceivability-to-possibility entailment but rather *results from* this argument.

In between 1996 and 2002, Chalmers explained that in his usage, «conceivability» and «logical possibility» are «*tied by stipulation*»; so that on this stipulation, «any variety of possibility for which conceivability does not imply possibility will then be a narrower class»[16]. Such an explanation

[10] *Ibid.* — P. 96-97.
[11] *Ibid.* — P. 65.
[12] *Ibid.* — P. 68.
[13] *Ibid.* — P. 67-68.
[14] *Ibid.* — P. 68.
[15] *Ibid.* — P. 73.
[16] *Chalmers D.* Materialism and the Metaphysics of Modality // Philosophy and Phenomenological Research. — 1999. — Vol. 59. — P. 477-8.

should be puzzling, because it makes nonsense of the use of conceivability as a test for logical possibility. (Accordingly, there is no sense to argue for the entailment from conceivability to logical possibility, if that entailment is a matter of stipulation.) This was corrected in Chalmers' later expositions of the zombie argument. In the corrected version, the stipulative link between conceivability and logical possibility is retained by defining (ideal negative) conceivability as the absence of *a priori* contradiction[17], and the proposition to use conceivability as a test for *logical* possibility is replaced with the argument for the claim that conceivability entails *metaphysical* possibility. This makes conceivability nearly synonymic with logical possibility, and makes the argument for the conceivability-entails-metaphysical-possibility claim tantamount to the argument for modal monism.

What is important here is that conceivability, in this objective sense, means the absence of *a priori* contradiction. Chalmers' reply to arguments like that of Frankish should be understood in this sense: it is obvious that there is no *a priori* contradiction in the notion of a phenomenal zombie; however, it is not obvious that there is no *a priori* contradiction in such notions as that of Frankish's anti-zombie.

However, a materialist like Frankish can challenge this claim by stating that for him, the notion of an anti-zombie *seems coherent*, and it is not for him more obvious that the notion of a zombie is coherent than that the notion of an anti-zombie is coherent. Can a zombist efficiently defend his point in face of this objection? I think that she can do it, by proposing a plausible explanation of where the incoherence may lie in the notion of an anti-zombie, and pointing out that there is no such a plausible explanation as to where the incoherence may lie in the notion of a zombie.

Basically, it is the general point made by Chalmers in *The Conscious Mind*:

> In general, a certain burden of proof lies on those who claim that a given description is logically *impossible*. . . . If no reasonable analysis of the terms in question points toward a contradiction, or even makes the existence of a contradiction plausible, then there is a natural assumption in favor of logical possibility[18].

[17] There is some complication, because Chalmers developed also the concept of positive conceivability and proposed two versions of the zombie argument, with negative conceivability and with positive conceivability. We can limit our discussion to the case of negative conceivability, because positive conceivability is a *stronger* notion — it is defined as the absence of *a priori* contradiction (that is, negative conceivability) plus detailed imaginability. If so, then if (as I will argue) an anti-zombie is not negatively conceivable, then it is not positively conceivable either.

[18] *Chalmers D.* The Conscious Mind. — P. 96.

For an anti-zombie, a reasonable analysis that points toward a contradiction in the notion (makes the existence of such a contradiction plausible) can easily be provided. Here, we begin with the notion of a purely physical copy of a human being — a creature that has all physical properties of a human being and has no other — nonphysical — properties. So far, all there is to this creature are physical properties, their bearers (having no other properties), such as atoms, arranged in complex spatial structures, and their physical interactions and movements; all these involve nothing subjective. Apparently, there are no subjective experiences in the picture. Then we attribute subjective experiences to this creature. It is at least very plausible that thereby we add to the initial notion (that of a purely physical copy of a human being) something that initially was not there (subjective experiences) and thereby we destroy this notion, since it forbids such additions. We can express this by the following simple formula:

a purely physical copy of a human being + subjective experiences ≠ a purely physical copy of a human being.

With the addition of subjective experiences, a purely physical copy of a human being ceases to be purely physical. The result of the addition is something that has other — nonphysical — properties. Thus, the idea of an anti-zombie (as a purely physical copy of a human being that has subjective experiences) is incoherent. That is, anti-zombies are inconceivable[19].

How can a physicalist (including functionalist) oppose this analysis? It would not do to oppose it with the claim that subjective experiences are identical with some brain states (or some functions realized by such states), and so if the latter are given, the former are there as well. Given the explanatory gap between the physical and the subjective (phenomenal), or

[19]The same case can be made in terms of supervenience. It is plausible that Frankish's description of the anti-zombie world as a world in which «consciousness is a physical phenomenon, supervening metaphysically on the world's microphysical features — in virtue of token identities, say» (*Frankish K.* The anti-zombie argument. — P. 654) is incoherent. If modal monism is accepted (and Frankish does not challenge it), then the metaphysical supervenience is a matter of *a priori* necessity or entailment. A metaphysically (logically) supervenes upon B if and only if B entails A as *a matter of conceptual relationship between A and B*. However, it seems clear that there is *no such entailment* from any possible fundamental (micro)physical facts to phenomenal facts. Whatever are the details of possible arrangements and dynamics of fundamental (micro)physical entities (on the assumption that such entities are not capable of having experiences, that there is nothing it is like to be a quark or a lepton), it seems clear that these would not entail there being mental subjects with their subjective experiences. If so (which is plausible, at least), then the idea that consciousness metaphysically supervenes on the physical is incoherent.

the hard problem of consciousness, the analysis provided is a reasonable analysis that points toward a contradiction in the notion (makes the existence of such a contradiction plausible). And that is all that is required of it.

Can this analysis be neutralized with a counter-analysis — a reasonable analysis that points toward a contradiction in the notion of a phenomenal zombie (makes the existence of such a contradiction plausible)? If we recognize that there is the explanatory gap from the physical to the phenomenal, and the hard problem of consciousness, we should also recognize that such a counter-analysis is impossible[20].

The decisive consideration here is the one to which Chalmers appealed in his formulation of the hard problem:

> The facts about experience cannot be an automatic consequence of any physical account, as it is conceptually coherent that any given process could exist without experience. Experience ... is *not entailed* by the physical[21].

Let us designate this claim as No Physical-to-Phenomenal Entailment, NPPE.

If NPPE is admitted, then there is «a reasonable analysis of the terms in question» that «points toward a contradiction, or even makes the existence of a contradiction plausible» for the case of an anti-zombie, and there is no such analysis for the case of a phenomenal zombie. Hence, the anti-zombie argument is impotent against the zombie argument.

On the other hand, if NPPE is denied, then the anti-zombie argument is just a red herring, because the denial of NPPE is tantamount to the denial of the conceivability of a phenomenal zombie (as well as the denial of there being the explanatory gap and the hard problem of consciousness).

It is no news at all (and does not require the idea of an anti-zombie to see) that one would hold that a phenomenal zombie is conceivable, only if one recognizes that there is the explanatory gap and the hard problem of consciousness. And, as I have argued above, those who recognize this should recognize also that a phenomenal zombie is conceivable whereas an

[20] There is a *prima facie* plausible analysis of how the idea of a phenomenal zombie can be incoherent that is *open for an interactionist dualist but not for a materialist.* If in the actual world, interactionist dualism is true, that is, immaterial consciousness causally contributes to the physical processes in the human brain, then a zombie would lack that causal factor, and the idea that in this case all physical processes could be just as with human beings seems *prima facie* incoherent.

[21] *Chalmers D.* Facing Up to the Problem of Consciousness // Journal of Consciousness Studies. — 1995. — Vol. 2. — P. 208.

anti-zombie is not. Some of them can still disagree with Chalmers that conceivability entails metaphysical possibility, but that is another story.

References

[1] *Frankish K.* The anti-zombie argument // Philosophical Quarterly. — 2007. — Vol. 57.

[2] *Marton P.* Zombies versus materialists: The battle for conceivability // Southwest Philosophy Review. — 1998. — Vol. 14. — P. 131-138.

[3] *Sturgeon S.* Matters of Mind: Conscious, Reason, and Nature. — London : Routledge, 2000.

[4] *Brown R.* Deprioritizing the A Priori Arguments Against Physicalism // Journal of Consciousness Studies. — 2010. — Vol. 17. — P. 47-69.

[5] *Chalmers D.* The Two-Dimensional Argument Against Materialism // Chalmers D. The Character of Consciousness. — New York : Oxford University Press, 2010.

[6] *Chalmers D.* Does Conceivability Entail Possibility? // Conceivability and Possibility / ed. T. Gendler and J. Hawthorne. — New York : Oxford University Press, 2002.

[7] *Chalmers D.* The Conscious Mind. — New York : Oxford University Press, 1996.

[8] *Chalmers D.* Materialism and the Metaphysics of Modality // Philosophy and Phenomenological Research. — 1999. — Vol. 59. — P. 477-8.

[9] *Chalmers D.* Facing Up to the Problem of Consciousness // Journal of Consciousness Studies. — 1995. — Vol. 2.

Надійшла до редакції 31 липня 2019 р.

Сепетий Дмитро Петрович
Кафедра суспільних дисциплін
Запорізький державний медичний університет
пр. Маяковського, 26
м. Запоріжжя
69035

Sepetyi Dmytro
Department of Social Disciplines
Zaporizhzhia State Medical University
Mayakovskogo ave., 26
Zaporizhzhia
69035

https://orcid.org/0000-0003-2110-3044

dmitry.sepety@gmail.com

10.31812/apd.v0i20.2598