



## 1. Introductory remarks

The dominant component of the modern “climate of opinion”<sup>1</sup> is the belief in the legitimacy of ontological naturalism, which can be defined as a statement that there exist no non-material substances or properties. On the basis of this definition (it can be called ‘substantive’) ontological naturalism is defined negatively, that is, by indicating the types of being whose existence is excluded, and is materialistic naturalism, because beings which are “excluded from existence” are those that are of non-material nature. Ontological naturalism, however, can also be defined in a different way than the above - as an ontology permitting the existence of only those types of being whose existence is postulated by the natural sciences. As can be easily seen, this definition has in fact a formal (and positive) nature, because on its basis, the specific “content” of naturalistic ontology is relativized to the present state of science or to the state of science in the final - “ideal” – stage of its development. At the present stage of science, ontological naturalism is indeed materialistic, although, as is known, entities whose existence is postulated by quantum mechanics, whatever exactly they may be - “ordinary” waves, probability waves or particles - are “material” in a sense distant from the colloquial understanding of this concept. However, an important and philosophically interesting question arises whether at a more advanced stage of the development of science, naturalism defined formally will not turn out to be closer to non-material ontology. In my further considerations, however, I assume, for

reasons of simplification, that both definitions of ontological naturalism - substantive and formal - are logically equivalent, so I pass over the problem of the possible differences between the “ontological content” of future scientific theories from the “ontological content” of contemporary theories. As part of the introductory remarks, it should also be mentioned that in the context of the philosophy of mind, materialistic naturalism leads to the rejection of dualistic views and to the acceptance of views based on the assumption of the so-called causal closure of the physical world (saying that for any physical event, if it has a sufficient cause for its occurrence, this cause has a physical nature, so that by determining the causal “origin” of any physical event, we will never go beyond the physical domain); these views are, for example, materialistic eliminativism (negating the existence of mental properties), theories of the identity of mental and physical properties, or theories of the supervenience of mental properties over physical ones.

I have mentioned that materialistic naturalism is an essential component of the modern “climate of opinion”. Naturally, this does not mean that materialistic naturalism is a position accepted by all thinkers; however, it is certainly the most popular position in the academic philosophical community (one could say that it created, „philosophical mainstream”). It seems, however, that its popularity is based on an extra-national basis: it does not follow that it has good philosophical justifications, but it is the result of (understandable) fascination with (undeniable) scientific achievements, which may lead to an unjustified belief that they justify the adoption of materialistic naturalism. Meanwhile, there exist good philosophical reasons for rejecting materialistic naturalism. One of the strongest reasons is, in my opin-

<sup>1</sup>This term is borrowed from Carl L. Becker (See C.L. Becker, *Państwo Boże osiemnastowiecznych filozofów*, transl. J. Ruzzkowski, Warszawa 1995, p. 9–29.

ion, an argument called „argument from the suicide of thought” (Gilbert Keith Chesterton), “the cardinal difficulty of naturalism” (Clive Staples Lewis) or “the argument from reason (Alvin Plantinga).. It will be the main subject of my reflections in this article. This argument, in its most general formulation, states that there is a conflict between the image of man as animal rationale and materialistic naturalism. The concept of animal rationale appearing in this argument contains two elements that can be called “procedural” and “effective (non-procedural)”. The first means that man is a reflective being, that is, guided by reasons in the course of shaping his beliefs and desires and of making his decisions; the second means that this process of being guided by reasons leads with substantial regularity, although of course not unfaithfully, to shaping true beliefs. In the analysis of this argument, I will distinguish two versions of it: non-evolutionary and evolutionary, and two of its interpretations: weaker, according to which this argument indicates the existence of a conflict between materialistic naturalism and the image of man as animal rationale, and stronger, according to which this argument not only indicates the existence of such a conflict, but additionally provides a reason for one of the sides of this conflict, as it also shows that materialistic naturalism is a self-defeating theory.

The first distinction has not appeared in the philosophical literature in a clear form (the two aforementioned versions are usually presented together), the second distinction has not appeared at all. However, they seem useful. The first one shows that the argument from the suicide of thought can formulate a conflict between materialistic naturalism and the image of a human being as animal rationale in two slightly different ways: the non-evolutionary

version tends primarily to prove that materialistic naturalism undermines the procedural component of the concept of animal rationale (and indirectly undermines the credibility of the effects of thinking, that is our beliefs), and the evolutionary version tends to show that materialistic naturalism undermines the non-procedural component associated with the tendency of animal rationale to formulate true beliefs (and therefore, directly undermines the credibility of our beliefs)<sup>2</sup>. The second distinction makes it possible to distinguish two different views of the philosophical implications of this argument.

## 2. Non-evolutionary version of the argument from the suicide of thought

I will start from citing several formulations of the argument from the suicide of thought in the non-evolutionary version:

- ◆ According to Gilbert K. Chesterton, it reads: “There is a thought that stops thought. That is the only thought that ought to be stopped”<sup>3</sup>. What thought, in the opinion of Chesterton, ought to be stopped? The one which says that our thoughts are no more than the activity of our brain since, if they were only this, , there would be no reason to believe

<sup>2</sup>Two subsequent points are based on a fragment of my book *Przeciw rozpaczy. O tragicznej wizji świata i sposobach jej przezwyciężenia* (Kraków 2014, p. 236–244); however, they constitute quite a significant modification of this fragment, consisting in distinguishing two interpretations of the argument from the suicide of thought and in connecting its two versions to the two elements of the concept of animal rationale (procedural and effective).

<sup>3</sup>G.K. Chesterton, *Orthodoxy*, New York 2010, p. 28.

that they are connected with reality at all. Consistent naturalism therefore leads, as Chesterton puts it, to „the suicide of thought“<sup>4</sup>.

◆ In a paper “When I am Dead” John B.P. Haldane formulated this argument in the following way

„For if my mental processes are determined wholly by the motions of atoms in my brain I have no reason to suppose that my beliefs are true. They may be sound chemically, but that does not make them sound logically. And hence I have no reason for supposing my brain to be composed of atoms” and: “if materialism is true, it seems to me that we cannot know that it is true. If my opinions are the result of the chemical processes going on in my brain, they are determined by the laws of chemistry, not of logic”<sup>5</sup>.

◆ C.P. Lewis presented this argument as follows: If minds are wholly dependent on brains, and brains on biochemistry, and biochemistry (in the long run) on the meaningless flux of the atoms, I cannot understand how the thought of those minds should have any more significance than the sound of the wind in the trees”<sup>6</sup>.

Now, I propose to distinguish two interpretations of this argument – the weaker and the stronger.

According to the weaker interpretation, this argument shows that it cannot be claimed at the same time that man is animal rationale (whose

fundamental distinction is thinking, which is an intentional and teleological process) and that materialistic naturalism, according to which thinking is essentially a causal-mechanical process, is true. This argument, therefore, forces us to choose between the image of man as animal rationale and materialistic naturalism<sup>7</sup>. Of course, materialistic naturalism can be compatible with a very thin version of the concept of animal rationale, according to which man is an instrumentally rational being, that is, he selects effective means to achieve his goals (technically: maximizes his utility function), and the question of his rationality is assessed not on the basis of the course of his thought processes, but on the observation of his behavior<sup>8</sup>.

However, one can ask why the image of a human being as animal rationale is contrary to materialistic naturalism (the above formulations of the argument from the suicide of thought do not provide a clear answer to this question). It is because determinism is, arguably, the key element of materialistic naturalism. This problem has been clearly presented by John Searle in his book *Rationality in Action*. According to Searle, the characteristic feature of rational action is that our deliberation and voluntary actions is not experienced by as caused by

---

<sup>7</sup>This interpretation is “weaker” in the sense that the second interpretation (according to which the argument from the suicide of thought undermines materialistic naturalism) is stronger. However, this is not a “weak” interpretation, because in its light, materialistic naturalism is also problematic; it can be convincingly argued that an indication of the existence of a conflict between materialistic naturalism and rationality is an argument against the former due to the fact that the view that we are stuck in a systematic illusion with regard to the real nature of our rationality seems to have little credibility.

<sup>8</sup>I omit the criticism from the cognitive science of the concept of man as an instrumentally rational being, since it is irrelevant to my considerations in this article. This criticism refers precisely to this “thin” concept of human rationality (instrumental rationality), which is not essential to my considerations of the classic concept of animal rationale.

---

<sup>4</sup>Ibidem, p. 25.

<sup>5</sup>As cited in: K.R. Popper, J.C. Eccles, *The Self and its Brain. An Argument for Interactionism*, Lon-don–New York 2006, p. 75.

<sup>6</sup>C.P. Lewis, ‘Is Theology Poetry’?, *augustinecollective.org*, p. 15.

causally sufficient conditions. Accordingly, Searle writes that "the operation of rationality presupposes that there is a gap between the set of intentional states on the basis of which I make my decision, and the actual making of the decision". This "gap" is an undeniable feature of our everyday experience. Let's imagine, for example, that we are in a restaurant and the waiter asks us what we want to order. As Searle notes, we cannot say, "Look, I am a de-

and desires cause"<sup>10</sup>, because such a refusal to use your freedom is itself only intelligible to you as an exercise of freedom"<sup>11</sup>. This "gap", if it is empirically real, must, according to Searle, perform the activity of the non-Hume self, i.e. the rational self which cannot be reduced to a cluster of perceptions. Searle's conclusion is consistent with the first interpretation of the argument from the suicide of thought: the experience of the "gap" does not show that

Naturalizm materialistyczny jest zasadniczym elementem składowym współczesnego 'klimatu opinii'. Nie znaczy to oczywiście, że naturalizm materialistyczny jest stanowiskiem przyjmowanym przez wszystkich myślicieli; jest jednak z pewnością stanowiskiem najbardziej popularnym w akademickim środowisku filozoficznym (tworzy, można by rzec, 'filozoficzny mainstream'). Wydaje się jednak, że jego popularność ma podłoże poza-racjonalne: nie wynika stąd, że ma dobre uzasadnienia filozoficzne, lecz stanowi rezultat (zrozumiałej) fascynacji (niezaprzeczalnymi) osiągnięciami nauki, mogącej jednak prowadzić do nieuprawnionego przekonania, że uzasadniają one przyjęcie naturalizmu materialistycznego.

terminist, che sarà, sarà. I will just wait and see what I order! I will wait and see what my beliefs

it is "empirically real", but that it is "psycho-

<sup>9</sup>J. Searle, *Rationality in Action*, Cambridge, Mass.–London 2001, p. 13.

<sup>10</sup>*Ibidem*, p. 14.

<sup>11</sup>*Ibidem*.

logically real<sup>12</sup>. In other words, Searle does not claim that his argument (which is essentially a version of the argument from the suicide of thought) shows that determinism is false; what he claims is that it points at the necessity of making a choice between the image of man as animal rationale and determinism, because determinism leads to the conclusion that this “gap”, or psychological indeterminism, is an illusion (and thus also rational deliberation is an illusion, since which is its presupposition). If we assumed that (the subjectively experienced) indeterminism at the psychological level is accompanied by determinism at the level of the functioning of the brain (i.e., that brain states underlie or are identical to mental states and remain in the relations of the deterministic causal influence with each other), it would mean that we defend epiphenomenalism, according to which subjective processes do not play any causal role, and our sense of agency is a systematic illusion<sup>13</sup>.

<sup>12</sup>Ibidem, p. 269.

<sup>13</sup>It is worth adding, however, that Searle formulates an additional argument (ibidem, p. 285), which, in his opinion, shows that determinism is very problematic, and hence that of two parts of the alternative constructed by him: animal rationale vs. determinism, stronger reasons speak for accepting the first disjunct. This argument says that epiphenomenalism is contrary to everything we know about evolution, because it means that the system of conscious decision-making, which is incredibly developed, complex, and consuming much energy, has no significance for the life and survival of the organism. This argument challenging epiphenomenalism also undermines determinism, because determinism leads to epiphenomenalism. I only invoke this interesting argument; I do not analyze it more closely, because it would divert me from the main topic of this paper. It is only worth adding here that according to Popper, one of the precursors to the argument from the suicide of thought was Epicurus, who formulated it, like Searle, not against materialism (he was a materialist himself), but against determinism: „He who says that all things happen of necessity cannot criticize another who says that not all things happen of necessity. For he has to admit that the assertion also happens of necessity” (As cited in: K.R. Popper, J. Eccles, *The Self and Its Brain*, p. 75). Epicurus, however, wanted to say more than Searle, namely, that determinism is a self-defeating view, and not, as Searle claimed, one that is contrary to the concept of

According to the second - stronger - interpretation of the argument from the suicide of thought, it does not only point at the necessity to choose between materialistic naturalism and the image of man as animal rationale, but it also shows that materialistic naturalism is an inconsistent, self-defeating view. More specifically: this argument shows that the naturalistic statement “materialistic realism” is true, although it is not formally incoherent (i.e., it is not internally contradictory), it is self-referentially incoherent, i.e. undermines itself, because in the light of what it preaches - materialistic naturalism - it is the result of a series of causal and mechanical events having nothing to do with reasons,-arguments put forward by the free mind to support a certain conviction<sup>14</sup>.

Therefore, even if materialistic naturalism is true, we will never be able to know that it is true, and the discovery of this view (assuming it to be true) can only be conceived as a result of a happy coincidence (that is, strictly speaking, of the biological/chemical/physical pro-

animal rationale and unlikely (as it implies an unlikely view - epiphenomenalism).

<sup>14</sup>The self-referential incoherence (inconsistency) of materialistic naturalism is similar to the inconsistency associated with the liar paradox and the inconsistency of cognitive relativism. There are, however, some differences between them that allow us to distinguish three slightly different forms of self-referential inconsistency; I suggest to call them “direct stronger”, “direct weaker” and “indirect”, respectively. The first (characteristic for the liar paradox) occurs when the truthfulness of a given sentence (for example the sentence “this sentence is false”) implies logically its falsehood, and its falsehood implies logically its truthfulness. The second (characteristic for a relativistic position, according to which no sentence is objectively true) occurs when the truth of a given sentence (for example “relativism is true”) implies logically its falsehood, but its falsehood does not imply logically its truthfulness. The third (characteristic for materialistic naturalism) occurs when the truthfulness of a given sentence (for example the sentence “materialistic naturalism is true”) does not logically entail its falsehood, but it provides strong reasons for its falsehood. The self-referential incoherence of materialistic naturalism is therefore indirect

cesses which caused the appearance in our minds of the conviction that materialistic naturalism is true). For example, a strong interpretation was adopted by Karl Popper who wrote that “although it does not show [the argument from suicide of thought – W.Z.<sup>15</sup>], that materialism destroys itself, I suggest that it shows that materialism is self-defeating: it cannot claim to be supported by rational argument”<sup>16</sup>.

It is worth noticing that Haldane withdrew this argument<sup>17</sup>, claiming that by formulating it, he erroneously assumed that action in accordance with natural laws cannot be an action consistent with the laws of logic. However, this “self-criticism” seems unjustified. The argument from the suicide of thought might be wrong if it were based on the assumption that something that works according to the laws of physics or chemistry cannot simultaneously act in accordance with the laws of logic (an example of an object that works simultaneously under both types of law is, for example, a com-

puter<sup>18</sup>). I used the phrase “might be wrong”, because it should be noted that no computer had discovered the laws of physics, written a novel, proposed an original philosophical argument or composed a sonata: acting in accordance with the laws of logic is neither a creative nor a self-aware action. However - regardless of what we think about the logical and creative capacities of calculating machines and their awareness - the key issue is that the argument is not based on the aforementioned assumption: it does not say that the naturalistic thesis is false because the naturalistic mind (meaning one whose thought processes have a mechanically-deterministic character) acts against the laws of logic. In a weaker interpretation, the argument states that assuming that the mind is such as described by materialistic naturalism, one cannot accept the image of man as animal rationale, and, in a stronger interpretation, that materialistic naturalism is self-referentially incoherent.

### 3. Evolutionary version of the argument from the suicide of thought

The evolutionary version of the argument from the suicide of thought was proposed by Alvin Plantinga; he called it an “evolutionary argument against naturalism”, because he based his skepticism towards the epistemic value of the beliefs of the naturalistic mind on a specific thesis of materialistic naturalism stating that the mind is a product of undirected natural selection<sup>19</sup>. In Plantinga’s opinion, there are no good reasons to assume that the mind which is

<sup>15</sup>Popper did not use this term.

<sup>16</sup>K.R. Popper, J. Eccles, *The Self and Its Brain*, p. 76. It should be added, however, that this argument does not occupy a central place in Popper’s criticism of materialistic naturalism. A more important role is played by the argument that material processes cannot provide logically correct thinking patterns, so these patterns must come from outside the material world (Popper “placed” them in the so-called world 3, a world different from the world of physical objects - world 1, and world of mental states - world 2). This Popperian argument resembles the criticism of psychologism carried out by Edmund Husserl, that the laws of logic are empirical generalizations of how we actually think. However, Husserl emphasized that psychologism leads to “species relativism” (“anthropologism”), that is, a View that there is no absolute truth, but only one relative to a given genre. It also means, according to Husserl, that psychologism is a self-defeating view. This last conclusion is consistent with a stronger interpretation of the argument from the suicide of thought.

<sup>17</sup>Publishing in 1954 an essay titled *I repent an error in Literary Guide*.

<sup>18</sup>See K.R. Popper, J. Eccles, *The Self and Its Brain*, p. 76.

<sup>19</sup>

the product of a naturalistically interpreted natural selection (i.e., the “undirected” selection) has a tendency to generate real beliefs; it can only be said that such a mind will have a tendency to generate beliefs conducive to survival and reproduction. Plantinga notes that thinkers who are hostile to theism, such as Thomas Nagel and Patricia Churchland, as well as those neutral towards theism, such as Charles Darwin also agree to this thesis. Plantinga invokes three suggestive quotes from these thinkers. Nagel writes: “If we came to believe that our capacity for objective theory were the product of natural selection, that would warrant serious skepticism about its results”<sup>20</sup>. Churchland claims:

„Boiled down to essentials, a nervous system enables the organism to succeed in the four F’s: feeding, fleeing, fighting and reproducing. The principle chore of nervous systems is to get the body parts where they should be in order that the organism may survive.....Improvement in sensorimotor control confer an evolutionary advantage: a fancier style of representing is advantageous so long as it is geared to the organism’s way of life and enhances the organisms chances of survival. Truth, whatever that is, definitely takes the hindmost”<sup>21</sup>. A similar thought was formulated by Darwin (in a letter to William Graham of 1881), who wrote that: “With me the horrid doubt always arises whether the convictions of man’s mind, which has been developed from the mind of the lower animals, are of any value or are at all trustwor-

thy. Would any one trust in the convictions of a monkey’s mind, if there are any convictions in such a mind”<sup>22</sup>.

Plantinga calls the first premise of his argument “Darwin’s doubt” and presents it in the form of a conditional probability:  $P(R/N \ \& \ E) = \text{low value}$ , where R means the phrase “our cognitive abilities are reliable”, and N & E means the conjunction of the sentences “naturalism is true”, “our cognitive abilities have originated as a result of the operation of natural selection”. The second premise is as follows: whoever accepts N & E and the first premise, has ipso facto a defeater of R (i.e., of the thesis about the credibility of our cognitive powers). The third premise states that the argument undermining R ipso facto undermines all other beliefs, including N & E (because these beliefs are created by cognitive powers whose credibility is undermined by the first premise). The fourth premise says that if N & E undermines N & E, N & E is a self-defeating, self-referentially incoherent thesis. Thus, the conclusion is that: N & E is a self-defeating, self-referentially incoherent thesis; as Plantinga puts it: „If someone accepts N & E and sees that  $P(R/N \ \& \ E)$  is low, then she have a defeater for N & E, a reason to reject it, a reason to doubt or be agnostic with respect to it”<sup>23</sup>. In other words, according to Plantinga, if the mind created as a result of “undirected” evolution (that is, naturalistically interpreted evolution), generates the conviction that the conjunction “materialistic naturalism” cum theory of evolution” is true, then he has no right to believe in the truth of this conviction. This is a stronger interpretation of this

<sup>20</sup>T. Nagel, Death (in:) T. Nagel, *Mortal Questions*, Cambridge 1979, p. 8; Nagel explicated this matter in the book *Mind and Cosmop. Why the Materialist Neo-Darwinian Conception of Nature is Almost Certainly False*, Oxford 2012.

<sup>21</sup>P. Churchland, *Epistemology in the Age of Neuroscience*, *Journal of Philosophy* 1987, Vol. 84, p. 548.

<sup>22</sup>As cited in: A. Plantinga, *Where the Conflict Really Liep. Science, Religion, and Naturalism*, p. 316.

<sup>23</sup>Ibidem, p. 345.

argument (assumed by Plantinga). The weaker interpretation is essentially limited to its first premise and states that we have to choose between the human image as animal rationale, whose thinking processes are directed towards the truth, and the conjunction of “materialistic naturalism cum theory of evolution”.

Let me now turn to the assessment of Plantinga’s argument. Undoubtedly, its essential point is the first premise. Plantinga claims that it is true, but it is not obvious. It may be objected that although, in the light of the naturalistically interpreted theory of evolution, the purpose of the human mind is not to generate real beliefs, but adaptive behaviors, the organism whose cognitive system tends to generate false beliefs will not effectively take actions serving the purpose of evolutionary goals. Plantinga considers this objection, but gives an unconvincing answer: he writes that he does not ask “how things are”, but how they would look if both evolution and materialistic naturalism were true.<sup>24</sup> Well, assuming that “undirected evolution” could at all lead to the creation of homo sapiens (this is the assumption with which Plantinga disagrees, but must accept it for the purposes of his “evolutionary argument”), it seems that things would be exactly the same or at least we have no reason to claim that they would not be the same. A more fruitful strategy of defense against this allegation is to divide beliefs into “relevant to evolutionary success” and “not relevant to evolutionary success” (e.g. metaphysical beliefs), because it is clear that, first, the thesis that the naturalistic mind has no tendency to generate real convictions of the second kind is convincing and, secondly, that it is precisely this type of conviction that is the subject of dispute in the context of which

Plantinga formulates his argument. In fact, Plantinga mentions this strategy, writing that it leads to a “weaker” version of his argument<sup>25</sup>. However, at the key point, it is as strong as the original version, because it justifies the thesis that the conjunction of materialistic naturalism and the theory of evolution cannot be rationally accepted. It seems, therefore, that Plantinga unnecessarily places emphasis on a more controversial (stronger) version of his evolutionary argument against materialistic naturalism (cum theory of evolution), since the weaker version undermines it equally effectively.

At the end of this point, it is worth considering the relation between the non-evolutionary and evolutionary version of the argument from the suicide of thought. The first premise of Plantinga’s argument is that, assuming that materialistic naturalism and evolutionary theory are true, the probability that our cognitive powers will generate true beliefs is low. The above-mentioned objection against this premise does not impact the version of the argument proposed by Haldane, Chesterton and Lewis. For this version does not say anything directly about the probabilistically measured degree of reliability of our cognitive abilities understood as a tendency to generate real beliefs: it only states that if our thinking process were as they presented by the materialistic naturalist, there would be a contradiction between the image of man as animal rationale and materialistic naturalism (weaker interpretation), or that there would be no reasons to accept the results of the cognitive activity of our mind, and thereby naturalism would be a self-defeating view (stronger interpretation). It should be noted, however, that the argument in Plantinga’s version would prove something

<sup>24</sup>Cf. *ibidem*, p. 335.

<sup>25</sup>31 Cf. *ibidem*, p. 348–349.

more and something less than the argument of Haldane, Chesterton and Lewis. It would prove something more than just a lack of rational justification for materialistic naturalism, assuming that it is real, namely it would demonstrate that the probability that the naturalistic mind would encounter a true outlook, is low (for the first premise of Plantinga's argument - the so-called Darwin's doubt – provides information precisely on the probabilistically measured degree of reliability of our cognitive abilities). It is quite surprising that Plantinga does not mention that his argument proves it (although it seems that this is a natural conclusion<sup>26</sup>); he argues that this argument makes it possible to conclude that one cannot rationally simultaneously accept materialistic naturalism and the theory of evolution, and that such a combination of beliefs is self-defeating<sup>27</sup>. So, he thinks in essence (though he does not formulate this explicitly), that his own argument proves something less than the argument of Haldane, Chesterton and Lewis<sup>28</sup>, the argument which, if it is accurate, shows that materialistic naturalism itself is a self-defeating view. This thesis is stronger because if one cannot rationally accept materialistic naturalism, one cannot rationally accept any conjunction of views of which materialistic nature is an element.

#### 4. Towards dualism

The argument from the suicide of thought in

<sup>26</sup>This is probably due to the fact that Plantinga does not distinguish between the evolutionary and non-evolutionary versions of the argument from the suicide of thought.

<sup>27</sup>Ibidem, p. 349.

<sup>28</sup>Although Plantinga does not compare his own argument with their argument; he writes only that a similar argument can be found in C.P. Lewis.

the non-evolutionary version shows, according to the weaker interpretation, that the image of a human being cannot be reconciled as animal rationale with materialistic naturalism, and according to a stronger interpretation, that naturalism is a self-referentially incoherent view. The evolutionary version of this argument seems more controversial than the non-evolutionary version, it is not clear whether Darwin's doubt is justified, but if the evolutionary version is correct, it provides additional support for a stronger interpretation of the non-evolutionary version (the argument from the suicide of thought in the evolutionary version leads directly to a strong interpretation, although, as I have mentioned above, its weak interpretation is possible). What are the additional implications of this argument? From the point of view of the very content of the argument from the suicide of thought argument, the direct implications concern the philosophy of mind, but before I present them, I will mention the consequences that this argument has according to Plantinga in the context of the discussion about the relationship between science and religion. According to Plantinga, this argument constitutes support for the thesis that: „There is superficial conflict but deep concord between science and theistic religion, but superficial concord and deep conflict between science and naturalism”<sup>29</sup>.

However, I will not continue to discuss this matter, because it goes beyond the theme of this article. As I mentioned in point 1, materialistic naturalism in the context of the philosophy of mind - a branch of philosophy particularly important from the point of view of the argument under consideration - leads to views based on

<sup>29</sup> A. Plantinga, *Where the Conflict Really Lies. Science, Religion, and Naturalism*, p. IX.

## Naturalism as the suicide of thought

the (naturalistic) assumption of the so-called causal closure of the physical world. This assumption leads to the rejection of the so-called mental causation; if the physical world is causally closed, the conscious experience is causally irrelevant<sup>30</sup>. It is therefore necessary to find, rejecting materialistic naturalism, the approach to the relation of the mind and the brain, which does without the assumption about the causal closure of the physical world and, as a result, allows mental causation.

In the philosophy of mind, two approaches of this kind have been proposed: non-emergent, substance dualism and emergent, substance dualism (the most sophisticated version of the first stance was proposed by Richard Swinburne<sup>31</sup>, of the second – William Hasker<sup>32</sup>). Both

assume that mind and body are two different substances - spiritual and material. Substance emergent dualism implies that the mind is the product of the brain (or more generally, of the body) and is not a separate element “attached from the outside” to the body. Therefore, it assumes that matter, after reaching a sufficiently high level of complexity, can lead to the “emergence” of the mind..

It is a common feature of emergent substance dualism and materialistic theories of mind, and the fundamental difference between non-emergent substance dualism and emergent substantive dualism. It should be emphasized, however, that the latter is a version of the so-called strong and not weak emergentism, which radically separates it from materialist theories. Weak emergentism is in fact a variant of materialistic naturalism, whereas strong emergentism assumes that the emergent properties are ontologically irreducible to the properties from which they “emerged”, and therefore are radically new in relation to them. Therefore, emergent substance dualism states that the mind is ontologically irreducible to the brain (body) from which it “emerged”, and therefore is radically new and has “causal powers” in relation to it. This is, of course, a thesis contrary to the thesis of the causal closure of the physical world. It should be emphasized that emergent dualism in Hasker's version is a

---

<sup>30</sup>Also see J. Kim, *The Myth of Nonreductive Materialism*, *Proceedings and Addresses of the American Philosophical Association* 1989, Vol. 63, No. 3, p. 31–47. Kim argues that the belief that the so-called non-reductive versions of materialism (for example Donald Davidson's anomalous monism) allow us to retain mental causation is a myth, although Kim's thesis is not used as an argument for a non-materialistic theory of mind, but for the reductionist materialist theories (a version of the materialistic theory defended by Kim says that the mind supervenes on the states of the brain).

<sup>31</sup>See R. Swinburne, *Mind, Brain, and Free Will*, Oxford 2013.

<sup>32</sup>See W. Hasker, *The Emergent Self*, London 1999. Non-emergent substance dualism and emergent substance dualism are, as it seems, the only stances which can be reliably defended upon rejection of materialistic naturalism. Other stances proposed in the philosophy of mind as an alternative to materialistic naturalism seem totally unconvincing. For example, I will quote two hypotheses considered by Searle in the context of his criticism of psychological determinism and epiphenomenalism (See J. Searle, *Rationality in Action*, p. 285–289) and I will mention the so-called property dualism. According to the first hypothesis considered by Searle, indeterminism at a neurobiological level corresponds to indeterminism at a psychological level. According to the second, consciousness is a feature of the entire system, which is the human brain, and although this system functions in a causal manner, it does not act on the basis of the deterministic principle (causally sufficient conditions). But Searle himself admits that the two hypotheses encounter serious difficulties: the first identifies mental acts with events from the quantum level, and thus in fact reduces freedom to quantum randomness. The second is very

---

vague: it is not known how the system can be causal, but not based on the principle of causally sufficient conditions. Searle is aware of the weakness of both hypotheses and ultimately refuses to accept any of them, acknowledging that none of them is “intellectually attractive”. However, he considers them more convincing than dualism. It is difficult to agree with this position. As far as the dualism of property (mental and physical) is concerned, it does not explain the features of animal rationale, whose explanation is the greatest advantage of substance duality: the unity of consciousness, the sense of subjectivity, free will. On the basis of property dualism d, these qualities are no less mysterious than on the basis of monistically materialistic theories of mind.

substance duality; it therefore assumes that it is not so much psychological properties which are emergent, but the psychological “individual”, who “emerges” as a result of a specific configuration of material components of the brain and nervous system. As Hasker aptly observes, only this assumption is consistent with our strong intuition that the mind shows the unity of consciousness and free will.

Substance dualism (in both versions) leads to rejecting the thesis of the supervenience of mental states on the physical ones. This thesis includes two distinct components, though usually presented together and not clearly separated: (1) There can be no difference between objects at the mental level without a difference between them at the physical level, although there may be a difference between them at the physical level without any difference at the mental level (the so-called multiple realization of mental states); therefore, two objects with identical physical properties must have identical mental properties<sup>33</sup>; (2) There exists a one-sided relationship between both levels - the dependence of the mental level on the physical level; therefore, it is changes at the physical level which cause changes at the mental level, not vice versa. Substance dualism implies that change at the mental level not caused by a change at the physical level is possible (it is usually assumed that a change at the physical level corresponds to the change at the mental level; however, this relationship is not necessary) and that changes at the physical level can be caused by changes at the mental level.

Substance dualism, therefore, is contradictory both to the second element of the thesis of supervenience (because it rejects the assumption of the unilateral dependence of the mental level on the physical level), as well as to the first element (because it allows the possibility of a difference between objects at the mental level without the difference between them at the physical level). Substance dualism leads to the conclusion that the physical properties of the organism “evolve” differently,

if they are accompanied by mental states, because the latter may causally affect physical states<sup>34</sup>. The acceptance of mental causation is consistent with the intuition that mental processes are irreducibly teleological and intentional, and as a result are irreducible to brain processes that have a nonteleological and non-intentional explanation, i.e., causal-mechanistic.

It is difficult to decide which of the presented versions of substance dualism presented is more convincing. A closer examination of this issue goes beyond the scope of my research in this article. It is worth mentioning, however, that the specific problem of the non-emergent version is the explanation of the origin of the mind (in order to solve this problem, its advocates usually enter teleological considerations); and the specific problem of the emergent version is to provide a law of nature, explaining how matter can “cause emergence” of an immaterial, radically new kind of being (the mind) within itself, and (if such law or laws could be formulated) an explanation of how one can rec-

<sup>33</sup>According to the slogan: no mental difference without a physical difference. If this change occurs within one possible world, supervenience is called “weak”; and if in all possible worlds - “strong”. On the subject of different understandings of supervenience, see particularly J. Kim, *Supervenience and Mind*. Selected Essays, Cambridge 1993.

<sup>34</sup>On the basis of materialistic naturalism, consciousness does not play any role in the manner in which matter operates: all brain processes have a causal-mechanistic explanation, which will not change if we assume that these processes are accompanied by consciousness.

oncile the conviction of the individual character of each mind (as it is presented in the form of law or laws of nature) with this general concept of “transition” from matter to mind<sup>35</sup>. The common problem of both versions of substance dualism is the explanation of the interaction between the mind and the brain, and in particular the indication of the mechanism of mental causation, that is, the exact description of how the mind affects the brain. However, this problem is not as serious as materialistic naturalists want to consider it. In order to solve it, modern dualists often turn to (with promising results) quantum mechanics. The scheme of the use of these achievements in the context of dualistic theories of the mind, postulating the ontological irreducibility of mental states (or at least some subset of mental events<sup>36</sup>) to neural processes can be presented, in a broad outline, in the following way<sup>37</sup>. The mind can affect the brain in two related ways: it can lead to the so-called reduction of the wave function describing the patterns of neuronal activity which are in the quantum superposition; as a result of this reduction (or “collapse”) one of the possible patterns of the neuronal scheme is “actualized” (in this approach, intentional

acts of consciousness leading to the physical reduction of the wave function are therefore analogous to the physical acts of measurement); the reduction, however, can also “support” specific quantum superpositions of neural activity patterns. In quantum theory of mind used by dualists<sup>38</sup> particular emphasis is placed on the role of attention as the basic act of the mind causally affecting the brain by reducing the wave function or maintaining the superposition of brain activity patterns. This second way in which the mind interacts with the brain is described by the so-called Quantum Zeno Effect. This effect consists in the fact that continuous observation of the unstable elementary particle “protects” it against decay. Thus, the evolution of the quantum system can be “frozen” by taking its continuous measurement by means of macroscopic devices. According to the presented hypothesis, the mind can affect the brain precisely through the Zeno effect, sustaining the brain in the superposition of certain states. In other words, in the context of the mind-body problem, the quantum Zeno effect means sustaining through the attention) of a certain “stream of consciousness” which would otherwise be dispersed. Therefore, the efforts of the conscious mind contribute to the creation of a specific pattern of neural activity, which then becomes a pattern of action<sup>39</sup>.

<sup>35</sup>These arguments against emergent substantive duality are formed by, among others, R. Swinburne (See R. Swinburne, *Mind, Brain, and Free Will*, p. 140–173)

<sup>36</sup>For example, R. Swinburne claims that only the so-called pure mental states, which he calls intentions (he identifies them with volitional states) and not e.g., emotions, have an irreducible nature.

<sup>37</sup>See for example J. Schwartz, H. Stapp, M. Beauregard, *Quantum physics in neuroscience and psychology: a neurophysical model of mind-brain interaction*, *Philosophical Transactions of the Royal Society B: Biological Sciences* 2005, Vol. 29, No. 360(1458), p. 1309–1327; H. Stapp, *Mind, Matter and Quantum Mechanics*, Berlin–Heidelberg 2004; H. Stapp, *Mindful Universe: Quantum Mechanics and the Participating Observer*, Berlin–Heidelberg 2007; K.R. Popper, J. Eccles, *The Self and Its Brain*; R. Swinburne, *Mind, Brain, and Free Will*, p. 113–117.

<sup>38</sup>It should be mentioned that also non-dualistic quantum theories of the mind are constructed. For example, according to the theory of mind presented above, the acts of consciousness are not identical with the reduction of the wave function, but they cause it; whereas in some other quantum theories of the mind, the act of consciousness is considered to be identical with the reduction of the wave function, and not to be its cause (the reduction is considered to be a purely random event). In the light of these last theories, the act of consciousness is not an act of the subject for which he can take responsibility, but an act “happening” to a subject that cannot be attributed to him according to any reasonable theory of responsibility.

<sup>39</sup>In this context, it is worth adding that research conducted on the border of quantum mechanics, philosophy of mind

In summary: a modern substance dualist can convincingly say that the mind acts on the brain as an immaterial cause through the laws of quantum mechanics. Therefore, it is not helpless in the face of the problem of communication of the spiritual mind with the material body which Descartes struggled ineffectively with, assuming - without good reasons - an eccentric hypothesis that this communication takes place through the pineal body<sup>40</sup>, and, as a consequence, drawing an easy, but unjustified objection on substance duality that such communication cannot take place at all. The presented (very adumbratively) hypothesis of the “quantum” interaction between the mind and the body is of course controversial. It has found many critics who, among other things, claimed that the fact that the mind in quantum models does not have its own wave function (because it is nonphysical), and yet it can affect the brain, leads to the conclusion that there may be any number of “spiritual minds” that can affect quantum systems. However, this accusation itself is inconclusive: it only indicates at the possible but not necessary consequence of this position. The allegation that this hypothesis undermines the assumption of a completely random nature of events from the quantum level is also inconclusive. In fact, if this hypothesis is true, a subset of quantum

events, namely those affected by our mind, is not random. This may simply indicate that the assumption of a completely random nature of events from the quantum level does not apply to all events of this kind. The dispute of the legitimacy of this hypothesis seems to be open and it will probably remain such for long. However, some general issues seem clear. First of all, certainly substance duality is not in contradiction with the achievements of science (for it is not known what such a contradiction might be, that is, what concrete scientific hypotheses would be conflicting or would be at all contradictory to it). Secondly, if the argument from the suicide of thought, which undermines materialistic naturalism, is apt, then the probability that one of the two versions of substance duality is true must be considered high. Thirdly, irrespective of philosophical arguments for and against substance duality, it should be noted that substance duality is also a common sense and intuitive image of the functioning of the mind, in contrast to materialistic monism. Accepting the latter would mean a revolution not only in our understanding of the mind, which is based on such - fictitious from the point of view of materialist monism - concepts as self, consciousness, free will, mental causation, but also in our understanding of the basic notions of ethics and law, and in particular, of the concept of responsibility, which, on the basis of materialistic monism, becomes very problematic. This fact is obviously not a direct argument for dualism (it cannot be logically ruled out that such a revolution brings us closer to the truth about our mind), but it would be extraordinary if it turned out that for centuries, we have been in a systematic illusion about the nature of our mind and the grounds of responsibility.

---

and cognitive psychology shows that the effectiveness of various cognitive-behavioral techniques (such as directed attention or cognitive redefinition of the situation) in regulating our emotional reactions, treatment of obsessive-compulsive disorders, depression or phobia, can be convincingly explained by adopting some version of the mind-brain dualism and, consequently, by acknowledging the causal effectiveness of mental states (See, e.g. J. Schwartz, P. Begley, *The Mind and the Brain: Neuroplasticity and the Power of Mental Force*, New York 2002).

<sup>40</sup>But also, at that stage of scientific development, not being able to find good arguments for the pineal hypothesis or for any other hypothesis explaining the interaction of *res cogitans* with *res extensa*.

## 5. Further conclusions

As noted above, the argument from the suicide of thought is directed against ontological naturalism in the materialistic version. It does not refer to ontological naturalism in the non-materialistic version (this is currently a purely hypothetical view, however, it cannot be ruled out that, at some further stage of the development of science, it will become an appropriate version of ontological naturalism in its formal interpretation), nor to naturalism, which can be called “soft” (contrary to materialistic naturalism - “hard”). Soft naturalism is an uncontroversial view that in the analysis of philosophical problems, one should refer to the achievements of science and that the ontology proposed as a result of this analysis cannot be contradictory to the results of science; this view, however, rejects the assumption that these achievements allow us to settle these problems, and the assumption that problems that cannot be resolved by appealing to science are pseudo-problems or apparent problems. The argument from the suicide of thought, however, has implications not only for the assessment of materialistic naturalism, but also for the so-called methodological naturalism. It seems that one can distinguish its two variants: methodological naturalism in the strict sense and methodological naturalism in the broad sense. The first states that only scientific methods should be used in solving scientific problems, that is, one should act as if materialistic naturalism was true, and therefore not accept any hypotheses postulating the existence of “non-natural” beings in the explanation of the natural world. The second says that solely scientific methods should be applied, not only in scientific research, but also in the philosophical considerations (the boundary between science and philosophy is thus blurred). It is based on the assumption of materialistic naturalism, according to which

all the “real” philosophical problems can be solved by means of scientific methods (if they cannot be solved in this way, it means that they are apparent problems or pseudo-problems); in other words, according to methodological naturalism in the broad sense, the ontology implied by these solutions will not only be consistent with the ontology of the best scientific theories, but also identical with it, that is, materialistic. Thus, it differs fundamentally from methodological naturalism in the strict sense: it presupposes that materialistic naturalism is true, whereas the latter only requires to act as if materialistic naturalism were true. Now, if the argument from the suicide of thought is apt, and therefore if materialistic naturalism is false, then the methodological naturalism in the broad sense should be rejected because it implies the truth of materialistic naturalism. Moreover, the argument from the suicide of thought does not undermine the methodological naturalism of the strict sense (since the latter does not imply the truth of materialistic naturalism), it appears, however, that it cannot be a categorical or “global” rule on the basis of scientific research: it is not adequate in these areas, in which science studies (or attempts to study) issues related to the problem of mind-body relations (such as, for example, free will, the specificity of human action, nature of consciousness), that is, those issues which are strictly connected to the controversy between the adherents of the materialistic and dualistic theories of the mind .

