

# Adaptivity and Truth. A Critique of Plantinga's Reasoning against Evolutionary Reliabilism<sup>1</sup>

**Andrea Fábiková**

Faculty of Arts, Comenius University, Bratislava

andrea.fabikova@uniba.sk

---

## Abstract:

The paper analyses the arguments put forward by Plantinga to justify his refutation of evolutionary reliabilism, i.e. the claim that the probability that the cognitive faculties, developed in the process of unguided evolution, are reliable is low. I argue that all the thought experiments offered by Plantinga to justify this thesis suffer from a common defect – they disregard the condition of evolution or fail to take it into account properly. In addition, I argue that pointing out the difficulties that naturalistic approaches have in explaining mental causation does not lead to Plantinga's conclusion that in a naturalistic world there would be no mental causation whatsoever.

**Keywords:** Alvin Plantinga, Darwin's doubt, evolutionary reliabilism, the evolutionary argument against naturalism

**DOI:** <https://doi.org/10.46854/fc.2021.3s62>

---

Looking at human cognitive faculties and their function from the evolutionary perspective, it seems natural to assume that adaptivity and truth are positively related. True beliefs are what helped our ancestors to survive. In the hunter-gatherer period, having a large number of true beliefs concerning edible and poisonous plants and local animal behaviour was crucial to people's lives and this knowledge was shared in the group and passed on to the next generation. The same holds true for us. It is certainly better for me to believe that the mushroom with an olive-green or yellowish cap and white gills under the cap, white annulus, and white volva at the base, known as *Amanita phalloides* or the death cap, is deadly poisonous, than for me to think it an exquisite delicacy.

---

1 This work was supported by the Slovak Research and Development Agency under Contract No. APVV-18-0178.

The contention of a positive link between truth and adaptivity lies at the heart of the philosophical position called “evolutionary reliabilism”. Plantinga’s attack on this principle was a key step in his journey to rejecting naturalism in its entirety. The Darwinian concept of evolution has become a strong explanatory tool in science; Dennett defines Darwin’s contribution to philosophy thus: “In a single stroke Darwin’s theory of evolution by natural selection united the realm of physics and mechanism on the one hand with the realm of meaning and purpose on the other.”<sup>2</sup> Plantinga, however, is prepared to argue that although evolution is *the main pillar of contemporary science*, it is by no means the pillar of the naturalistic worldview.<sup>3</sup> In his famous evolutionary argument against naturalism (EAAN)<sup>4</sup> Plantinga attempts to demonstrate that to combine evolutionary theory with naturalism is self-referentially incoherent, and so these doctrines cannot rationally be accepted together. The main line of argumentation goes roughly like this: supposing our cognitive faculties evolved through the process of undirected evolution, then the probability of these faculties being reliable is low. Hence, if we accept evolutionary theory and naturalism, we have reason to doubt the reliability of our cognitive faculties. And if the reliability of our cognitive faculties is suspect, the same applies to what they deliver. But naturalism and evolutionary theory are themselves produced by our cognitive faculties. Therefore, “my belief that naturalism and evolution are true gives me a defeater for that very belief.”<sup>5</sup>

The EAAN has been the subject of extensive discussion, but Plantinga, having responded to dozens of objections raised by well-known scholars, considers that the “EAAN seems to me to emerge unscathed – or if a bit scathed, then at least bloody but unbowed.”<sup>6</sup> At the same time, the EAAN does not

---

2 Dennett, D., Darwin’s “Strange Inversion of Reasoning”. In: Avise, J. C. – Ayala, F. J. (eds.), *In the Light of Evolution: Volume III: Two Centuries of Darwin*. Washington, The National Academies Press 2009 [accessed on: 17. 2. 2021]. Available at: <http://www.nap.edu/catalog/12692.html>, pp. 343–344.

3 Plantinga, A., *Where the Conflict Really Lies: Science, Religion, and Naturalism*. Oxford, Oxford University Press 2011, p. 310 (hereafter *Where the Conflict Really Lies*).

4 The first formulation of this argument can be found in Plantinga’s book *Warrant and Proper Function* (New York, Oxford University Press 1993; hereafter *Warrant and Proper Function*). Since then, the EAAN has been restated by Plantinga many times and has been the subject of extensive argumentational exchanges between Plantinga and other scholars, and so naturally it has undergone some refinement, but not a substantial overhaul, as I see it. Plantinga himself refers in his later works to the earlier wordings without qualification (see Plantinga, A., Content and Natural Selection. *Philosophy and Phenomenological Research*, 83, 2011, No. 2, p. 435 [hereafter Content and Natural Selection]).

5 Plantinga, A., *Where the Conflict Really Lies*, p. 314.

6 Plantinga, A., Reply to Beilby’s Cohorts (hereafter Reply to Beilby’s Cohorts). In: Beilby, J. K. (ed.), *Naturalism Defeated? Essays on Plantinga’s Evolutionary Argument Against Naturalism*. Ithaca, Cornell University Press 2002, p. 205 (hereafter *Naturalism Defeated?*).

seem to have converted any naturalists (and Plantinga probably did not expect it to). In this paper I do not intend to examine the whole argument, and I focus only on the first premise and the reasons put forward to substantiate it. I shall argue that, despite all Plantinga's thought experiments offering justification, his general claims suffer from a common deficiency – they disregard *evolution* or fail to take it into account properly. Moreover, I argue that merely pointing out the difficulties that naturalistic approaches have in explaining mental causation does not lead to the conclusion that *in a naturalistic world, there can be no mental causation whatsoever*.

### Darwin's Doubt

The first premise of the EAAN says that the conditional probability (P) that our cognitive faculties are reliable (R), given evolution (E) and naturalism (N), is low:  $P(R/N\&E)$  is low.<sup>7</sup> For Plantinga *naturalism* is “the belief that there aren't any supernatural beings – no such person as God, for example, but also no other supernatural entities, and nothing at all like God.”<sup>8</sup> *Evolution* is an abbreviation for the claim that “we and our cognitive faculties have come to be in the way proposed by the contemporary scientific theory of evolution.”<sup>9</sup> A *reliable* cognitive faculty “must deliver at least 3 times as many true beliefs as false: the proportion of true beliefs in its output is at least three quarters.”<sup>10</sup> Finally, asserting *the conditional probability* of a proposition means considering its probability to be true under some specified circumstances: “the conditional probability of one proposition p on another proposition q is the probability that p is true *given that*, on the condition that, q is true.”<sup>11</sup>

Plantinga also refers to this questioning of whether cognitive faculties developed over the course of “blind” evolution are in fact reliable as “Darwin's doubt”, taking his inspiration from the following sentence in Darwin's letter to William Graham:<sup>12</sup> “With me, the horrid doubt always arises whether the

---

7 Plantinga, A., *Where the Conflict Really Lies*, p. 317.

8 Plantinga, A., Introduction: The Evolutionary Argument Against Naturalism: An Initial Statement of The Argument. In: Beilby, J. K. (ed.), *Naturalism Defeated? Essays on Plantinga's Evolutionary Argument Against Naturalism*. Ithaca, Cornell University Press 2002, p. 3 (hereafter Introduction: The Evolutionary Argument Against Naturalism).

9 Plantinga, A., *Where the Conflict Really Lies*, p. 317.

10 *Ibid.*, p. 332.

11 *Ibid.*, p. 317, emphasis in original.

12 I believe that Darwin's doubt was in fact about something else – it is apparent from the context of the letter that he was concerned with “intuitive” metaphysical beliefs rather than cognitive faculties. In the absence of God, who could have infused our mind with metaphysical truths, our metaphysical intuitions about the world are in themselves – without any further critical ex-

convictions of man's mind, which has been developed from the mind of the lower animals, are of any value or at all trustworthy. Would any one trust in the convictions of a monkey's mind, if there are any convictions in such a mind?"<sup>13</sup>

Plantinga approaches the problem as follows. From the perspective of the Christian religion, our cognitive faculties are reliable because they have been created by God as such. In other words, we were created in the image of God, as creatures able to acquire knowledge.<sup>14</sup> But, "if our cognitive faculties have originated as Dawkins thinks, then their ultimate purpose or function (if they *have* a purpose or function) will be something like *survival*."<sup>15</sup> Natural selection operates directly upon behaviour only, not *beliefs*. If beliefs are of any interest to natural selection, then it is solely because of the relationship between beliefs and behaviour. Certain cognitive mechanisms could be selected during the evolutionary process only insofar as they have some effect on behaviour. And so, when Plantinga states that the probability that cognitive faculties are reliable if naturalism and evolutionary theory is true is low, he means that "our having reliable faculties *isn't* guaranteed by or even particularly probable with respect to adaptive behaviour."<sup>16</sup> Of course, this claim needs further argument since it goes against what is usually believed, as noted above.

According to Plantinga, the overall probability (R/N&E) should be calculated as "the weighted average of the probabilities of R on N&E&C and N&E&-C (weighted by the probabilities of C and -C on N&E)",<sup>17</sup> where C is the proposition that the content of beliefs is causally efficacious and -C represents the denial of the causal efficacy of the content of beliefs. Following this distinction, two main lines of reasoning can be identified. The first line is aimed at demonstrating that there is a causal gap between *the adaptivity* of beliefs and their *truthfulness*. In the second line of argumentation, Plantinga asserts that the *content* of beliefs (and therefore also their truth value) is completely out of reach of natural selection. I will now treat these lines of reasoning separately in the following sections.

---

amination – dubious. Cf. Darwin, C. R., To William Graham, July 3<sup>rd</sup> 1881. In: *Darwin Correspondence Project*, "Letter no. 13230" [accessed on: 19. 2. 2021]. Available at: <https://www.darwinproject.ac.uk/letter/DCP-LETT-13230.xml>.

13 Darwin, C. R., To William Graham, July 3<sup>rd</sup> 1881. Plantinga cites this passage almost every time he brings up the EAAN.

14 Plantinga, A., Introduction: The Evolutionary Argument Against Naturalism, p. 2.

15 Plantinga, A., *Warrant and Proper Function*, p. 218, emphasis in original.

16 Plantinga, A., Introduction: The Evolutionary Argument Against Naturalism, p. 5, emphasis in original.

17 *Ibid.*, p. 10.

## Adaptive False Beliefs

Plantinga's first observation with regard to natural selection is that it leads organisms to *adaptive behaviour*, i.e. to improve their *fitness*, which is "a measure of the chances that one's genes are widely represented in the next and subsequent generations."<sup>18</sup> From this, it is clear that natural selection does not aim at *truth*, at least *not directly*. Still, we may think the truthfulness of beliefs is the usual reason why actions are adaptive. To this Plantinga remarks: "Our having evolved and survived makes it likely that our cognitive faculties are reliable and our beliefs are for the most part true, only if it would be impossible or unlikely that creatures more or less like us should behave in fitness enhancing ways but nonetheless hold mostly false beliefs."<sup>19</sup> This assertion is followed by several thought experiments aimed at demonstrating that in fact it is possible for there to be creatures similar to us whose behaviour is adaptive and most of whose beliefs are false.

Before introducing and analysing the thought experiments proposed by Plantinga, it is important to note that they could not be about *us*. Similarly, in developing a critique of Plantinga's position and formulating counterexamples we cannot draw on *our own experience*. The reason is simple: as suggested above, Plantinga thinks *our cognitive faculties* are quite reliable but at the same time, he does not think that we have developed through a process of unguided evolution. Therefore, none of the situations that we have experienced, or the scientific observations or theories relating to humans that could serve as confirmation of the close relationship between the adaptivity of our behaviour and the truth of our beliefs, can be used as counterexamples. Although some scholars do not in fact accept the restriction,<sup>20</sup> as a *reply to Plantinga's argument* I consider this methodologically incorrect. The EAAN as a whole is designed to "undercut" naturalism and therefore to reject the idea that we have developed through "blind" evolution. But if we take our own experience or data obtained through human research as evidence of the fact that human cognitive faculties are reliable under naturalism, then the assumption is that naturalism holds for us.

Let us take a look at Plantinga's thought experiments that are supposed to demonstrate the possibility of unreliable, yet adaptive cognitive faculties (or false but fitness-enhancing beliefs) on N&E&C.

---

18 Plantinga, A., *Naturalism Defeated?*, p. 4.

19 Plantinga, A., Introduction: The Evolutionary Argument Against Naturalism, p. 5.

20 See e.g. Dennett, D., Darwin's "Strange Inversion of Reasoning", pp. 347–353; and Fales, E., Darwin's Doubt, Calvin's Calvary. In: Beilby, J. K. (ed.), *Naturalism Defeated? Essays on Plantinga's Evolutionary Argument Against Naturalism*. Ithaca, Cornell University Press 2002, pp. 48–49.

A. The source of the first set of examples is a fact Plantinga puts forward – that behaviour is caused not only by beliefs, but also by other factors, such as *desire, suspicion, doubt, approval and disapproval, fear*.<sup>21</sup> To demonstrate how this supports his case, Plantinga chooses examples where combinations of desires and false beliefs lead to adaptive behaviour. He introduces Paul – a prehistoric hominid, and the possible ways that he could avoid being eaten by a tiger (i.e. act adaptively) on the basis of false beliefs (in combination with desires). I shall mention just the first two of them since all the examples are based on the same principle.

1. Paul “likes the idea of being eaten, but when he sees a tiger, always runs off looking for a better prospect, because he thinks it unlikely that the tiger he sees will eat him.”<sup>22</sup>
2. Paul thinks “the tiger is a large, friendly, cuddly pussycat and wants to pet it; but he also believes that the best way to pet it is to run away from it.”<sup>23</sup>

This course of reasoning falls well short of establishing the desired thesis and contains several flaws. First, all the examples come from the realm of the conceivable. But if we are to consider the conditional probability of the reliability of cognitive faculties and the conditions are *evolution* and *naturalism*, then the examples have to stem not from what is *conceivable*, but from what *could have originated in the process of evolution*. It is highly improbable that Paul does not have sufficient information regarding tigers’ behaviour if tigers are part of the natural environment of Paul and his community. Although Plantinga makes the probability of R conditional on naturalism and evolution, he does not include their principles in his thought experiments. I consider this to be a general error in all of his examples. Furthermore, since those beliefs are not necessarily connected to the accompanying desires, they are adaptive only by chance in particular situations. As soon as Paul loses his interest in petting the “pussycat”, or postpones the fulfilment of his desire to be eaten, or simply changes his mind, his beliefs will prove non-adaptive. But a belief can be labeled adaptive only if it can be repeatedly proved. What is more, behind the adaptive but false beliefs one has to identify a general belief-forming mechanism. When Dennett and McKay in their study “The Evolution of Misbelief”<sup>24</sup> look for beliefs that are adaptive but in-

---

21 Plantinga, A., Introduction: The Evolutionary Argument Against Naturalism, p. 8.

22 Plantinga, A., *Warrant and Proper Function*, p. 225, emphasis in original.

23 Ibid.

24 McKay, R. T. – Dennett, D. C., The Evolution of Misbelief. *Behavioral and Brain Sciences*, 32, 2009, No. 6, pp. 493–561 (hereafter *The Evolution of Misbelief*).

correct, they are attempting to find incorrect *systematically adaptive* beliefs that are part of the design of the cognitive system. Also as Ramsey points out, *no one denies the possible existence of false beliefs which, “coupled with certain desires, would produce adaptive behaviour in certain conditions.”*<sup>25</sup> He sets out three conditions that such a cognitive mechanism would have to meet: 1. it could come about through evolution, 2. it generates mostly false beliefs and 3. it proves adaptive.<sup>26</sup> Plantinga’s first argument does not fulfil any of these; in the first place because he does not describe a general cognitive mechanism.

In the second group of examples, Plantinga reflects on the need for a more systematic account. He suggests that we should imagine a general mechanism for the formation of false beliefs – members of a hypothetical population refer to things only by using *definite descriptions*, which are false, i.e. not satisfied by anything.<sup>27</sup> Paul thinks everything is conscious and refers to everything using the description “*That so-and-so conscious being,*”<sup>28</sup> or Paul believes that “all the plants and animals in his vicinity are witches, and his ways of referring to them all involve definite descriptions entailing witchhood.”<sup>29</sup> Finally, the third example is that of “a tribe of cognitively gifted creatures believing that everything (except God Himself) has been created by God,”<sup>30</sup> and we are encouraged to imagine that these beliefs are false, i.e. naturalism holds for them. As Plantinga specifies further, “all their beliefs are properly expressed by singular sentences whose subjects are definite descriptions expressing properties that entail the property of creaturehood,”<sup>31</sup> where “creaturehood” means “having been created by God”. Importantly, Plantinga proposes that “their definite descriptions work the way Bertrand Russell thought definite descriptions work.”<sup>32</sup>

What is wrong with these examples? Although at first sight they seem to be a shade more convincing than the first set, I shall try to show that they suffer from similar maladies. Plantinga disregards the condition of evolu-

---

25 Ramsey, W., *Naturalism Defended*. In: Beilby, J. K. (ed.), *Naturalism Defeated? Essays on Plantinga’s Evolutionary Argument Against Naturalism*. Ithaca, Cornell University Press 2002, p. 20 (hereafter *Naturalism Defended*).

26 *Ibid.*

27 Plantinga, A., *Introduction: The Evolutionary Argument Against Naturalism*, p. 9.

28 *Ibid.*, emphasis in original.

29 *Ibid.*

30 Plantinga, A., *Reply to Beilby’s Cohorts*, p. 260.

31 *Ibid.*

32 *Ibid.* The third example was introduced in Plantinga’s reply to the objections posed by Ramsey. The last condition mentioned in particular is a direct response to Ramsey’s suggestion that the definite description should be analysed by means of Kripkean causal theory of reference. Ramsey, W., *Naturalism Defended*, pp. 26–27.

tion again, i.e. presents to us a situation that is *conceivable*, but it is dubious whether it is probable from the evolutionary viewpoint. He does not consider the question of whether such a community could have evolved in the evolutionary process, nor does he ask whether their way of thinking about the world would be adaptive. In an argument supporting the probability of such scenario it would be appropriate to submit some partial real examples, e.g. an instance of language in which reference can be made using only definite descriptions, an instance of language that does not contain any general statements, or an instance of language in which all the definite descriptions contain one and the same predicate. Since all the beliefs have to be expressed in *singular sentences containing definite descriptions*, members of the hypothetical tribe would have only very limited knowledge, which is problematic from the viewpoint of adaptivity. Furthermore, it is truly odd that the language would contain demonstrative pronouns, as is evident from the form of the definite descriptions Plantinga sets out (see above), but that those pronouns could not be used independently; that is, they could say “that witch is blooming” but not “that is blooming”. Moreover, members of the hypothetical tribe would not be able to reflect on the accuracy of the predicate contained in the definite descriptions. Not only would it be impossible to deny these properties and say that something is *not a witch*, or that it is *not a creature*, or that it is *not conscious* – because that would lead to a contradiction – but if a given description is the only medium of reference to a particular thing, it is not even possible to consider whether the description applies to the thing or not.

Although I find Plantinga’s example of a cognitive system that is adaptive yet unreliable to be mistaken, it does not automatically mean that the whole thesis, Darwin’s doubt, is under serious threat. As Plantinga remarks,<sup>33</sup> even the concession that  $P(R/N\&E\&C)$  is high does not by itself disprove his claim that  $P(R/N\&E)$  is low. Let us recall that Plantinga determines the probability of  $R/N\&E$  as “the weighted average of the probabilities of  $R$  on  $N\&E\&C$  and  $N\&E\&-C$  (weighted by the probabilities of  $C$  and  $-C$  on  $N\&E$ )”<sup>34</sup> Therefore, to dismiss the thesis that  $P(R/N\&E)$  is low, it is not enough to undermine the assertion that  $P(R/N\&E\&C)$  is low. Besides, it should be demonstrated that  $P(C/N\&E)$  is high, or that  $P(-C/N\&E)$  is low. However, Plantinga is convinced that  $P(-C/N\&E)$  is high, because, as he puts it, “it is extremely hard, given materialism, to envisage a way in which the content of a belief could get causally involved in behaviour.”<sup>35</sup>

33 Plantinga, A., Introduction: The Evolutionary Argument Against Naturalism, pp. 9–10.

34 See footnote 17 and the corresponding place in the text.

35 Plantinga, A., Introduction: The Evolutionary Argument Against Naturalism, p. 10.



## Causal Gap between Material and Mental

Plantinga considers the problem of the (in)ability of materialism to account for the causal impact of belief contents on behaviour (and for the causal laws between the neurophysiological and the mental) in more detail especially in his later works.<sup>36</sup> He either *thinks of naturalism as including materialism*<sup>37</sup> or alternatively ponders on materialism in reaction to *the conditionalisation problem*<sup>38</sup> – but for our present purposes the difference is not significant. Plantinga considers materialism in two forms, as *reductive* and *non-reductive* materialism. Let us begin with the latter.

Plantinga defines non-reductive materialism (NRM) as the theory that content properties supervene on neurophysiological (NP) properties: “for any content property C that a neural structure can have, there is an NP property P such that if a neural structure has the content property C, it has P, and conversely, any neural structure that has P also has that content property C.”<sup>39</sup> NRM is not considered here as a complex theory; rather, Plantinga investigates whether the sole fact of the supervenience of the content on NP properties could ensure its causal impact on behaviour and again supports his line of reasoning with a thought experiment. We are invited to imagine a hypothetical species cognitively similar to us – they have beliefs and change them, make inferences etc. They live in a world where no God exists, i.e. naturalism holds for them. Their beliefs are neurological structures, complex enough to generate content, which at the same time serve as reliable indicators – a concrete “structure is a reliable indicator of that kind of predator: it arises when and only when there is a such a predator in the middle distance.”<sup>40</sup> Yet we have no reason to expect that the proposition determined by the NP structure is true, states Plantinga.<sup>41</sup> Certain NP properties determine the content

---

36 E.g. in the last part of his book *Where the Conflict Really Lies*, pp. 318–339, in the paper *Content and Natural Selection*, pp. 437–445, or in the response to Paul Draper *Against “Sensible” Naturalism* (Plantinga, A., *Against “Sensible” Naturalism*. 2007 [accessed on: 21. 2. 2021]. Available at: [https://infidels.org/library/modern/alvin\\_plantinga/against-naturalism.html](https://infidels.org/library/modern/alvin_plantinga/against-naturalism.html)).

37 Plantinga, A., *Where the Conflict Really Lies*, p. 326.

38 Plantinga, A., *Content and Natural Selection*, p. 439. The conditionalisation problem was originally formulated by Richard Otte in his critical study *Conditional Probabilities in Plantinga’s Argument*. Otte states here that even if  $P(R/N\&E)$  were low, we could have some further evidence (O) such that if we add it to the conditions, then  $P(R/N\&E\&O)$  would be high. Otte, R., *Conditional Probabilities in Plantinga’s Argument*. In: Beilby, J. K. (ed.), *Naturalism Defeated? Essays on Plantinga’s Evolutionary Argument Against Naturalism*. Ithaca, Cornell University Press 2002, p. 137.

39 Plantinga, A., *Where the Conflict Really Lies*, p. 324.

40 *Ibid.*, p. 330.

41 *Ibid.*, pp. 330–331.

of beliefs and they also cause adaptive behaviour. But according to Plantinga, supervenience does not ensure that the associated content of adaptive NP properties is true or that it has any impact on the behaviour: “whether or not that content is *true* makes no difference to fitness.”<sup>42</sup> He argues that on NRM it is possible that a person from the naturalistic world could avoid stepping into a bathtub with an alligator (i.e. act adaptively) while believing that “the alligator is a mermaid, or even that he’s sitting under a tree eating mangoes.”<sup>43</sup> All the person needs is “indicators and other neural structures that send the right messages to his muscles.”<sup>44</sup> According to the proposed scenario NRM obviously collapses into semantic epiphenomenalism. Natural selection shapes the NP properties to be adaptive, i.e. to cause adaptive behaviour, but the associated content may be false, or may be true; it does not really matter.

As before, I consider this argumentation to be flawed, since the conditions of *evolution*, i.e. *natural selection* and *naturalism*, have not been considered *properly*. Yes, they have been taken into account. As Plantinga clarifies, his argument does not necessarily disqualify *materialism* – since there are also theists who are materialists and could believe that, since God has created us as knowers, he has established “psychophysical laws of such a sort that successful action is correlated with true belief.”<sup>45</sup> The argument should therefore be effective particularly against those who endorse the naturalistic view of evolution. But from the perspective of evolutionary naturalism, how can one explain the *supervenience* of the content “*I am sitting under a tree eating mangoes*”<sup>46</sup> on NP properties that caused the person to avoid stepping into a bathtub? And if we are endowed with adaptive indicators and adaptive NP properties of beliefs, how can we account for the emergence of mental content at all, with no obviously adaptive function?

Plantinga’s own solution to the problem of causality between mental and material is reminiscent of Descartes: the “causal laws linking NP properties with content properties in such a way that the beliefs in question would be... mostly true” have been *instituted by God*.<sup>47</sup> Indeed, the example described above with the alligator and mangoes recalls the Cartesian evil demon who randomly (or intentionally wrongly) attaches ideas to neurophysiological states. But in naturalism there is no place for this evil demon. Thus it is exceedingly difficult to see how one could reasonably explain, from the point of

---

42 Ibid., p. 327, emphasis in original.

43 Plantinga, A., *Against “Sensible” Naturalism*.

44 Ibid.

45 Plantinga, A., *Where the Conflict Really Lies*, p. 339, footnote 29.

46 See footnote 43.

47 Plantinga, A., *Against “Sensible” Naturalism*.

view of evolutionary naturalism, how a situation in which there is a bathtub and an alligator (but no tree or mango nearby, and nor is the subject eating anything) could give rise to the mental content “I am sitting under a tree eating mangoes”. Even if Plantinga is right that NRM allows for such a situation (because of the lack of an adequate theory of mental causation), there are still two more conditions – naturalism and evolution – which disqualify such cases. And this is what Plantinga has overlooked.

Before the final conclusion, let us take a brief look at how Plantinga treats RM in terms of its ability to establish the causal impact of mental contents on behaviour. In reductive materialism, as the name suggests, content properties are reducible to NP properties.<sup>48</sup> Plantinga follows a similar path as he does with NRM and once again ends up with semantic epiphenomenalism.<sup>49</sup> He points out that among the NP properties of a belief is “the property of having such and such a proposition as its content,”<sup>50</sup> but the fact that the NP properties of a belief are adaptive gives us no reason to assume that the associated content is true: “the content doesn’t have to be true, of course, for the neuronal structure to cause the appropriate kind of behaviour.”<sup>51</sup>

As with NRM, Plantinga indicates the failure of RM to account for the causal relationship between NP properties and mental content, and his examples cannot be disregarded just because they contradict our experience (for Plantinga *our* cognitive faculties, as designed by God, are reliable). However, since the first premise of the EAAN – “Darwin’s doubt” – concerns the low probability of the cognitive faculties being reliable under the conditions of *naturalism* and *evolution*, (and materialism is conceived either as the only admissible theory of mind for naturalism, or as another condition), Plantinga’s argumentation should have contained the principles of evolutionary theory as well. All the examples should reflect not only what would be acceptable on materialistic principles,<sup>52</sup> but also on evolutionary principles – which they do not. Moreover, it seems that the conditionalisation problem

---

48 Plantinga, A., *Where the Conflict Really Lies*, p. 323.

49 Boudry and Vlerick call the view Plantinga ascribes here to naturalists “*arbitrary content labeling*”, arguing it is even stronger and stranger than semantic epiphenomenalism. Boudry, M. – Vlerick, M., Natural Selection Does Care about Truth. *International Studies in the Philosophy of Science*, 28, 2014, No. 1, p. 70.

50 Plantinga, A., *Where the Conflict Really Lies*, p. 334.

51 *Ibid.*

52 There have also been objections to Plantinga’s way of reasoning based on the fact that he reduces materialism to a single thesis, instead of *taking reductive materialism seriously and assuming “the full strength of reductive materialism”*. Ye, F., Naturalized truth and Plantinga’s evolutionary argument against naturalism. *International Journal for Philosophy of Religion*, 70, 2011, No. 1, p. 33 (hereafter Naturalized truth and Plantinga’s evolutionary argument against naturalism).

(mentioned earlier)<sup>53</sup> might well apply here. The fact that a general theory on the relationship between NP and mental features of beliefs allows for semantic epiphenomenalism does not mean that materialistic theories of mental content do not contain other important features capable of preventing it.

Now let us take a slightly different perspective and accept that Plantinga has highlighted an important difficulty that materialism has in accounting for mental causation.<sup>54</sup> Would that be enough to substantiate the claim that in a naturalistic world there would be no mental causation? First, it is important to determine which thesis Plantinga has proved: that *no materialistic theory could possibly explain mental causation* or that *neither the postulate of supervenience of mental properties on NP properties of beliefs (alternatively the postulate of the identity of mental and NP properties) and the most popular current naturalistic theories of content*<sup>55</sup> are able to explain mental causation properly. Although I presume that Plantinga believes that materialism is incapable of explaining mental causation – since he himself relies here on supernatural explanation<sup>56</sup> – his arguments support only the latter thesis.

Second: is the fact that *the naturalistic theories are currently not capable of ascertaining causality between the mental and NP properties of beliefs* sufficient reason to claim that *in a world that can be truthfully described by naturalism and contemporary evolutionary theory* (some version of it) *the probability that mental content has a causal effect on behaviour is low*? I believe not, obviously. It has not been demonstrated that RM or NRM lead *inevitably* to semantic epiphenomenalism. As mentioned earlier, Plantinga did not consider the full-blown theories which may contain other constraints on the relationship between mental and NP properties of beliefs, and therefore he has at best proved that RM and NRM *allow for* semantic epiphenomenalism.<sup>57</sup> If, then, semantic epiphenomenalism is not a necessary companion of materialism, or of naturalism, the current lack of a proper naturalistic account of mental causation gives us no reason to suppose that there would be no

---

53 See footnote 38.

54 However, as Ye points out, the problem of explaining mental causation could be posited as a separate argument against materialism. But “the concern here is merely whether his evolutionary argument contains any new challenge against materialism”. Ye, F., Naturalized truth and Plantinga’s evolutionary argument against naturalism, p. 33, emphasis in original.

55 In his reactions to the conditionalisation problem, Plantinga considers *the indicator semantics, functionalism, and teleosemantics* as well as RM and NRM. Plantinga, A., Content and Natural Selection, pp. 445–458.

56 See above, footnote 47 and the corresponding place in the text.

57 Cf. Novotný, D., How to Save Naturalism from Plantinga? *Organon F*, 14, 2007, No. 1, p. 38: “... a naturalistic account of mental causation has not been shown to be impossible but (at best) currently nonexistent.”

mental causation in a naturalistic world. Here I shall try to make a tentative distinction between an *epistemological* and *metaphysical* reading of the argument, presenting it using the example of the EAAN. On a metaphysical reading it says: if contemporary evolutionary theory and naturalism are true, the probability that our cognitive faculties are reliable is low. On an epistemological reading: if we are convinced of the truth of contemporary evolutionary theory and naturalism, we have little reason to believe that our cognitive faculties are reliable. In this case, the epistemological reading is correct. With the first premise of the argument, “Darwin’s doubt”, the opposite is the case. The intention is not to make an epistemological claim: if you believe that naturalism and contemporary evolutionary theory are true, then you should admit that the probability of R is low. Rather, the concern is “metaphysical”: if naturalism and evolutionary theory are true, the probability<sup>58</sup> of R is low. But his arguments do not substantiate this thesis. Even if contemporary naturalistic theories do not establish mental causation, it does not follow that in a naturalistic world there would be no mental causation. More generally, if we are currently unable to explain a phenomenon, that does not mean it is inexplicable in principle or does not exist.

On the basis of what has been said, I conclude that Plantinga has not provided us with good reasons for accepting the thesis that  $P(R/N\&E)$  is low. All the same, nothing has been said that justifies the claim that the probability is high. Naturalists agree on the fact that natural selection truly “cares” about adaptivity in the first place – sometimes at the expense of truth.<sup>59</sup> Although the question is beyond the scope of this paper, I am convinced that in developing an answer, cultural evolution should also be taken into account. As McKay and Dennett put it, “cultural evolution can have played the same shaping and pruning role as genetic evolution.”<sup>60</sup>

<https://orcid.org/0000-0002-7554-4606>

---

58 The probability here denotes essential randomness; when the belief content is causally inert, it is truly random whether adaptive NP properties would carry true or false content.

59 E.g. a group of adaptive misbeliefs identified by Dennett and McKay comprising some kinds of positive illusion. See McKay, R. T. – Dennett, D. C., *The Evolution of Misbelief*, pp. 505–509.

60 *Ibid.*, p. 508.