ARTÍCULOS

Logos. Anales del **Seminario de Metafísica**
ISSN: 1575-6866

EDICIONES COMPLUTENSE

https://dx.doi.org/10.5209/asem.88829

EVAAN: An empirical verification argument against naturalism

Ward Blondé[1]

Recibido: 22 de mayo de 2023 / Aceptado: 17 de noviembre de 2023

Abstract. Alvin Plantinga's evolutionary argument against naturalism (EAAN) claims that if both naturalism (N) and evolutionary theory (E) are true, then all our beliefs are unreliable (premiss 1). Consequently, given N&E, the belief in N&E is unreliable (premiss 2) and N&E is self-defeating (conclusion). My empirical verification argument against naturalism (EVAAN) is more cautious and improves Plantinga's EAAN by withstanding a rejoinder of the evolutionary naturalist to premiss 1. EVAAN claims that metaphysical beliefs are unreliable, given N&E (premiss 1a). This anticipates the evolutionary naturalist's claim that empirical verification plays a crucial role in peer assessments and sexual selection of intelligence, and that, therefore, N&E makes empirically verifiable beliefs often reliable (premiss-1 rejoinder). However, even then it can be argued that the belief in N&E is unreliable, given N&E (premiss 2), because N&E is metaphysical (premiss 1b) and therefore not empirically verifiable. My EVAAN distinguishes reliably verifiable intelligence from metaphysical intelligence and claims that, if N&E is true, humans are lacking metaphysical intelligence. This paper also contains an argument against Plantinga's EAAN, by supporting the premiss-1 rejoinder.
Keywords: Evolutionary argument against naturalism, Alvin Plantinga, Evolutionary reliability, Evolution by natural selection, Epistemology, Supernaturalism, Intelligence, Just-so stories.

How to cite: Blondé, W. (2023) "EVAAN: An empirical verification argument against naturalism", en *Logos. Anales del Seminario de Metafísica* 56 (2), 345-362.

[1] Investigador independiente
wardblonde@gmail.com

1. From EAAN to EVAAN

This paper provides an argument that the conjunction of naturalism (N) and evolutionary theory (E) is self-defeating. Here, I use the following definitions: Natural entities are those that can be revealed as explanatory by the methodologies of the empirical sciences (Forrest 2000). N is metaphysical[2] naturalism: all and only the natural entities exist in reality (Mahner 2012). Furthermore, metaphysical beliefs are beliefs about the existence or non-existence of certain classes of logically possible, non-natural (or supernatural) entities.

Blondé's (2015, 2016, 2019) and Blondé and Jansen's (2021) accounts of supernaturalism will be taken as leading examples.[3] Their supernaturalists can make the claim that naturalists are supernaturally predestined to believe in N. In response to this claim, naturalists should be able to prove N. The belief in N often goes together with the belief in E. As a part of the main argument in this paper, I show that, if N&E is true, evolutionary naturalists are empiricists about the ontology of reality by their own nature: methodologies that prove premisses that are not empirically verifiable are rejected. All of this would not be a problem if N&E itself was an empirically verifiable premiss. However, this is not the case.

The empirical verification argument against naturalism (EVAAN) in this paper is based on Plantinga's (1993, 2011) evolutionary argument against naturalism (EAAN). EAAN's essence can be reconstructed as follows:

Premiss 1: If N&E is true, then all our beliefs are unreliable.[4]
Premiss 2: If N&E is true, the belief in N&E is unreliable.
Conclusion: N&E[5] is self-defeating.[6]

Premiss 1 of Plantinga's EAAN is a strong claim for many, because it rejects the reliability of all beliefs, if N&E is true. In this way, EAAN is preaching mainly to those who are already converted to the falsehood of N&E. In particular, it does not appeal to those who hold the following two beliefs: N&E is true and some of our beliefs are reliable. With the following variation of Plantinga's EAAN, my EVAAN attempts to get more people on board in the argument against N(&E):

Premiss 1a: If N&E is true, metaphysical beliefs are unreliable.
Premiss 1b: N&E is a metaphysical belief.
Premiss 2: If N&E is true, the belief in N&E is unreliable.
Conclusion: N&E is self-defeating.

---

2 While metaphysical naturalism is about what exists, methodological naturalism is a practice in empirical research.

3 See Section 4.1 for their exposition.

4 Plantinga's formulation is that, if N&E is true, the probability that our cognitive faculties are reliable is low or inscrutable.

5 It will be shown in Section 5.2 that an argument against N&E is in the first place an argument against N, rather than against E.

6 Plantinga goes even further than this conclusion and claims that N&E is irrational and unacceptable, as a result of it being self-defeating. However, I believe that this requires an additional argument against N&E within the core premisses of Plantinga's EAAN, because some self-defeating theories may be rational and acceptable. For example, John's theory that John is not smart is self-defeating, but may be rational and acceptable for other reasons.

EVAAN's premiss 1a is a consequence of EAAN's premiss 1. Therefore, my EVAAN is strictly more cautious than Plantinga's EAAN. In particular, EVAAN's premiss 1a is resistant to the following rejoinder of the evolutionary naturalist to EAAN's premiss 1:

> Premiss-1 rejoinder: If N&E is true, beliefs that are empirically verifiable, are often reliable.

The damage that this rejoinder can do to Plantinga's EAAN is the rationale behind this paper about EVAAN. Even though an *a priori* proof of the premiss-1 rejoinder requires an evolutionary 'just-so' story (Gould 1978), I will focus on the claim that the premiss-1 rejoinder cannot easily be disproved *a priori*, in contrast to Plantinga's just-so story. This means that this paper will contain two arguments: first, an argument against Plantinga's EAAN that seems to be lacking in the current literature, and second, my EVAAN argument that replaces EAAN.

After a section with background (Section 2), three sections follow that analyze and support the premiss-1 rejoinder (Section 3: against Plantinga), premiss 1a of my EVAAN (Section 4), and the rest of EVAAN's premisses (Section 5), respectively. After that come a discussion (Section 6) and conclusions (Section 7).

2. Background

In this paper, evolutionary reliability is the claim that human beings develop reliable beliefs if N&E is true. E, in this case, does not require any supernatural interventions in order to create reliable beliefs. Plantinga argues against this kind of evolutionary reliability (or at least, against any variant of it in which N must be true) in his premiss 1, and so do Stich (1990), Sage (2004), and Slagle (2021). Their arguments essentially distinguish truth-reliability from fitness reliability (Sage 2004, p. 95). In order to be evolutionarily adaptive (fitness enhancing), organisms only have an advantage from fitness reliable beliefs causing adaptive behavior, what we could call *survival of the fittest behavior*. For example, if N&E is true, a being could have the belief that seeing a tiger signals the start of a running competition (Plantinga 1993, p. 225), because this belief, even though not truth reliable, enhances its behavioral fitness. Another example is that of a cautious cognitive faculty that often creates the false belief that a predator is nearby (Sage 2004, p. 97). Again, this cognitive faculty can be fitness reliable, without being truth reliable, because it results in adaptive behavior. It can, therefore, be argued that, if N&E is true, our beliefs are fitness-reliable, however, our epistemic goals require them to be truth reliable.

Many have argued in favor of evolutionary reliability (O'Connor 1994, Fales 1996, Beilby 2002, Boudry and Vlerick 2014, Deem 2018). W.V. Quine made the famous quote that 'creatures inveterately wrong in their inductions have a pathetic but praiseworthy tendency to die before reproducing their kind' (O'Connor 1994, p. 528). O'Connor's (p. 531) main objection to Plantinga's EAAN is that it disconfirms any competing account of our origins, including theism (T). In reaching this conclusion, O'Connor uses Plantinga's claim that P(R), the probability that our cognitive faculties are reliable, is 1 or very near to 1, while P(R/T), this probability given theism, must, according to O'Connor, be certainly less than 1.

Fales (1996, p. 432) overturns Plantinga's argument and argues that Plantinga is committed to a Neo-Darwinian account that provides strong reasons for expecting general cognitive reliability. Reliability with respect to theism, on the other hand, would be unknowable. As I will show, a general cognitive reliability does not necessarily suffice to support N&E. It needs to be reliable on the domain of metaphysics.

Boudry and Vlerick (2014, p. 68) argue that 'true beliefs are overall better guides to action in the world than false beliefs,' and that, therefore, 'natural selection will weed out neural structures that give rise to false beliefs.' The modifier 'overall' seems to circumvent Plantinga's and Sage's counterexamples in which false beliefs are equally good guides to the world. On the other hand – and that is the point of this paper – it does not exclude the possibility that certain classes of true beliefs, such as beliefs that are metaphysical but true, are not better guides to action in the world than false beliefs.

Deem points at a flaw in both Stich's and Plantinga's challenges to evolutionary reliability. He argues that the challenges rely on a species-neutral claim about what natural selection favors, in order to derive a species-specific conclusion about the evolution of a particular trait. Deem (2018, p. 222) substitutes 'true beliefs' by the 'showy plumage' of peacocks in the claim of Stich and Plantinga:

> Natural selection does not care about *showy plumage*, but only about adaptive behavior.

However, this does not explain why peacocks tend to have showy plumage.

According to Miller (2000) and Arneth (2009), the human intelligence is driven by sexual selection. Sexual selection traits often serve as fitness indicators (Miller 2000, p. 260): the beauty and attractiveness of prominent, genome-encoded traits, such as the songs of birds, the peacocks' plumage, the antlers of stags, etc., are indicators that also less prominent, genome-encoded traits are healthy, which proves reproductive fitness. Sexual selection traits become particularly important for intelligent beings, because of their superior capabilities to assess such traits. According to Wilson et al. (2017, pp. 1-2), sexual selection explains not just traits of the human body, but also behavior, such as art, music, humor, sports, and, indeed, human intelligence itself. Sexual selection of intelligence is universal in humans: both males and females assess and display intelligence (Miller 2000, p. 260). However, as I will argue in Section 4, the intelligence favored by N&E does not warrant that beliefs are truth-reliable on all epistemic domains.

3. Against Plantinga's EAAN

My EVAAN is a more cautious argument than Plantinga's EAAN. In order to show why this caution is required, I will first strengthen a restricted evolutionary reliability of our cognitive faculties by showing that the premiss-1 rejoinder (If N&E is true, empirically verifiable beliefs are often reliable.) can withstand Plantinga's challenge of evolutionary reliability, if certain intelligence-related evolutionary hypotheses[7]

[7] In this paper, an *evolutionary hypothesis* is a hypothesis about a property (such as intelligence) of an organism that became increasingly important throughout certain stages of the organism's evolutionary history.

are accepted. First of all, we should avoid relying on empirical evidence and show that there are *a priori* reasons, compatible with N&E, that can be given in support of the premiss-1 rejoinder. Otherwise, the supernaturalist can claim that empirically verifiable beliefs are only reliable due to supernatural interventions that have shaped the *a posteriori* evidence. The focus will therefore be to show that the premiss-1 rejoinder cannot easily be disproved *a priori*. If this fails and even empirically verifiable beliefs can be shown to be unreliable if N&E is true, premiss 1 of Plantinga's EAAN and, consequently, premiss 1a of my EVAAN, remain unchallenged.

In what follows, reliably verifiable (RV) intelligence is defined as the state of having beliefs that are often true if they are verifiable in a reliable manner. I contend that RV-intelligence coincides roughly with the classical 'IQ' intelligence but does not warrant 'metaphysical' intelligence. After all, an IQ test typically does not assess reasoning capabilities that are not reliably verifiable, such as about theism or supernaturalism. People with metaphysical intelligence should be able to consider and derive metaphysical beliefs that can be emotionally and mentally challenging to hold, such as being watched and judged continuously, being in a simulation, being the only person with a real consciousness, having an eternal life, or having observed a miracle. Metaphysical intelligence is therefore probably a combination of IQ intelligence and a type of intelligence that is related to emotional intelligence (Łowicki and Zajenkowski 2017).

After introducing four evolutionary hypotheses about RV-intelligence, the preliminary conclusion is established that having RV-intelligence is fitness enhancing for humans, and is therefore favored by E,[8] if N&E is true. Then I will argue that empirically verifiable beliefs can often be reliably verified to be true. From there it can be argued that, if N&E is true, E favors cognitive faculties that often produce true beliefs if the beliefs are empirically verifiable. This shows that EAAN's premiss 1 can be disproved, such that my EVAAN is needed to rescue EAAN's premiss 2 and conclusion.

3.1 Four evolutionary hypotheses about RV-intelligence

If N&E is true, then throughout their evolution, humans could increasingly:

> Hypothesis (1): develop – initially often false – beliefs they can express through propositions in a language.
> Hypothesis (2): analyze whether an RV-proposition is true or false with some – initially low – degree of certainty.

Furthermore, if N&E is true, then throughout human evolution, RV-intelligence has been increasingly assessed – with an initially low rate of success:

> Hypothesis (3): by peers when it comes to granting somebody a role of importance.
> Hypothesis (4): by sexual partners in the context of sexual selection.

---

[8] Fitness-enhancing traits can evolve via adaptation, but also as by-products, or so-called spandrels (Gould and Lewontin 1979), or via constructive neutral evolution (Stoltzfus 1999).

Supporting an evolutionary hypothesis may seem like presupposing what one wants to establish. However, it is about showing that a certain phenomenon becomes *increasingly* important throughout evolution. This is often done by demonstrating the relationship with evolutionary hypotheses about properties that evolve simultaneously. The following sections contain such demonstrations.

3.2 Having RV-intelligence is fitness enhancing for humans

With peer assessment of RV-intelligence I mean the phenomenon that the members of a human social group feel the need to assess the RV-intelligence of other members. While non-human animals may assess each other's strength and size to determine the role of an individual within the group, humans have, according to hypothesis 3, increasingly often opted for RV-intelligence. In the case two members are competing for an important role, such as the leadership of the group, a verification that their RV-beliefs are often *true*[9] may be an important criterion for humans to determine who will become the leader. This is to say that human leaders better have RV-intelligence.

As a result of getting an important role, RV-intelligent group members acquired more power, wealth, and fame for themselves and their families, than humans without RV-intelligence. This in turn provided them and their families with more attractive partners, more offspring, and more resources to raise their offspring.

According to hypothesis 4, sexual selection of RV-intelligence increasingly warrants attractive partners even more directly. Also here, reliable verification is required: the innate conviction of birds that a certain song is beautiful, does not test the reasoning capabilities of other birds. By reasoning about things that can be reliably verified, humans can impress their sexual partners. This shows that having RV-intelligence has been increasingly fitness enhancing for humans.

Admittedly, humans have increasingly assessed many aspects of RV-beliefs during their evolutionary history, apart from the truth of these beliefs. This includes whether beliefs were fitness enhancing or merely useful. Note, however, two things: first, I only need to show here that the truth-values are an important aspect of RV-beliefs for N&E-beings, rather than the only aspect, and second, aspects like evolutionary fitness and usefulness are less reliable as a test criterion, because they make the analysis both more complex and subjective for the assessor.

An analysis about evolutionary fitness can diverge in many directions and does not appear to be reliably determinable in most cases. For example, the belief that swans are blue is reliably verifiable in just one way, while it can be found to be fitness enhancing or fitness diminishing in many ways: the belief might be comical, creative, unique, mysterious, betraying a bad eyesight, and, evidently, betraying a low RV-intelligence. The last way highlights my point that the analysis of evolutionary fitness is strictly more complex than the analysis of truth.

An analysis about usefulness is also subjective and more difficult to make. For example, the belief that two very similar species of birds are in fact the same species may be a useful belief with respect to communication, however, it cannot be reliably determined whether believing in the distinction between the two species is either less or more useful. Here again, truth provides a single, objective criterion.

---

[9] According to Plantinga, a naturalist about the mind cannot use mentalistic language and must, therefore, even abandon the concept of truth.

In conclusion, there are various and complex ways in which the evolutionary fitness and the usefulness of RV-beliefs can be assessed, while there is always a unique and reliably verifiable way to assess their truth-values. This indicates that truth is a unique property of RV-beliefs that can be exploited as a separate but important test in assessing mental fitness. Therefore, in assessing each other's RV-beliefs, humans have increasingly often opted for a reliable verification of their truth-values. All of this holds if N&E is true.

Evolutionary psychologists might object here that this is an *a priori* just-so story, warned for by Gould. However, the aim in this section is to refute Plantinga's *a priori* story, which is less refined.

3.3 The case of empirical verifiability

As said, the RV-intelligence of a human was increasingly assessed by verification that this human had true beliefs. There is one obvious method to verify the truth-value of a belief, namely empirical observation of the world. Therefore, despite the fact that empirical verification may sometimes be fallible, there was a selective pressure on having true beliefs if they were empirically verifiable, because having such beliefs increasingly optimized the number of times a human was assessed as RV-intelligent. Moreover, having such beliefs also increasingly improved the ability to assess the truth-value of the empirically verifiable beliefs of others. As a result, if N&E is true, humans developed cognitive faculties that often produced true beliefs if the beliefs were empirically verifiable. Therefore, if N&E is true, empirically verifiable beliefs are often reliable. This is the premiss-1 rejoinder. Because metaphysical beliefs are not empirically verifiable, the rejoinder is in accordance with premiss 1a of my EVAAN, however, it is against premiss 1 of Plantinga's EAAN. With that, the evolutionary naturalist has demonstrated that a restricted evolutionary reliability cannot easily be disproved *a priori*, starting from four evolutionary hypotheses.

3.4 A discussion of the restricted evolutionary reliability

With the following discussion I hope to corroborate further that a restricted evolutionary reliability of our cognitive faculties cannot as easily be disproved as Plantinga believes. The often counter-intuitive examples that worked against evolutionary reliability can be refuted as a result of a combination of empirical observation and, according to hypothesis 2, a conceptual analysis of the empirical facts that humans are increasingly able to make. For the example that seeing a tiger signals the start of a running competition, humans have increasingly been able to observe and/or analyze that such a running competition is, for example, lacking a finish line and a preparation phase. Therefore, the belief about the running competition has increasingly been verified as false, and humans who communicate about this belief were increasingly regarded as not RV-intelligent. This, in turn, made these humans less evolutionarily adapted. Whereas Plantinga assumes that beliefs, given N&E, were only shaped by the question whether they resulted in fitness-enhancing behavior, it now turns out that it also matters whether they were communicated or the target of interrogations and assessments by other humans.

An objection that could be made is that N&E favors beliefs that merely cohere with the beliefs of others, rather than favoring empirically verifiable beliefs that are

true. The response is that RV-intelligent people can devise empirically verifiable experiments that support their true beliefs. If an RV-intelligent person can convince a whole community of a new belief through an experiment, then that person is regarded as especially RV-intelligent. Moreover, RV-intelligent people can explain their true beliefs in simpler terms to people that are less RV-intelligent.[10] Therefore, developing beliefs that are different from the beliefs of others is evolutionarily beneficial if the beliefs are both empirically verifiable and true.

Let us investigate the objection once more that the four intelligence-related evolutionary hypotheses seem to presuppose what they are trying to establish. Consider for example the process of finding a mate: individuals who were recognized as RV-intelligent are claimed to be favored as mates over individuals who were not. But does this not presuppose that humans are already RV-intelligent prior to the generation of sexual selective pressure on RV-intelligence?

The answer is that the four hypotheses are about four interdependent processes[11] that, like so many evolutionary processes, evolved – and still evolve – gradually together, and with increasing importance. Even for intelligent animals, some prototypes of these processes are already taking place. Therefore, claiming that the intelligence-related evolutionary hypotheses are presupposing what they have to establish is taking a stance against any form of evolution. An example of this would be the claim that 'peacock hens were increasingly attracted by showy plumage at some stage in their evolution' presupposes what it has to establish, namely showy plumage.

Plantinga's argument against evolutionary reliability (premiss 1) may still hold for most non-human animals, because these can barely communicate about their beliefs, and they are typically not assessed by their intelligence. A cow may believe that an electric fence is a poisonous snake without becoming less evolutionarily adapted. A human who has such a belief and mentions it, is judged negatively by his or her peers and sexual partners for this false belief.

Even though it is possible to show that a restricted evolutionary reliability cannot easily be disproved, it cannot be proved, either. Without making use of any *a priori* evolutionary hypotheses, we would need a supercomputer that could simulate the whole evolution of species on Earth, including the currently unsolved problem of the origin of life (abiogenesis), and including, for example, the increased energy consumption of animals with bigger brains. Such a supercomputer cannot be built in the foreseeable future.

Another method, which avoids the supercomputer, would be to use *a posteriori* evidence (from fossils, biology, genomics, etc.) to prove the compatibility of a restricted evolutionary reliability with N&E. However, with this circular method we risk importing evidence that has a supernatural origin, instead of a purely natural one. Therefore, the question is not whether the restricted evolutionary reliability can be proved *a priori*, but whether it can be disproved *a priori*. My EVAAN is more cautious than Plantinga's EAAN in this respect: a restricted evolutionary reliability might not be disprovable in this way.

---

[10] I will argue that, if N&E is true, humans have a hard-wired resistance against the method of explaining in simpler terms when the results are not empirically verifiable.

[11] The related processes of evolving linguistic capabilities (hypothesis 1), analytic capabilities (hypothesis 2), etc.

4. Proving premiss 1a of EVAAN

In this section I show that premiss 1a of my EVAAN (If N&E is true, metaphysical beliefs are unreliable.) is true. I already proposed empirical observation as a method to verify beliefs of other humans. However, some logically verifiable metaphysical beliefs, such as beliefs about theology and cosmological natural selection (Smolin 1992, Blondé 2016, 2019), are not empirically verifiable. By showing that premiss 1a is true, the unrestricted evolutionary reliability (If N&E is true, then our beliefs are reliable on all epistemic domains.) is refuted.

I propose two methodologies that both prove EVAAN's 1a: one via *a posteriori* evidence and one via an *a priori* just-so story. The *a posteriori* methodology consists of observing the peer-reviewed literature to find out whether non-complex metaphysical theories are always easily agreed upon by evolutionary naturalists,[12] or whether some of these theories are controversial among them. The status of these theories holds independent of whether N&E is true or false, and therefore if N&E is true. The *a priori* methodology seeks to find out to what degree true metaphysical beliefs have been fitness enhancing or fitness diminishing for humans throughout their evolutionary history. The *a posteriori* methodology only shows that premiss 1a is apparently true. The *a priori* methodology also provides insight in why 1a is true, if its evolutionary just-so story is deemed plausible.

4.1 *A posteriori* evidence to prove 1a

In this subsection I show that non-complex metaphysical beliefs are often controversial in scholarly debates among contemporary evolutionary naturalists. If these beliefs were reliable and if N&E were true, then evolutionary naturalists should easily find agreement about such beliefs, because they are non-complex. Because this is not the case, premiss 1a, that metaphysical beliefs are unreliable if N&E is true, is apparently true.

Four logically verifiable, non-complex,[13] metaphysical theories will be shown to be controversial among contemporary evolutionary naturalists (independent of N&E being true or not): the exceptional-point-of-view problem, the application of the Sleeping Beauty problem to the anthropic principle, cosmological natural selection, and the multiverse hypothesis. In all these cases, evolutionary naturalists hesitate to accept the conclusions that follow from logical reasoning alone, and they stick by empirical verification of the conclusions.

These four theories are inspired by Blondé's (2015, 2016, 2019) and Blondé and Jansen's (2021) accounts of supernaturalism, which jointly make the following claims:

1. Every logically possible entity exists in reality.
2. Physical brain matter that generates conscious experiences outweighs other types of physical entities with respect to mass and volume.

---

[12] In fact, my EVAAN is already supported by showing that there is disagreement about metaphysical beliefs among metaphysicians. However, a stronger case can be built by replacing my 1a with: "If N&E is true, even non-complex metaphysical beliefs of evolutionary naturalists are unreliable." Indeed, if even they cannot agree about such non-complex problems, then N&E is in deeper trouble.

[13] The four theories are much easier to understand than most theories in physics.

3. Cosmological natural selection is the link between an almost empty, low-dimensional external world and a high-dimensional reality that is filled with complexity and dense brain matter. Due to evolutionary conservation, more complex entities always more efficiently reproduce and/or simulate the less complex entities on which they became evolutionarily dependent.

This shows that there are metaphysical accounts that give clear-cut answers to the four theories, and this is how it should be for non-complex theories. However, the point of EVAAN's 1a is that N&E does not provide such answers.

4.1.1 The exceptional-point-of-view problem

The exceptional-point-of-view problem of Blondé and Jansen (2021) is a subjective problem of consciousness that is about the question how it can be explained that at least one intelligent person experiences the consciousness of a relatively tiny, intelligent brain in the middle of an enormous universe that is indifferent about which physical entities it brings about according to the laws of physics. According to Blondé and Jansen, the probability to find at least one intelligent consciousness in the universe is not equal to the probability that you find yourself as a first-person, intelligent consciousness. They make a calculation based on equiprobable mass: if a proverbial arrow is shot randomly in the matter of the observable universe, the probability to hit intelligent matter (brain matter or CPU matter) is less than 10^{-28}. This extremely[14] small probability, if N is true, is the maximum probability that you have a first-person experience.

There are two supernaturalistic solutions to this problem: either we have a supernatural soul that attracts the proverbial arrow, or the majority of the physical matter in reality consists of brain matter that generates intelligent consciousness. These solutions are not empirically verifiable and evolutionary naturalists are left with different options on how to solve the problem. If a selection on the basis of equiprobable mass is rejected, was our consciousness selected among equiprobable conscious animals (Nagel 1974)?[15] Equiprobable brain size-moments (Pereira 2017, p. 4-5)?[16] Or equiprobable intelligent beings (Leslie 1992)?[17] These different options have been advocated by different philosophers in a naturalistic context, which makes the problem controversial among evolutionary naturalists. This is an indication that beliefs that are not empirically verifiable (hence metaphysical beliefs) are apparently unreliable, if N&E is true.

4.1.2 The Sleeping Beauty problem

The Sleeping Beauty problem was brought to the attention of the philosophic community by Adam Elga (2000, p. 143):

---

14 A probability of 10^{-2} already makes N improbable on statistical grounds.
15 Nagel reminds us that also bats have a consciousness.
16 Pereira uses this criterion in his Super-Strong Self-Sampling Assumption.
17 According to Leslie, we were randomly selected from all intelligent beings.

> Some researchers are going to put you to sleep. During the two days that your sleep will last, they will briefly wake you up either once or twice, depending on the toss of a fair coin (Heads: once; Tails: twice). After each waking, they will put you to [sic] back to sleep with a drug that makes you forget that waking. When you are first awakened, to what degree ought you believe that the outcome of the coin toss is Heads?

For philosophers, heads and tails represent two theories about all of reality that are equiprobable a priori, whereas the three awakenings from sleep are births of observers in the universes that are predicted by the theories. Elga demonstrates that the answer is 1/3, because the three possible awakenings are equiprobable. After all, if the experiment is repeated many times and you had to bet for money each awakening, you would make most profit by assuming that the answer is 1/3. Nevertheless, there is a school of thought that claims the answer is 1/2 (Lewis 2001; Bostrom 2007, p. 59). Bostrom, a leading scientist in this debate, even proposes a hybrid model in which the answer is neither 1/2 nor 1/3, but something in between. His main argument is the following thought experiment about a presumptuous philosopher (Bostrom 2007, p. 64):

> *Presumptuous philosopher*
> It is the year 2100 and physicists have narrowed down the search for a theory of everything to only two remaining plausible candidate theories, T1 and T2 (using considerations from super-duper symmetry). According to T1 the world is very, very big but finite and there are a total of a trillion trillion observers in the cosmos. According to T2, the world is very, very, *very* big but finite and there are a trillion trillion trillion observers. The super-duper symmetry considerations are indifferent between these two theories. Physicists are preparing a simple experiment that will falsify one of the theories. Enter the presumptuous philosopher: "Hey guys, it is completely unnecessary for you to do the experiment, because I can already show to you that T2 is about a trillion times more likely to be true than T1!"

And then it turns out, so Bostrom warns, that the presumptuous philosopher is wrong. In other words, Bostrom denies the rules of logic in order to safeguard the metaphysical intuition that empirical verifiability has priority over logical reasoning. If, on the other hand, we follow the logical evidence of Sleeping Beauty where it leads, we have to conclude that there are an infinite number of observers in reality. Bostrom (2007, pp. 64-65) admits that this is the case. However, he refuses to accept this conclusion as a certainty, again, because it is not an empirically observed fact.

David Lewis, who also opposed Elga, was probably worried that his possible worlds would no longer be equiprobable, which could falsify his theory of modal realism via empirical methods (Lewis 1986). We can conclude that the Sleeping Beauty problem is clearly controversial among scientists in general, and, therefore, among evolutionary naturalists, because its logic-based outcome is not empirically verifiable in cosmology. If even purely mathematical outcomes are altered by some evolutionary naturalists because of empirical verifiability, this is a strong indication that premiss 1a of my EVAAN is apparently true.

4.1.3 Cosmological natural selection

Cosmological natural selection is a theory proposed by the evolutionary naturalist physicist Lee Smolin (1992), who argues that the whole observable universe can reproduce and is subjected to evolution by natural selection within a multiverse. However, this idea is rejected by other evolutionary naturalists because the empirical verification of it is currently controversial. Rothman and Ellis (1993, p. 203) remark that Smolin's cosmological natural selection merely mimics biological evolution. It is ironic that the *a priori* structure of E is rejected by some of those who believe N&E is true. Apparently, E is only accepted by evolutionary naturalists as far as it is empirically verifiable. Outside of that domain, there is disagreement among evolutionary naturalists. When cosmological natural selection is applied on a multiverse that contains all the logical possibilities, it is clearly a metaphysical theory, because it entails the existence of multi-dimensional agents that hide for our evolutionarily conserved empirical sciences (Blondé 2019). Again, this hints that premiss 1a is apparently true.

4.1.4 The multiverse hypothesis

The hypothesis of a multiverse, or an ensemble of worlds, has been controversial among evolutionary naturalists since it was first proposed (Carter 1974, Ellis 2008, Kragh 2009). The main reason why an evolutionary naturalist must accept the multiverse hypothesis, is that our observable universe, considered in isolation, has laws and fundamental parameters that are highly fine-tuned to enable intelligent life.[18] If only the laws or parameters are modified slightly, a universe follows that is devoid of complexity, let alone intelligent life. If, on the other hand, a sufficiently large multiverse exists in which other universes exist with different parameters, then the fine-tuning of our universe can be explained by an observation selection bias: we can only exist in one of the few fine-tuned universes, so we observe a fine-tuned universe. This explanatory method is called the anthropic principle, and evolutionary naturalists need it to avoid a supernatural fine-tuner.

In spite of this, many evolutionary naturalists reject the multiverse hypothesis because it is not empirically verifiable (Kragh 2009). The evidence-based cosmologist George Ellis (2008, p. 2.35) writes the following:

> The multiverse is a reasonable theoretical explanation of the fine-tunings, but this does not help in observationally confirming the hypothesis. The issue here is, which is more important in cosmology: theory (explanation) or observations (tests against reality)? The essential proposal is that one should downgrade observational testing in favor of theory – a dangerous road to take.

Also here, evolutionary naturalists are in doubt about which side to rely on with respect to what exists in metaphysics: empirical observation or logical theory. This shows again that, apparently, if N&E is true, metaphysical beliefs are unreliable.

---

[18] A multiverse that explains fine-tuning is in a grey area with respect to being metaphysical. However, if even such a multiverse is controversial among evolutionary naturalists, then certainly also a multiverse that contains universes that are not required to explain fine-tuning.

That proves premiss 1a of my EVAAN via *a posteriori* evidence in the philosophic and scientific literature.

4.2 An *a priori* just-so story to prove 1a

I propose three *a priori* explanations of why premiss 1a of my EVAAN is apparently true: if N&E is true, 1) metaphysical beliefs are not likely to impress peers and sexual partners, 2) spending thoughts on metaphysical beliefs requires often-scarce energy, and 3) applying logical rules to derive metaphysical beliefs can be mentally destabilizing. In all these cases, evolution has either avoided the formation of cognitive faculties that generate metaphysical beliefs, or it gave rise to cognitive faculties that suppressed or overturned this kind of belief content. As a result of that, if N&E is true, our brain is designed such that metaphysical beliefs must be unreliable. This shows that premiss 1a must be true and explains why it is apparently true.

With respect to the first explanation, if N&E is true, an ancestor A who tried to impress peers or sexual partners with a metaphysical belief B that is logically verifiable, but not empirically verifiable, was often faced with the claim that ancestor A was making a reasoning error or even using the wrong logic, in the style of Lewis and Bostrom, who deny Elga's 1/3 solution to a problem as simple as that of Sleeping Beauty. Because ancestor A had no empirical proof, nor any kind of experiment, the attempt to impress failed, and the cognitive faculty responsible for generating belief B proved to be useless.

The second explanation builds on the fact that the human brain accounts for no less than 20% of the calorie consumption in the human body (Raichle and Gusnard 2002). Therefore, it must have been a favorable evolutionary strategy not to create cognitive faculties that are preoccupied with beliefs that have no clear evolutionary benefit, in order not to consume scarce energy. However, most metaphysical beliefs do not have a clear evolutionary benefit, because they are about those parts of reality with which evolutionary humans do not interact directly, if N&E is true.

What regards the third explanation, it must be said that, if N&E is true, applying logical rules to derive metaphysical beliefs can result in delusions, paranoia, and suicidal behavior, especially if beliefs are derived that affect the individual emotionally (Mujica-Parodi et al. 2000). One example of this is thought broadcasting: the belief that other people can hear an individual's thoughts (Pawar and Spence 2003). Other examples are the belief that one is possessed by demons, that one will go to a hell as a consequence of taking certain actions, and that one is being watched and judged continuously. In other words, the evolutionary naturalist remains shielded from the potential complexity of an imagined supernatural reality. This is a blessing with respect to mental health, but it prohibits metaphysical intelligence, if N&E is true.

The close relationship between metaphysical beliefs and mental disorder has been used as an argument against supernaturalism (Dawkins and Ward 2006). However, it sustains premiss 1a of my EVAAN, because it follows from the truth of N&E that E will invert or weed out such beliefs. As a result of that, the evolutionary naturalist is left with metaphysical intuitions that are misguided or poorly developed, and thus unreliable.

Again, these three *a priori* explanations form only a just-so story in evolutionary psychology. However, I contend that it is a more refined, and therefore better,

story than the just-so story evolutionary naturalists use to defend an unrestricted evolutionary reliability (If N&E is true, then our beliefs are reliable on all epistemic domains.). For example, Boudry and Vlerick (2014, p. 68) just posit that 'natural selection will weed out neural structures that give rise to false beliefs,' because 'true beliefs are overall better guides to action in the world than false beliefs.'

As a conclusion of this section, premiss 1a of my EVAAN (If N&E is true, metaphysical beliefs are unreliable.) has been shown to be true via both an *a posteriori* and an *a priori* methodology. The *a priori* methodology demonstrates that 1a is true and explains why 1a is apparently true according to the *a posteriori* methodology. If N&E is true, our evolutionary history is apparently so determinative that some even deny logic-based conclusions in order to hold on to their empiricism-based worldviews.

5. Strengthening EAAN

5.1 Premiss 1b of EVAAN

Now I will demonstrate that premiss 1b of my EVAAN (N&E is a metaphysical belief.) is true. I defined N as metaphysical naturalism, namely the claim that all and only the natural entities exist in reality. N is therefore a belief in the non-existence of the class of the logically possible, non-natural entities. According to my definition of a metaphysical belief, N is a metaphysical belief. This shows that N and, as a consequence, the conjunction N&E, are metaphysical beliefs. This proves EVAAN's premiss 1b.

5.2 Premiss 2 and conclusion of EVAAN and EAAN

From premisses 1a and 1b of my EVAAN it follows that if N&E is true, the belief in N&E is not reliable. This is premiss 2, shared by both EAAN and EVAAN. Then the shared conclusion follows that N&E is self-defeating: it proves its own unreliability.

Because it is difficult to explain the biological evidence record while N is true and E is false, EAAN and EVAAN are in the first place arguments against N, rather than against E. The opposite scenario, N being false while E is true, only increases the resources to explain biology as compared to N&E, by adding supernatural explanations. Therefore, EVAAN, just like EAAN, is an evolutionary argument against naturalism.

With that I have strengthened the conclusion of Plantinga's EAAN: even if it has to be admitted that empirically verifiable beliefs are often reliable if N&E is true, contra premiss 1 of Plantinga's EAAN, then still it can be shown that the belief in N&E is not reliable if N&E is true and that, therefore, N&E is self-defeating. This is in accordance with premiss 2 and the conclusion of EAAN and EVAAN.

6. Discussion

There are many types of reasoning: inductive and deductive reasoning, analogical reasoning, generalization, abstraction, etc. What my EVAAN reveals is that, if N&E

is true, humans are bad in applying these reasoning rules correctly to metaphysics, because this can result in brain states that are evolutionarily harmful. Whereas reasoning may be reliable in verifying beliefs about the empirically accessible domain, it breaks down as soon as a metaphysical conclusion is reached. Four examples of such a breakdown have been given in Section 4.1.

An objection to my EVAAN could be to invert my *a priori* arguments for premiss 1a:

Premiss-1a inversion: If N&E is true, then some beliefs about the supernatural are fitness enhancing.

The idea of premiss-1a inversion is to reduce beliefs about the supernatural to their presumed evolutionary origins, in order to discredit them (Boyer 2007, Peoples et al. 2016). Instead of being the result of a correctly applied logical reasoning, supernatural beliefs would then be merely hardwired in our biological predisposition. First of all, premiss-1a inversion cannot make supernaturalism self-defeating, because it starts from the truth of N&E. Second, even if this premiss is true, it further supports my premiss 1a (If N&E is true, metaphysical beliefs are unreliable.). We have to distinguish two kinds of beliefs that are metaphysical: those that are fitness diminishing and those that are fitness enhancing. These two kinds of beliefs can be held by a single individual at the same time. In both cases, N&E ensures that humans have certain hardwired beliefs about metaphysics, which corroborates that they are unreliable, if N&E is true.

A second objection could be the claim that, if supernaturalism were true, there would be no controversies about the truth of N&E. Here, I argue that evolutionary naturalists may be supernaturally predestined to believe in N&E, for example, because they have an evolutionarily conserved role to play in the historical development of the sciences (Blondé 2019).

A third objection could be the proposition that metaphysical naturalism does not even require a proof, because of its connection with methodological naturalism and (among other things[19]) 'the massive amount of knowledge gained by it' (Forrest 2000). It has to be admitted that methodologically distinguishing entities that are revealed by the empirical sciences from entities that escape these sciences is a useful scientific endeavor.[20] However, this has nothing to do with the bold claim of metaphysical naturalists that the latter entities are absent from reality. Such a claim certainly requires a proof.

7. Conclusions

The crucial foundation below my EVAAN is its premiss 1a (If N&E is true, metaphysical beliefs are unreliable.), for which both an *a posteriori* and an *a priori* methodology are provided as a proof. The *a posteriori* methodology consists of four non-complex, logically verifiable theories that are not empirically verifiable,

---

[19] Forrest mentions four reasons: (1) the demonstrated success of methodological naturalism, (2) the massive amount of knowledge gained by it, (3) the lack of a method or epistemology for knowing the supernatural, and (4) the subsequent lack of evidence for the supernatural. I summarized a method for knowing the supernatural in Section 4.1.

[20] Plantinga (2001) believes otherwise: if metaphysical naturalism is rejected, then methodological naturalism should also be rejected.

and the observation that evolutionary naturalists cannot find agreement about such theories. They are the exceptional-point-of-view problem, the Sleeping Beauty problem, cosmological natural selection, and the multiverse hypothesis. The *a priori* methodology provides explanations of why premiss 1a is apparently true according to the *a posteriori* methodology: N&E has never favored cognitive faculties that are preoccupied with metaphysical beliefs, because such faculties 1) are not likely to impress peers and sexual partners, 2) result in an unnecessary energy consumption, and 3) can be mentally destabilizing.

It has to be acknowledged that the *a priori* proofs of my premiss 1a (see above) and my premiss-1 rejoinder (If N&E is true, empirically verifiable beliefs are often reliable.) are evolutionary just-so stories, or hypotheses that may not be entirely accurate. However, they are clearly more refined – and therefore better – than the just-so stories of the premisses they refute respectively: the unrestricted evolutionary reliability (If N&E is true, then our beliefs are reliable on all epistemic domains.) and Plantinga's premiss 1 (If N&E is true, then all our beliefs are unreliable.). My EVAAN, therefore, provides a more nuanced middle way between these two extremes.[21]

8. Bibliographic references

Arneth, B. M. (2009), 'Sexual selection and intelligence: Does sexual reproduction drive the evolution of intelligence?' *Bioscience Hypotheses*, 2(4): 209-212.

Beilby, J. K., ed., (2002), *Naturalism Defeated?: Essays on Plantinga's Evolutionary Argument Against Naturalism*. Cornell University Press.

Blondé, W. (2015), 'An evolutionary argument for a self-explanatory, benevolent metaphysics', *Symposion*, 2(2): 143-166.

Blondé, W. (2016), 'Can an eternal life start from the minimal fine-tuning for intelligence?', *Philosophy and Cosmology*, 17(17): 26-38.

Blondé, W. (2019), 'EMAAN: An evolutionary multiverse argument against naturalism', *Symposion*, 6(2): 113-128.

Blondé, W. and Jansen, L. (2021), 'Proving God without dualism: Improving the Swinburne-Moreland argument from consciousness', *Metaphysica*, 22(1): 75-87.

Bostrom, N. (2007), 'Sleeping beauty and self-location: A hybrid model', *Synthese*, 157(1): 59-78.

Boudry, M. and Vlerick, M. (2014), 'Natural selection does care about truth', *International Studies in the Philosophy of Science*, 28(1): 65-77.

Boyer, P. (2007), *Religion Explained: The Evolutionary Origins of Religious Thought.* UK: Hachette.

Carter, B. (1974), 'Large number coincidences and the anthropic principle in cosmology', *Confrontation of cosmological theories with observational data*, Springer, Dordrecht, 291-298.

Dawkins, R., and Ward, L. (2006), *The God Delusion.* Boston: Houghton Mifflin Company.

Deem, M. J. (2018), 'A flaw in the Stich–Plantinga challenge to evolutionary reliabilism', *Analysis*, 78(2): 216-225.

---

[21] I wish to thank Ludger Jansen, Michael Vlerick, and several anonymous reviewers for helpful suggestions and comments.

Elga, A. (2000), 'Self-locating belief and the Sleeping Beauty problem', *Analysis*, 60(2): 143-147.

Ellis, G. (2008), 'Opposing the multiverse', *Astronomy & Geophysics*, 49(2): 2-33.

Fales, E. (1996), 'Plantinga's case against naturalistic epistemology', *Philosophy of Science*, 63(3): 432-451.

Forrest, B. (2000), 'Methodological naturalism and philosophical naturalism: Clarifying the connection', *Philo*, 3(2): 7-29.

Gould, S. J. (1978), 'Sociobiology: The art of storytelling', *New Scientist*, 80(1129): 530-33.

Gould, S. J., and Lewontin, R. C. (1979), 'The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme', *Proceedings of the Royal Society of London B*, 205(1161): 581-598.

Kragh, H. (2009), 'Contemporary history of cosmology and the controversy over the multiverse', *Annals of Science*, 66(4): 529-551.

Leslie, J. (1992), 'The Doomsday argument', *The Mathematical Intelligencer*, 14(2): 48-51.

Lewis, D. K. (1986), *On the Plurality of Worlds*. Vol. 322. Oxford: Blackwell.

Lewis, D. (2001), 'Sleeping Beauty: Reply to Elga', *Analysis*, 61(3): 171-176.

Łowicki, P., and Zajenkowski M. (2017), 'Divine emotions: On the link between emotional intelligence and religious belief', *Journal of Religion and Health*, 56: 1998-2009.

Mahner, M. (2012), 'The role of metaphysical naturalism in science', *Science & Education*, 21(10): 1437-1459.

Miller, G. (2000), 'Sexual selection for indicators of intelligence', In *Novartis Foundation Symposium*, 260-270, Chichester; New York; John Wiley.

Mujica-Parodi, L. R., Malaspina, D., and Sackeim, H. A. (2000), 'Logical processing, affect, and delusional thought in schizophrenia', *Harvard Review of Psychiatry*, 8(2): 73-83.

Nagel, T. (1974), 'What is it like to be a bat', *The Philosophical Review*, 83(4): 435–50.

O'Connor, T. (1994), 'An evolutionary argument against naturalism?' *Canadian Journal of Philosophy*, 24(4): 527-539.

Pawar, A. V., and Spence, S. A. (2003), 'Defining thought broadcast: Semi-structured literature review.' *The British Journal of Psychiatry*, 183(4): 287-291.

Peoples, H. C., Duda, P., and Marlowe, F. W. (2016), 'Hunter-gatherers and the origins of religion', *Human Nature*, 27(3): 261-282.

Pereira, T. (2017), 'An anthropic argument against the future existence of superintelligent artificial intelligence', *arXiv preprint*, arXiv:1705.03078.

Plantinga, A. (1993), *Warrant and Proper Function.* Oxford University Press.

Plantinga, A. (2001), 'Methodological naturalism', *Intelligent design creationism and its critics: Philosophical, theological, and scientific perspectives*, 339-361.

Plantinga, A. (2011), *Where the Conflict Really Lies: Science, Religion, and Naturalism.* OUP USA.

Raichle, M. E., and Gusnard, D. A. (2002), 'Appraising the brain's energy budget', *Proceedings of the National Academy of Sciences*, 99(16): 10237-10239.

Rothman, T. and Ellis, G. F. R. (1993), 'Smolin's natural selection hypothesis', *Quarterly Journal of the Royal Astronomical Society*, 34: 201.

Sage, J. (2004), 'Truth-reliability and the evolution of human cognitive faculties', *Philosophical Studies*, 95-106.

Slagle, J. (2021), *The Evolutionary Argument Against Naturalism: Context, Exposition, and Repercussions.* Bloomsbury Publishing.

Smolin, L. (1992), 'Did the universe evolve?' *Classical and Quantum Gravity*, 9(1): 173.

Stich, S. P. (1990), *The Fragmentation of Reason: Preface to a Pragmatic Theory of Cognitive Evaluation.* Cambridge, Mass: MIT Press.

Stoltzfus, A. (1999), 'On the possibility of constructive neutral evolution', *Journal of Molecular Evolution*, 49(2): 169-181.

Wilson, M. L., Miller, C. M., and Crouse, K. N. (2017), 'Humans as a model species for sexual selection research', *Proceedings of the Royal Society B: Biological Sciences*, 284(1866), p.20171320.