These lecture notes include commentary on Plantinga’s argument on the irrationality of agnosticism in his *Warranted Christian Belief*, and then elides to his defense of the Evolutionary Argument against Naturalism in his *Where the Conflict Really Lies*.

Recap from the argument against agnosticism
1. If I do not know or believe anything about my origins, then I cannot determine R
2. I do not know or believe anything about my origins
3. Therefore, I cannot determine R
4. If I cannot determine R, then I should be uncertain about the beliefs that my faculties produce.
5. Therefore, I should be uncertain about the beliefs that my faculties produce.
6. One of the beliefs my faculties produce is (2).
7. Therefore, if I believe in (2), I should be uncertain of (2). I.e., (2) is self-defeating, if I believe it, it is irrational to believe it.

Here is a second extended explanation:

The first premise might simply be the thesis of agnosticism, according to which,
1. I do not know or believe anything about my origins/purpose/provenance. Notice that this claim itself is making a claim, it is a belief, namely, a belief about what I believe or know. So, agnosticism does involve beliefs, even if the belief content is not about God’s existence directly.
2. If I do not know or believe anything about my origins/purpose/provenance, then I cannot determine whether my cognitive faculties are reliable (following Plantinga, let that be abbreviated as R).

Premise 2 is defended on analogy with the radio. The radio represents my cognitive faculties, the English sentences the radio is expressing represents the beliefs produced by our faculties. If I have no idea where or why the radio is, I have no reason for trusting the truth of its outputs. It follows that,

3. Therefore, I cannot determine R
4. If I cannot determine R, then I should be uncertain about the beliefs that my faculties produce. I have no reason for thinking that they are true, just as I have no reason for thinking that the radio is expressing true claims.

5. Therefore, I should be uncertain about the beliefs that my faculties produce.
6. One of the beliefs that my faculties produce is agnosticism itself.

(Thomas challenged this claim in that agnosticism is not a belief at all. Good point. But the agnostic has to be committed to something. If asked "do you believe that theism is true?" the
agnostic must respond "no" and that response is itself a belief: i.e., a belief that I don't believe that theism is true. A second point is that in order to know something I must believe it. So, if I do not even believe that P (whatever P is), then I cannot know it either. Likewise for warranted, or rational, or evidentially supported belief since they all require belief as well. So, if agnosticism is simply the absence of belief, then an agnostic cannot know, or be warranted or rational etc. in her agnosticism. It comes to the same conclusion just a bit quicker.)

In any case, I think Plantinga wants to conclude from all of this that, 7. Therefore, if one is an agnostic, she has no reason for trusting the outputs of her cognitive faculties, which would include the beliefs that support her agnosticism

I say beliefs that support her agnosticism on the condition that agnosticism is not a belief itself, but remains supported by other beliefs. If it is not supported by other beliefs then for it to count as rational or warranted, it must be properly basic. I do not know of any agnostic who is prepared to say that agnosticism is a basic belief since that requires assuming that her cognitive faculties are reliable. And this argument shows that this assumption is unwarranted.

Lecture on the EAAN argument

Introductory thought experiment:
Drug Case 1:
Suppose you have just been discharged from the hospital after having received treatment for concussion. You notice on your drive home that foxes are passing across your visual field. After a few occasions of this image, you mention to your friend who’s picked you up from the hospital, that you keep on seeing foxes crossing the street. Your friend responds to your report saying, “oh, yes, I forgot to tell you. The doctor informed me that the treatment for your concussion is a weird drug that may cause you to hallucinate seeing foxes crossing streets. In preclinical trials it causes such hallucinations in some patients up to 80% but in others as low as 10%.” Assuming your friend is not known to be a huckster or mountebank, what does her report do to your belief that there are in fact foxes crossing the road? (Notice, your friend does not confirm or deny that there are in fact foxes that are crossing the road.)

- Commentary: there are 2 lessons to draw from the case. First, even if the perception of the foxes comes with a veridical phenomenology - you really feel like you see foxes - your friends report should lesson your confidence in the belief that there are foxes crossing the road. 2nd, you don’t know whether you’re in the 80% or the 10%. That is, in some patients 8 out of 10 perceptions of foxes are hallucinations and in others it is one out of 10. Even if you don’t know which type of patient you are and how the drug is affecting you, you still have reason for doubting your belief that foxes are crossing the road.
Case 2:
Recipe: construct a case where the possible hallucination cannot be verified from the first-person perspective. Consider a case where the drug affects our moral intuitions.

Suppose as above, that you’ve been discharged from the hospital having received treatment for concussion. The drug, however, causes a suppression of certain emotions associated with anger. And suppose as above your driving home with a friend and your friend sees a motorist cut off an elderly driver complete with sign language. Your friend exclaims that this is an unjust action against the elderly driver. You don’t see it as such; you just see a motorist who was in a hurry and the elderly driver was in her way. The elderly driver should not have gotten in her way.

Commentary: here’s a case where the empirically verifiable facts of the matter are agreed-upon. Disagreement occurs at a level of moral judgment. Since moral judgment involves value perception and values are not empirical entities, there is no fact of the matter that can adjudicate your differences of judgment. Suppose, however, that your friend now informs you that your treatment involves a drug that compromises your emotions of empathy and anger. (Your friend further explains the empirical evidence in moral psychology associating empathy and anger with the ability to perceive violations of justice.) Here too, your moral judgment that the elderly person has done something wrong, or at least, you fail to see that the other driver did something wrong, comes with a phenomenology as clear as ever. But knowing that you have received this drug and believing in the cognitive science of moral perception, you should hold your judgment that there was no wrongdoing with suspicion.

The point of these 2 cases held in contrast side-by-side, is to illustrate that a clear and distinct phenomenology still does not justify maintaining one’s belief in those cases. And things get worse when the fact of the matter is not verifiable by all parties. One’s moral cognitive apparatus has to be finely tuned in order to pick up on the moral signals in one’s environment. Absent the right apparatus, in this case, certain emotions associated with justice judgments, one will not form those very judgments. But that hardly means that injustice has not occurred. Knowing that one’s apparatus has been compromised in this respect justifies suspicion.

In both cases, You no longer have a reason for trusting your perceptual (moral or otherwise) faculties. We usually assume that our cognitive faculties are reliable, and rightly so, they work! Plantinga’s argument wants to suggest that if we take seriously the story as to how our cognitive faculties formed on the typical evolutionary naturalistic count, we are not going to have a reason for trusting them. We only have reason for thinking that the cognitive faculties will conduce to survival, but not truth. [Read from WPF, the quotations from Darwin, and Churchland].

I. The argument
Let R = our cognitive faculties are reliable. N = naturalism. E = our cognitive faculties have come to be by way of mechanisms proposed by contemporary evolutionary theory, namely, random genetic mutation. Let P (X/Y) = The probability of X given Y.

1. P(R/N&E) is low.
2. Anyone who believes N&E, and believes (1), has a defeater for R. (One has reason for doubting R or for being suspicious of R).
3. Anyone who has a defeater for R has a defeater for any other belief she thinks she has, including N&E itself. (This is because R refers to the reliability of one’s cognitive faculties and one’s beliefs are produced by those faculties.)
4. This defeater cannot be defeated itself. (This is because the Defeater pertains to the reliability of one’s faculties. Using one’s judgments or beliefs to justify the reliability of one’s faculties is using the very faculties that are in question, pp. 340-341).
5. Therefore, anyone who believes N&E and (1), has no reason for believing N&E; i.e. N&E is self-defeating, if you believe it you have no reason for believing in it.

II. What the argument is not saying:

- Plantinga is not arguing against evolutionary theory (E)
- He is not arguing for theism
- He is not arguing that naturalism (N) is false
- He is arguing that belief in (E & N) supplies an undercutting defeater for N. That is, belief in naturalism is irrational.

III. In defense of the premises

The focus of the reading today was in defense of premise 1. The person who is committed to N&E has to give us a reason for thinking 2 things: that the purpose of our cognitive systems is to serve us with true beliefs and that they do in fact furnish us with mostly true beliefs.

Importantly, one cannot assume that naturalism is actually the right metaphysical worldview. This is important because Plantinga grants that our faculties are reliable and that many of our beliefs are true. The temptation would be for the naturalist to infer inductively from the truth of many of our beliefs, and the reliability of our faculties more generally to the judgment that we hit the evolutionary lottery. We got epistemically lucky, he might say. There is, thus, no reason to doubt the reliability of one’s cognitive faculties given E&N. But it could be equally true that theism is true and that evolutionary mechanisms selected for reliable cognitive faculties with God’s help. And that is why we have the reliable faculties that we do. For the naturalist to debunk the theistic explanation, he must show that N predicts the formation of reliable cognitive faculties without appeal to how things are. So one cannot infer from R, to the claim that ‘E&N selected for reliable cognitive faculties.’ The inference would have to assume that N is true. Rather, we must look at what N predicts without appeal to how things are.
IV. The probability thesis (premise 1)

Why think that the probability thesis is true? Given strictly E and N, we have 4 predictions.¹

1. Evolutionary mechanisms would not give rise to cognitive faculties that produce beliefs at all. All that counts is that the nervous system produces behavior by the organism that is adaptive and conducive to survival. (Think of an ant colony, or even certain types of plants - such as the Venus flytrap - both of which are examples where adaptive behavior occurs without the supposition that any beliefs are involved.)

2. Evolutionary mechanisms would give rise to cognitive faculties that produce beliefs - beliefs with representational content - but those beliefs are not involved in the organism’s behavior. (This option seems to be the only option available to the materialist, whether of the non-reductive or reductive varieties.)

3. Evolutionary mechanisms would give rise to cognitive faculties that produce beliefs, those beliefs could be involved in the organism’s behavior, but are maladaptive. (On this 3rd option we can understand that the belief content could cause maladaptive behavior, but only those organisms whose neurophysiological states actually cause the organism to be in the right place at the right time survive. Only those whose neurophysiological states cause adaptive behavior, not the belief content, survive.) See slide 5.

4. Evolutionary mechanisms would give rise to cognitive faculties that produce beliefs, those beliefs are involved in causing an organism’s behavior, and that behavior is adaptive.

These are the 4 logically possible options that E and N would predict. The first point to observe is that the probability of one’s beliefs being true are zero on the first option, and very low on options 2 in 3. Overall, the probability that evolutionary mechanisms would produce faculties that in turn produce true beliefs is very low.

(A) Option 2

Individually, options 2 and 4 are the only possible options where evolutionary mechanisms might produce reliable cognitive faculties. Consider option 2 first. Plantinga spends most of his time arguing that on option 2, we have no reason for thinking R. Why? Assume that naturalism entails materialism with respect to the mind. There are 2 versions of materialism regarding the ontological status of one’s beliefs. On materialism, one’s beliefs have a neurophysiological component to them as well as propositional content. Beliefs represent states of affairs (e.g. it is raining outside) or ideas (e.g. naturalism is true). They have what Plantinga refers to as mental content. Mental content is distinct from neurophysiological properties for the simple reason that representations can be either true or false, but neurophysiological states cannot be true or false. One neuron cannot say to another neuron I’m true and you are false. On non-reductive materialism, the neurophysiological states cause mental content, but mental content is not identical to that neurophysiological state. Hunger is caused by an empty stomach, low blood sugar levels etc., But hunger is not identical to, it is not ontologically the same thing as, an

¹ From Alvin Plantinga’s “Naturalism defeated,” (unpublished manuscript, 1994).
empty stomach, low blood sugar levels etc. Hunger is a feeling I have, not a state of my digestive system even if it is caused by the latter.

Because on materialism, the neurophysiological state causes the mental content, we have no reason for thinking that the mental content is involved in causing one’s adaptive behavior. All that matters is that the neurophysiological state causes the right behavior. Mental content is invisible, if you will, to the evolutionary mechanisms. They can be anything whatsoever. So long as they are not involved in the production of behavior, it doesn’t matter whether they are true or false. Plantinga states, “But as long as the NP properties are adaptive, it doesn’t matter, for survival and reproduction, what content is determined by those NP properties. It could be true content; it could be false content; it doesn’t matter” (2010, 331). Further on, Plantinga supposes that the probability of any given belief (it’s mental content) being true is .5. But if I have 3 beliefs, each of which has a probability of being true at .5, the probability of all 3 beliefs being true together is .125. Suppose I have 1000 independent beliefs. The probability of all being true is $10^{-58}$ (333). So, given option 2, the probability of R/N&E is really low.

(B) Option 4

Now, one might think that we get epistemically lucky on option 4, but this is not the case. For Plantinga, it is important to understand that it is both belief and desire that factors into the etiology of one’s behavior. One might believe that everything other than a human being (e.g. tigers and wolves etc.) is a witch and desire to avoid witches. Such a belief-desire pair would be enormously adaptive if the person sees a tiger but believes (falsely) that the tiger is a witch. We can think of any number of belief-desire pairs according to which the belief turns out false but the behavior such a pair causes turns out to be adaptive. Let your imagination run wild. Plantinga supposes that the early hominids like very much to be eaten by a tiger but always run off looking for a better prospect, “because he thinks it unlikely that the tiger he sees will eat him. This will get his body parts in the right place so far as survival is concerned, without involving much by way of true belief” (1993, 225). Naturalists themselves think that belief in God is false, but many evolutionary psychologists believe that it is adaptive. Suppose I see Sam getting eaten by a tiger. It is not necessary, given that visual presentation, that I form the belief that tigers are dangerous animals for which I should avoid. Adaptive behavior might just as well be served by the belief that tigers are the embodiment of the gods and periodically come down to take people, like Sam, back up into heaven. Pair that belief with the desire “I don’t want to go to heaven right now” and one gets the adaptive behavior of running from the tiger. These examples might strike you as extravagant, but their point is that strictly given E&N, we have no reason for supposing that the mental content of these beliefs would not actually be the case. If we think of the logical possibilities of how belief-desire pairs can be adaptive, the objective probability that the beliefs in those pairs are in fact true is quite low.

- To the riposte that adaptive behavior requires true belief, as with the zebra and frog examples canvassed by Plantinga, all that’s required is that the organism have accurate indicators that cause adaptive behavior. Plantinga’s the examples of anaerobic marine bacteria (page 328 FF) are instructive. Such bacteria have magnetosomes that tell them
where magnetic north is, and this causes them to travel to deeper depths where they would otherwise die in the oxygen-rich waters on the surface. Clearly, these magnetosomes are not themselves beliefs, nor do we need to suppose that they cause beliefs. Rather they can function as indicators. That is, we need not suppose that adaptive behavior is caused by propositional content of one’s mental states being accurate.

V. Premise 2

If a naturalist comes to believe that premise (1) is true, she suffers a Defeater for her belief in naturalism. Return to the 2 cases introducing this argument. If one finds out that there is no reason given her metaphysical views that her cognitive faculties will bequeath true beliefs, she has a reason for doubting the truth of those very beliefs.

But can one argue in the following way: “I believe in naturalism, and the typical evolutionary mechanisms purported to explain the evolution of my cognitive faculties. But I also believe that my faculties are reliable. Therefore, unguided evolutionary mechanisms produced reliable cognitive faculties. I have no reason for doubting their outputs.”

Michael Bergmann supposes that we can argue in this way and, in particular, the 2nd premise (of this argument) can be justified based on common sense. Belief in R can be formed in the basic way. I have non-propositional evidence for the truth of R. What is that evidence? The veridical phenomenology that accompanies many of my beliefs. My memories of the not-too-distant past just feel right and correct. My perceptions in normal lighting conditions absent any known mechanisms of hallucination just feel right and correct. My a priori judgments about simple arithmetic, or logical truths such as modus ponens just feel right and correct.

There are several points to consider vis-à-vis Bergman’s reply to Plantinga.

1. Irrelevance: Bergman’s key claim is that belief in R can be properly basic, although Bergman himself only says that belief in R is formed “in the basic way,” and on the “basis of common sense.” The examples he uses are perceptual and memorial beliefs to motivate the idea that I cannot doubt the veridicality of those beliefs. Even if Bergman is right on all these points, it still does not touch the EAAN argument because naturalism is not a view that is formed by one’s perceptual judgments. It is a philosophical judgment, a metaphysical view that arguably requires a college of faculties over and above one’s perceptual faculties. I have no reason given all of what Bergman has said, to think that R is true with respect to those college of faculties. I might think that my perceptual judgments are in fine order, but be cautious about my metaphysical views.

2. There is something deeply suspicious about Bergman’s move. The basic pattern is a Moorean shift. A Moorean shift is so-called because it replaces a modus ponens argument with a modus tollens argument or vice versa. A skeptical argument might run as follows: I know that I have hands only if I know that I’m not a brain in a vat. I do not
know that I am not a brain in a vat. Therefore I do not know that I have hands. The shift would be to argue instead as follows: I know that I have hands. If I know that I have hands then I know that I’m not a brain-in-a-vat. Therefore I know that I’m not a brain-in-a-vat. Notice that both arguments except the principle that knowing that I have hands requires knowing that I’m not a brain-in-a-vat.

Bergman’s reply appears to make a similar move. For Bergman, that my faculties are reliable is held constant as a premise in his reply. The form of his argument is as follows. My faculties are reliable. My faculties would be reliable only if they were produced by objectively improbable, but unguided evolutionary events. (The consequent grants the first premise in the EAAN argument.) Therefore, my faculties were produced by objectively improbable but unguided evolutionary events. You see! Instead of using the conclusion as a reason against R, Bergmann holds belief in R constant, and argues that you and I hit the evolutionary lottery with respect to our cognitive faculties.

Now I can state my suspicion. Suppose you are a naturalist and you think that the evidential argument from evil is a really good argument. Your naturalism is philosophically defended and in good epistemic order, you might think, because there is no God, and your reasons trace back to the evidential argument. Suppose the theist replies to your evidential argument claiming that “the propositional evidence against God’s existence is strong given the apparent gratuitous evils in the world. However, I have non-propositional evidence that God exists.” Suppose further that the theist adopts a Moorean shift on your argument from the apparent gratuitous evils in the world. That is, instead of arguing to God’s nonexistence, the theist holds constant the belief that God exists and concludes instead that for any apparent gratuitous evil, there must be a purpose for such evils. If you were a naturalist, and the theist responded to you in this way, you would be flummoxed. That would be a bad response mainly because it sounds so dogmatic, and unresponsive to potential defeaters. Likewise, if the naturalist asserts R, and this is the best route by which to repartee the EAAN argument, then the naturalist is committing a move she herself considers a cognitive sin in other trivially different contexts.

Recall that Plantinga grants that our faculties are reliable. But he thinks this because he believes that God guided the evolutionary mechanisms so that they would bequeath reliable cognitive faculties. That is how things stand in the actual world. Plantinga would agree, that is, that many of our beliefs are true and the ones that are true have a certain phenomenology to them; but he believes this, thanks be to God! One cannot argue from the actual standpoint to the conclusion that naturalism selected for reliable cognitive faculties. The reason why is because the actual standpoint could be a God governed world and that the reliability of our cognitive faculties are attributable to God’s agency not blind, unguided evolutionary processes.
Plantinga’s argument is best understood as assuming a metaphysically neutral standpoint *from the start*. One should not assume either that naturalism is true nor that it is false. Rather, one has to proceed through the argument asking oneself what *would* unguided evolutionary mechanisms produce (vis-à-vis reliable cognitive faculties).

There is an interesting symmetry to how the argument can go at this point. One can argue as follows:

| a. My faculties are reliable. |
| b. My faculties would be reliable only if they were designed by God. |
| c. Therefore, my faculties were designed by God. |
| d. My faculties are reliable. |
| e. My faculties would be reliable only if they were produced by objectively improbable, but unguided evolutionary events. |
| f. Therefore, my faculties were produced by objectively improbable but unguided evolutionary events. |

Now, the Naturalist might think that the argument a-c is a bad argument. But whatever criticisms are levelled against it could in principle be levelled against d-f, for example, the naturalist might think that the existence of God is implausible – but so is the likelihood of having reliable cognitive faculties on E&N. But the Naturalist *needs* d-f to rebut Plantinga’s second premise – the one Bergmann attacks. Why? Because Bergmann’s naturalist believes that R is true and forms it “in the basic way” as he says. But one can rely on common sense *after entertaining EAAN*, only if one thinks that common sense is a reliable belief producing mechanism that is designed to produce true beliefs. And one can only believe that (while holding to E&N) if one thinks that we hit the evolutionary lottery with regard to our cognitive faculties (i.e., f, in the argument d-f).

So, by parity, any Naturalist critique of a-c would apply to d-f. And if so, the Naturalist who critiques a-c has no reason for avoiding the second premise of EAAN (since a sound and valid d-f argument is needed to rebut the second premise). If the naturalist finds nothing to critique about a-c, then she has a reason for being a theist. So, either the naturalist has no reason for rebutting premise (2) of EAAN, in which case she has an undefeated defeater for N; or she has a reason for rebutting (2) but she also has an undefeated reason for theism as well.

The reasonable “move” might be to be agnostic. But then we encounter Plantinga’s argument for the self-defeating nature of agnosticism. It appears that without any argument for theism, Plantinga has given arguments that theism is the most rational of the options available. That seems to break the impasse even if we do not import arguments for/against God’s existence.

The End
A few miscellaneous points from previous drafts of these notes:

- Now the key point is to understand how Plantinga’s argument avoids the response that false belief would lead to maladaptive behavior. Since our behavior is adaptive, that would be reason for thinking that our cognitive faculties are reliable; they are delivering true beliefs.
- Plantinga’s riposte is to decouple behavior from belief. The nature of this decoupling is fairly straightforward in the reading for today. Plantinga makes a distinction between neural structures having 2 different types of properties, namely, their neuro-physiological properties and their mental content properties. The basic idea is that since the belief content is caused by the neurophysiological properties (on both non-reductive and reductive materialism) and it is the neurophysiological properties that cause behavior, beliefs would not necessarily figure in to the etiology of any action. Belief content would not have any purpose; they might be understood as unintended byproducts of chance evolutionary mechanisms. So, the argument that false beliefs would lead to maladaptive behavior needs to assume that belief content would affect an organism’s behavior given N&E. But there is no reason to suppose that on N&E think of an ant colony. Ants do not form beliefs, at least it is implausible to suppose so. And yet their behavior is adaptive. No doubt, organisms would interact with their environments on Plantinga’s view. But, he thinks that given N&E, we only have reason for thinking that the neurophysiological properties of a long-standing neural event would cause behavior, not the content of one’s beliefs. Again, explaining the adaptability of an ant colony does not have to include reference to true beliefs. What is more, not only is this hypothesis an option given N&E, but that N entails materialism, and materialism entails that the neurophysiological properties of a long-standing neural event cause a belief’s content. The content of a belief cannot be part of the etiology of an organism’s behavior.

- But there are other routes Plantinga has exploited in the past to motivate this very premise. Assume, somehow, that beliefs do figure in the causal chain of one’s behavior given strictly naturalism. What is the probability that an adaptive belief will be true given the suggestion? For Plantinga, it is important to understand that it is both belief and desire that factors into the etiology of one’s behavior. One might believe that everything other than a human being is a witch and desire to avoid witches. Such a belief desire pair would be enormously adaptive if the person sees a tiger but believes (falsely) that the tiger is a witch. We can think of any number of belief-desire pairs according to which the belief turns out false but the behavior such a pair causes turns out to be adaptive. Let your imagination run wild. Plantinga supposes that the early hominids like very much like to be eaten by a tiger but always run off looking for a better prospect, “because he thinks it unlikely that the tiger he sees will need him. This
will get his body parts in the right place so far as survival is concerned, without involving much by way of true belief” (1993, 225).

- To be more precise, Plantinga entertains 5 possibilities regarding the relationship between naturalism and the functioning of our cognitive faculties.
  o Naturalism might predict that our beliefs are not causally connected to behavior at all, and this is because naturalism might predict that we don’t form beliefs!
  o It might predict that beliefs are epiphenomenal effects of causes, those causes of which also cause behavior. On this supposition we do have beliefs, but they are not connected to behavior in the right way. Here they are causally connected to behavior but not as causes themselves. orthodox biologists believe that behavior, however complex, is governed entirely by biochemistry and that the attendant sensations--fear, pain, wonder, love--are just shadows cast by that biochemistry, not themselves vital to the organism's behavior . . . ." Smith,
  o Belief is syntactically connected with behavior – a belief’s neurophysiological properties cause belief. In the previous three cases, belief is invisible to selection processes.
  o Belief is both syntactically and semantically connected with behavior, but is maladaptive.
  o Belief is connected with behavior and is adaptive.

Possible repartees:
(i) The argument points out that one has a defeater for R, but not an undefeated defeater. One could argue that the undercutting defeater provided by premise (1) is defeated by other evidence.
(ii) The conclusion of self defeat only follows if one accepts both N&E and premise (1). Why can’t one simply dismissed premise (1) by conjoining N&E with a standard of rationality according to which one should avoid self-defeating networks of belief.
(iii) Others?