

DESIGN DISCOURSE AND THE COGNITIVE SCIENCE OF DESIGN

JEROEN DE RIDDER\*

1. *Introduction*

Much of Alvin Plantinga's *Where the Conflict Really Lies* (2011) will contain few surprises for those who have been following his work over the past decades. This — I hasten to add — is nothing against the book. The fact alone that his ideas on various topics, which have appeared scattered throughout the literature, are now actualized, applied to the debate about the (alleged) conflict between science and religion, and organized into an overarching argument with a single focus makes this book worthwhile. Moreover, I see this book making significant progress on two opposite ends of the spectrum of views about science and religion. On the one end, we find the so-called new atheists and other conflict-mongers. Compared to the overheated rhetoric that oozes from their writings, this book is a breath of fresh air. Plantinga cuts right to the chase and soberly exposes the bare bones of the new atheists' arguments. It immediately becomes clear how embarrassingly bare these bones really are. On the other end of the spectrum are theologians and scientists who envisage harmony and concord between science and religion. However, it is not uncommon to see harmony and concord attained in less than desirable ways. Traditional orthodox Christian belief undergoes significant revisions to the point of becoming hardly recognizable as Christian belief, or one is invited to swallow philosophical positions clothed in vernacular that is far from easy to grasp, let alone digest. Plantinga's commitment to Christian orthodoxy and his unwavering analytical approach constitute a major improvement here. Regardless of whether or not one ultimately agrees with him, Plantinga makes it clear how he sees deep concord between science and orthodox Christianity and he does so in a commendably straightforward style.

My topic for this essay, however, is an element of Plantinga's narrative about which he seems to have had a change of mind over the years. It is how to understand and evaluate teleological arguments, such as biological design arguments from irreducible complexity and the fine-tuning argument. In the next section, I sketch out how Plantinga proposes to recast design arguments as 'design discourse'. In section 3 I consider whether evolutionary theory spells trouble for design discourse and argue that it does not. In section 4, I discuss results from cognitive science that shed light on how we form design beliefs. In sections 5 through 7, I consider various ways in which these results might be used to mount an objection to design discourse and argue that none of them succeeds. Section 8 concludes the paper.

---

\* Jeroen de Ridder, Faculty of Philosophy, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, the Netherlands. Email: g.j.de.ridder@vu.nl.

## 2. *Plantinga on Design Discourse*

Plantinga devotes Chapters 7 and 8 of *Where the Conflict Really Lies* to a discussion of two arguments from design that have received a fair bit of attention in recent years: Chapter 7 scrutinizes the cosmological fine-tuning argument and the first section of Chapter 8 is about Michael Behe's biological argument from irreducible complexity (Behe 1996). This is not the occasion to review these arguments and the discussion about them.<sup>1</sup> What is surprising about Plantinga's take on these arguments is that he ultimately finds them wanting. Although he discusses them sympathetically, he is led eventually to the sober conclusion that "the [fine-tuning argument] offers some slight support for theism" (Plantinga 2011, 224) and that Behe's biological design argument "is by no means airtight" (ibid., 231). As far as I can tell, this is something of a departure from an earlier point of view. In the widely circulated lecture notes on 'Two Dozen (or So) Theistic Arguments',<sup>2</sup> for instance, the fine-tuning argument occupies a prominent place in the list, as does a slightly different design argument from the general order of the universe and the laws of nature.

More interesting than the question whether or not Plantinga has changed his mind about the force of design arguments is his novel take on them. Rather than thinking of design arguments as explicit arguments with distinct premises, a conclusion, and an explicit logical pattern connecting the two, we can think about the conclusions of design arguments as beliefs that are formed directly in us.<sup>3</sup> People acquire such beliefs not through an inference with explicit discursive steps that has as its conclusion that something is the product of design, but as beliefs that force themselves upon their minds immediately, without the mediation of an explicit argument or inference pattern. His proposal is that, in this respect, design beliefs are analogous to perceptual beliefs, memory beliefs, or beliefs about what other people think, feel, or want. When we see a tree outside our window, we don't run a quick inference or argument in our heads to the effect that, say, the visual imagery we are currently experiencing needs explanation and that the best explanation for it is the fact that there is an external world that contains a tree which causes our visual experience. No, we simply form the belief that there is a tree in a basic way, i.e., unmediated by any discursive steps. Plantinga's proposal, then, is to consider design beliefs as being formed in a similar way:

The belief that something or other is the product of design is not formed by way of inference, but in the basic way; what goes on here is to be understood as more like perception than like inference. (2011, 245)

This is not to say that design beliefs pop up into our heads at random without any sort of external prompting. Rather, they arise in certain circumstances in response to observations or other experiences. Naturally, an encounter with a

---

<sup>1</sup> In addition to Plantinga's chapters, interested readers may want to consult Pennock (2001), Manson (2003), Dembski & Ruse (2004), Collins (2009), and Monton (2009).

<sup>2</sup> Also available as the appendix to Baker (2007).

<sup>3</sup> Cf. also Ratzsch (2003), whom Plantinga credits for inspiration.

familiar human artifact, as in William Paley's well-worn example of a watch found on a heath, would constitute circumstances in which forming a design belief is the appropriate response, but, says Plantinga, so could encounters with elements of the natural world, such as the human eye, the bacterial flagellum, or the fine-tuning of the universe. On this suggestion, authors such as William Paley and Michael Behe who appear to be proposing design arguments can be interpreted differently. Rather than developing explicit arguments, they can be seen as engaging in *design discourse*: they attempt to describe features of the natural world, which, when appreciated by observers with properly functioning cognitive faculties, ought to elicit design beliefs.

Suppose this is right and we indeed sometimes form design beliefs in the basic way. A pertinent question then is whether it is plausible that design beliefs thus formed can have positive epistemic status, i.e., whether they can be justified, have warrant, or even amount to knowledge. Here Plantinga draws on his earlier work (Plantinga 1993; 2000) and suggests that basic design beliefs can indeed enjoy plenty of warrant and even constitute knowledge, provided they are produced by properly functioning cognitive mechanisms that are used in the environment for which humans were designed (by God or evolution) according to a design plan successfully aimed at the production of true beliefs. Again, in this respect they are on a par with perceptual beliefs, memory beliefs, and beliefs about other minds. Plantinga doesn't argue that we indeed possess the relevant cognitive mechanisms for perceiving design, but leaves it at the conditional claim that if we do have such mechanisms, then basic design beliefs can have positive epistemic status.

This proposal for construing design arguments as design discourse makes a real difference. First, Plantinga thinks that design arguments aren't very powerful. At best, they offer some slight support for theism. Hence, when a design belief is formed as the conclusion of a design argument, it will have little by way of positive epistemic status. As we just saw, however, a design belief that is formed in the basic way can have much more positive epistemic status and can even constitute knowledge.

Secondly, basic beliefs are criticized in a different way than beliefs formed on the basis of arguments. To criticize a belief of the latter sort, one has three options: (a) attack the premises of the argument by showing that they are false or that we lack reason to accept them, (b) attack the validity of the argument by showing that the conclusion doesn't follow logically from the premises, or (c) attack the conclusion directly by arguing that we have independent reasons to think it is false, without showing that there is anything wrong with the specific argument given for the conclusion. If one wants to criticize basic beliefs, however, options (a) and (b) are not available for obvious reasons. Instead, basic beliefs are undermined by defeaters. A defeater for a belief is an experience or a belief one comes to have which takes away the positive epistemic status of one's initial belief.<sup>4</sup> Defeaters come in the following two kinds: rebutting and undercutting. A rebutting defeater for a basic design

---

<sup>4</sup> See Bergmann (2004) and Plantinga (2011, 164–167) for this characterization of defeaters.

belief would be a belief to the effect that the thing one thought to be designed is in fact not designed. An undercutting defeater for a design belief would be a belief to the effect that the cognitive mechanism by which one formed one's design belief was not reliable.

This second difference opens up a novel way of thinking about objections to design arguments. All those objections that seek to attack the premises of design arguments or that attempt to show that the logical form of a design argument is deficient suddenly become irrelevant, since design beliefs formed in the basic way do not depend on explicit premises and a specific logical form. Instead, we ought to consider whether objections to design arguments amount to rebutting or undercutting defeaters. This, then, is what I will do in what follows.

### 3. *Evolutionary Theory as a Potential Defeater*

The first potential defeater comes from standard neo-Darwinian evolutionary theory itself. As Plantinga rightly observes, it is often suggested that Darwin's theory of evolution by natural selection rendered belief in supernatural design obsolete (Plantinga 2011, 252–254). Evolution is, to use Richard Dawkins' famous metaphor, a blind watchmaker. It can account for all the appearance of design in the natural living world without invoking an actual intelligent designer.<sup>5</sup> If this is correct, then evolutionary theory would seem to constitute a rebutting defeater for design beliefs because it entails that, contrary to appearances, those features of the living world that may give rise to design beliefs are in fact not designed. Hence, design beliefs are undermined by a rebutting defeater.

This line of thought is too quick. Evolutionary theory only constitutes a rebutting defeater if the claim that something evolved in the manner proposed by standard evolutionary theory entails the falsity of a design belief — i.e., if it entails that that thing is not designed. This is not the case. For, as Plantinga has shown in Chapter 1 and as others before him have stressed, it is no part of standard evolutionary theory that evolution is unguided.<sup>6</sup> To claim otherwise is to put a metaphysical spin on evolution. Although mutations are often said to be random, this means nothing more or less than that there is no physical mechanism of any kind connecting mutations with their benefits for the survival of the organism in which they occur.<sup>7</sup> Mutations being random in this sense is wholly compatible with the idea that some mutations are intended and caused by God so that the resulting features of the living world are designed after all.

Next, Plantinga discusses the possibility that evolutionary theory provides an undercutting defeater by showing “that it is *possible* that the structures and traits

---

<sup>5</sup> Dawkins (1986) and Dennett (1995) famously develop captivating narratives along these lines.

<sup>6</sup> Here, Plantinga's position is similar to that of the proponents of 'theistic evolution', such as Collins (2006), Ayala (2007), and Alexander (2008).

<sup>7</sup> Cf. Sober (2011) for an extended argument to this effect.

in question have come to be by way of unguided evolution” (2011, 254). Given the above characterization of what an undercutting defeater is — to wit, a belief or experience one comes to have which casts doubt on the reliability of the mechanism by which one’s original belief was formed — I wasn’t entirely sure what to make of this. Even if evolutionary science has established that it is possible that seemingly designed features of the world came about through unguided evolution, it doesn’t follow that the mechanism by which we form design beliefs is unreliable. The mere fact that it is possible that some feature evolved unguidedly has no tendency of showing that it did so evolve. But even if it did show this, this would constitute a rebutting rather than an undercutting defeater for design beliefs.

Perhaps, then, we should understand Plantinga’s suggestion as follows. Evolutionary science shows (1) that it is possible that every seemingly designed feature of the world came about by means of unguided evolution and (2) that, if this possibility were actual, we would still form design beliefs about many features of the natural world. If this is correct, one might take it to show that the cognitive mechanism responsible for design beliefs is unreliable.

But this isn’t right either. Firstly, the assumption that is really driving the idea that our cognitive mechanisms are unreliable is the second one about what our beliefs would have been under the assumption of unguided evolution. Regardless of whether or not this assumption is correct, it certainly doesn’t follow in any straightforward way from evolutionary science. Hence, if we have an undercutting defeater on our hands here, it doesn’t come from standard evolutionary theory. Secondly, much depends on how the reliability of cognitive mechanisms is to be analyzed. Epistemologists have debated this at great length, but for now one brief consideration will have to suffice. There is good reason to think that the reliability of a cognitive mechanism should not be analyzed in terms of what its outputs would be in possible worlds that are very different from ours, since this quickly leads to the problematic conclusion that none of our actual cognitive mechanisms are reliable. Supposing that our perceptual beliefs would be much the same in possible worlds in which a Cartesian demon deceives us, it would follow that our actual perceptual cognitive mechanisms are unreliable because in such worlds most of our perceptual beliefs would be false. This is clearly the wrong result and such analyses of reliability should therefore be avoided. So, the mere fact (if indeed it is one) that we would still form design beliefs in worlds very different from our world — namely, worlds in which unguided evolution happened — cannot be taken to show that the cognitive mechanisms by which we form design beliefs in the actual world are unreliable. Perhaps tinkering with possible definitions of reliability can circumvent this problem, but since we will have occasion to explore this in more detail in the next section, I want to leave it at this for now.

The conclusion to draw at this point is that (*pace* Plantinga) standard evolutionary theory does not provide us with an undercutting defeater for

design beliefs.<sup>8</sup> In so far as it shows that it is at least possible that all features of the living world could have come about through a process of unguided evolution, it shows that our design beliefs may be false. This, however, is at best a partial defeater of the rebutting kind. Its exact strength depends on whether a compelling case can be made that it is not merely possible that unguided evolution has produced all apparent design, but that this is not prohibitively improbable, and perhaps even somewhat plausible. Someone who seeks to make such a case, however, runs into daunting difficulties, as Plantinga rightly observes (2011, 255–256). Not only is detailed knowledge of complete paths through evolutionary history leading from primitive life forms all the way to highly complex features, such as the human eye or the bacterial flagellum, simply unavailable to us at present. But even if it were available, whether evolution really was unguided would still be a question that empirical science cannot answer.<sup>9</sup>

#### 4. *The Cognitive Science of Design Beliefs*

In this section, I want to look at another way in which evolutionary science, or at least evolution-inspired science, could lead to undercutting defeaters for design beliefs. This is a topic that Plantinga does not consider but which has been getting a lot of press lately.

Developmental and experimental psychologists try to shed light on the human cognitive apparatus. Sometimes, their work is supplemented by that of evolutionary psychologists who attempt to explain various structural and near-universal features of human cognition by looking at our evolutionary history. In this section, I will review some results from these fields that are particularly relevant to design beliefs.<sup>10</sup> In the next sections I will explore how this work might be thought to provide an undercutting defeater for design beliefs.

Studies with young (preschool) children show that they are likely to consider various types of natural objects, such as plants, animals, rocks, the earth, or the sky, as being made by God, when given the choice between saying that these objects were made by humans, that they were made by God, or that no one knows (Petrovich 1999). Other studies are in line with this. Margaret Evans (2001) asked American children in the age groups five to seven years old and eight to ten years old from both Christian and non-Christian backgrounds to rate their agreement with the following origin accounts: creation, evolution, artificialism (the belief that things were made by humans), or emergentism (the belief that things just appeared). The creation account was the most popular by far.

---

<sup>8</sup> The Evolutionary Argument Against Naturalism, which Plantinga discusses in Chapter 10, does imply that someone who accepts naturalism and evolutionary theory thereby obtains an undercutting defeater for all of her beliefs, but that is not what is at issue here.

<sup>9</sup> See also Van Woudenberg & De Ridder (ms) for an argument that design hypotheses are exceedingly difficult if not impossible to disconfirm by empirical science.

<sup>10</sup> I rely on Barrett (2004) and De Cruz & De Smedt (2010) for this overview.

Such findings dovetail very nicely with Deborah Kelemen's work (2003; 2004), which suggests that thinking in terms of purpose comes naturally to both children and adults. Humans have a propensity to think that traits of living beings have some functional role to fulfill. They exist because they are beneficial to either the organism itself or to other organisms. While most adults limit teleological thinking to specific traits of living organisms, young children also have a propensity to ascribe purposes to entire objects and organisms. When asked what lions are for, they answer things like 'To go in the zoo'. Similarly, they will say that clouds are for raining and that pointy rocks have their shape to prevent people from sitting on them or to allow animals to scratch themselves with them. Young children are promiscuous teleologists, says Kelemen. The tendency towards teleological reasoning lessens around the age of ten to twelve, presumably because formal education provides children with compelling non-teleological explanations for many phenomena. Nonetheless, it never quite disappears. Educated adults, when forced to make speedy decisions, show a heightened acceptance of teleological explanations (Kelemen & Rosset 2009). This remains true even for highly educated people, such as physicists working at research universities (Kelemen et al. 2013). Similarly, patients with Alzheimer's disease also display a re-emerging preference for teleological explanations (Lombrozo et al. 2007).

Such findings are significant in relation to design beliefs in the following ways. First, there are conceptual connections between teleology, intelligent agency, and design. Any object that has been designed has one or more purposes for which it was designed and thus displays teleology. Teleology is an indicator of design. Being created by an intelligent agent is also closely conceptually linked with design. Not everything that intelligent agents create is designed, because they sometimes create things as unintentional byproducts of their activities (e.g., footsteps in the snow or saw dust). But many things that intelligent agents create are indeed intentionally designed. Being created by an intelligent agent is thus another strong indicator for design.

Secondly, these conceptual connections are borne out by empirical studies. When young children have been told that stains on a piece of paper have been created intentionally by someone, they are much more likely to identify them as representations of something rather than just stains (Gelman & Ebeling 1998). In another experiment, adult test subjects were presented with an object which looked like it could either be an artifact or something that was the product of chance. They were then told two different accounts of how the object was brought about. Subjects who had been exposed to the account that someone had intentionally modified the features of the object to make it the way it was identified it as an artifact, i.e., a product of design (Gelman & Bloom 2000). Similarly, Newman et al. (2010) found that young children and even babies expect that only agents, and not inanimate processes, can create order. Hence, the information that something has been created by an intelligent agent strongly predisposes people to think of it as a product of design.

Taken together, the above results imply that people are predisposed to form design beliefs about the natural world. They have an inclination to think of

natural objects as being created by an intelligent agent and they naturally ascribe purposes promiscuously to natural objects. Both of these features — being created by intelligent agency and teleology — are indicators of design. The step to design beliefs is very small and natural indeed. Hence, we might say people are natural born design believers. Of course, in normally functioning adults, such tendencies are curbed in various ways by education and conscious reflection, but the empirical findings do suggest that the natural inclination to form design beliefs doesn't quite go away.

##### 5. *Undercutting Design Beliefs I: Track-Record Unreliability*

These are the facts — according to recent psychology at least. How do we get from them to undercutting defeaters for design beliefs? The suggestion must be that the psychological evidence establishes that the cognitive mechanisms by which we form design beliefs — let's call them design-belief-forming mechanisms or DBMs for short — are somehow lacking in an epistemically important quality. There are different ways to make this more precise.<sup>11</sup>

The first way to put a spin on the evidence is to take it to show that our DBMs are unreliable. They produce too many false positives, i.e., beliefs that something or other has been designed when this is not the case. A belief that the mechanism by which a certain class of beliefs is formed is unreliable constitutes an undercutting defeater for those beliefs. If I believe that my eyesight is severely impaired, I cannot continue to hold on to visual beliefs that I have formed through the use of my own eyes. By analogy, I should give up my design beliefs (or at least hold them less confidently) once I learn that my DBMs are unreliable.

Straightforward as it is, the suggestion that the psychological evidence shows that design beliefs are unreliably formed is hard to maintain. What the evidence actually tells us is that young children and patients with Alzheimer's disease have a strong tendency to form design beliefs about items in the natural world and that adults — even highly educated ones — who are forced to make hasty judgements tend to form design beliefs. Even if we grant that the beliefs formed in these groups indeed include many false positives (more on that in due course), however, this nowhere near shows that the DBMs of educated adults who are functioning in normal circumstances are unreliable. For these mechanisms include — in addition to what is perhaps an overly enthusiastic tendency to form design beliefs stemming from deep down in our primitive animal brains — various correctives to this tendency that are put in place by critical reflection and formal education. This is exactly what the psychological evidence confirms where it shows that older children are less prone to form design beliefs about items in the natural world. To make a general case for the unreliability of DBMs, the evidence would have to show

---

<sup>11</sup> Inspiration for the following discussion is drawn from Murray (2009) and Thurow (2013) where analogous debunking arguments for religious beliefs are scrutinized.



that educated adults in normal circumstances regularly form mistaken design beliefs. Such evidence has not been produced.

Another problem with the claim that our DBMs are unreliable is that it is difficult to support without begging the question. To say that these mechanisms produce too many false positives is to assume that design beliefs about various or even all items in the natural world are actually false. But for Plantinga and others who want to defend the possibility that design beliefs sometimes amount to knowledge, this is exactly what is at issue. They think that there are items in the natural world that have been designed by God, either by means of a process of guided evolution or directly through special divine action. Design beliefs formed about those items would actually be true beliefs rather than false positives. In saying that our DBMs are unreliable, the objector is already assuming that there is no design in the natural world, which was the very point of contention. In a discussion about the legitimacy of design discourse, the total absence of design in the natural world cannot be assumed from the outset.

Observe, furthermore, that pro-design authors like Paley and Behe are not proposing that we form design beliefs indiscriminately about any and all organs or body parts. They try to describe specific qualities — e.g., sophisticated functional integration or irreducible complexity — that only some traits and structures possess. They license forming design beliefs only in these special cases. Without these strictures it might have been easier to make the charge of unreliability stick, because there are many items for which we do have strong reasons to suppose that they could have evolved without guidance or intervention. But since the defenders of design discourse do not propose that design beliefs are appropriate across the board, the point made in the previous paragraph stands: to argue that our DBMs are unreliable when operating on the special class of items that defenders of design discourse attempt to delineate is impossible without begging the question.

#### 6. *Undercutting Design Beliefs II: Modal Unreliability*

Let's turn to a second interpretation of the psychological evidence which would — if successful — show it to be an undercutting defeater for design beliefs. The evidence suggests that the tendency to form design beliefs is a cross-cultural universal feature of the human cognitive make-up. Surely, then, it must have been evolutionarily advantageous or a byproduct of something that was evolutionarily advantageous. This demonstrates that the truth of design beliefs doesn't enter into the explanation of why we form them. And this is bad from an epistemic point of view because in order for a belief to amount to knowledge, it must be the case that the truth of the belief in question enters into the explanation of why we hold it. To put it another way: if beliefs are to constitute knowledge, they must track the truth. Design beliefs fail this requirement, because no matter whether they are true or false, we still would have held them because our doing so contributed to survival (or piggybacks on something that did). This kind of modal unreliability is an

epistemic defect of DBMs and constitutes an undercutting defeater for design beliefs, or so the objection goes.

In response, it must be observed that this line of reasoning assumes that humans would have evolved DBMs and thus the tendency to form design beliefs even if a naturalistic account of the origin and evolution of humans were correct. This, of course, is an assumption that the defender of design discourse is not likely to buy into without further ado. Classical theists like Plantinga hold that there wouldn't even have been a universe, let alone evolved intelligent life, were it not for God's creating and sustaining activity. Moreover, proponents of design discourse are also unlikely to grant the more specific assumption that unguided evolution will lead to anything like intelligent beings such as humans. So someone who wants to employ the above line of reasoning to show that there is an undercutting defeater for design beliefs faces the burden of arguing that unguided evolution could produce human beings and, even worse, the burden of arguing that a naturalistic account of the origins of the universe is plausible. Such claims are typically taken for granted by staunch evolutionists and naturalists, but it should be clear that assuming them in the current discussion about the epistemic status of design beliefs begs the question.

But perhaps there is a way around this response. I claimed that the objection must assume that we would have had DBMs and would have formed design beliefs even if a naturalistic account of human origins had been correct. But, the objector might push back, she doesn't really need to assume something this strong. To make the modal unreliability charge plausible, all that is required is for us to evaluate what we would have believed if there had been no design and we would still have existed and used DBMs to form beliefs, regardless of whether such a state of affairs is possible or plausible. DBMs being what they are — mechanisms for forming design beliefs — it seems clear that in such a situation we would still have formed design beliefs. Thus, so the modified objection goes, DBMs are unreliable in an epistemically problematic way and beliefs formed through them are defeated.

Unfortunately for the objector, however, there is a deeper problem with the proposed test for reliability, at least in so far as it is supposed to track an epistemically relevant defect of a belief-forming mechanism. The problem is that the test cannot be correct in general, because it is subject to counterexamples: cases of beliefs that amount to knowledge which are formed by a belief-forming mechanism that fails the proposed test. To see this, it will be useful to have the test before us more clearly. Suppose that a belief that  $p$  is produced by a belief-forming mechanism (BM). Put in general terms, the test claims that, in order for our true belief that  $p$  to amount to knowledge, the following counterfactual has to be true of it:

(T) If  $p$  had been false and we were to form a belief about whether or not  $p$  using BM, we would not have believed that  $p$  is true.<sup>12</sup>

---

<sup>12</sup> Those familiar with contemporary epistemology will recognize (T) as a version of Nozick's sensitivity principle (Nozick 1981, Chapter 3).

Consider the following case.<sup>13</sup> Suppose I bike to the train station every morning and leave my bike in a bike rack at the station. I've been doing so for years, I always double-check that my bike is locked, and the train station is in a safe part of town with virtually no chance of bicycle theft. I take the train to work and, while idling on the train, I form the belief that my bike is where I left it, in a bike rack at the train station. It should be uncontroversial that I know my bike is in the rack.<sup>14</sup> Nonetheless, my belief fails (T). For if, against all odds, someone took my bike (thus rendering my belief false), I would still have formed the same belief using the same belief-forming mechanism.

So (T) fails as a general test: the fact that beliefs formed by a BM fail (T) does not entail that they cannot be cases of knowledge. The modal unreliability captured by (T) is not a defect that necessarily undermines the positive epistemic status of a belief. Hence, even if it is true of design beliefs formed by DBMs that they fail (T), this does not automatically show that they are confronted with an undercutting defeater.

### 7. *Undercutting Design Beliefs III: Restricted Modal Unreliability*

Fair enough, the design skeptic might retort. The test is not correct when it is taken as fully general. There are exceptions: some true beliefs formed by some BMs fail (T) and are cases of knowledge nonetheless. But it doesn't follow from this that a belief's failing (T) is *never* an epistemically relevant defect. Perhaps the test is fine when it is applied to a more restricted domain of beliefs or BMs. The test could get at an epistemically relevant defect for some types of beliefs or some types of BMs. For instance, suppose I suffer from face blindness. I form the belief that I see Anna across the room from me. Surely, the fact that I still would have formed this belief if Bernadette, Chris, or Diana had stood across the room prevents me from knowing that I see Anna.

In view of this insight, the objector might try to support the contention that failing (T) is a knowledge-undermining epistemic defect for the case of design beliefs by first offering a fully fleshed-out account of the distinction between cases in which failing (T) is and is not an epistemic defect and then arguing that design beliefs fall on the bad side of this distinction. Unfortunately, I do not see what such an account could look like, nor am I aware of any proposals in the literature. Instead, then, the objector might attempt a more modest strategy. She might try to support her claim that failing (T) leads to an undercutting defeater for beliefs produced by DBMs by presenting analogous cases and arguing that our verdict in those analogous cases would be that there is an undercutting defeater. It is this strategy that I will pursue here to develop a final modified version of the unreliability objection.

<sup>13</sup> The case is structurally equivalent to a case discussed in Sosa (1999).

<sup>14</sup> Anyone tempted to deny knowledge here should realize that doing so in effect denies the possibility of inductive knowledge in general. This is an unacceptably radical form of skepticism.

Suppose I have an app on my iPad that tells me the location of the planets in our solar system.<sup>15</sup> It does so because it has been programmed with equations that capture the laws of planetary motion and correct information about the initial locations of the planets at some point in the past. Suppose I form the belief that Mars will be in the Libra constellation tonight by consulting the app. This belief fails (T). Had Mars been at a different location and had I formed a belief about its location using my iPad, I still would have believed it to be in the Libra constellation, since Mars changing its location does not directly affect the way the app is programmed. In this case, however, it seems clear that I do know the location of Mars, since the app is in fact reliable.

Now consider a variant on this case. I don't have an iPad but I stumble upon a device that tells me the locations of the planets. It gives the exact same information for the location of Mars on said night. I examine the device carefully and note that it doesn't seem to communicate with any outside sources of information and doesn't have any sensors with which it could track planetary motions. This gives me reason to think that beliefs formed by looking at this device would fail (T): if Mars had been at a different location, the device would still have led me to form the belief that it is in the Libra constellation. Yet in this case, it would seem that I do not know the location of Mars, even though the device may well be perfectly reliable.

The relevant difference is that in the first case I possess independent background information that speaks to the reliability of the app. I know that the people who programmed the app have access to reliable information about the locations of the planets and their courses of motion. I know that, if something had been amiss with the app, other users would have reported this and an updated version would have soon come out. No such information is available to me in the second case, at least not when I first find the device. I may acquire confirming evidence for the device's reliability later by comparing it with other sources of information about the locations of the planets or perhaps even by taking it apart and verifying that it has been built and programmed in a reliable way. But as long as I don't have such information and I do have reason to think beliefs formed by looking at the device satisfy (T), those beliefs have an undercutting defeater.

Here, then, is the objection in its final modified form. Unfortunately for the proponent of design beliefs, beliefs formed by DBMs are like those in the found device case, since we don't have any independent means of verifying their reliability. We have neither background knowledge from which we can infer the reliability of DBMs, nor do we have access to how DBMs work — i.e., we don't know exactly which inputs produce design beliefs as outputs or how inputs are converted into outputs. Moreover, we are not in a position to collect confirming evidence about their reliability because we have no independent checks for design beliefs. DBMs are all we have to go on, since, as Plantinga argued, design arguments provide only weak support for design beliefs. So finally we see that design beliefs are confronted with an undercutting defeater,

---

<sup>15</sup> The cases are inspired by Thurow (2013, 88).

because they satisfy (T) and we lack independent reason to think that they are reliable.

Note that endorsing this objection does not commit the objector to the implausible claim that we first need positive evidence for the reliability of any BM before we can trust its outputs as constituting knowledge. Such a claim would quickly lead to widespread skepticism, since we lack such evidence for all of our basic BMs.<sup>16</sup> The point is more restricted: if we have reason to think that the beliefs produced by a BM fail (T) and we lack positive evidence for the reliability of this BM, then beliefs produced by this BM are faced with an undercutting defeater and do not amount to knowledge. Since we have excellent reason to think that most beliefs produced by mechanisms such as sense perception, memory, expert testimony, and rational intuition satisfy (T), the objector can grant that there is nothing wrong with trusting these BMs to produce knowledge, even if we have no independent evidence for their reliability.<sup>17</sup>

Is there anything the friend of design discourse can say against this lengthy and sophisticated charge? I want to consider two possible responses. She might submit that all this talk about modal unreliability and (T) in particular has taken us down a wrong path. Modal conditions on knowledge are mistaken. Inspiration for this response may be drawn from Plantinga's own account of knowledge (1993; 2000). On this account, a true belief amounts to knowledge provided it has enough warrant. This, in turn, is analyzed as follows:

[A] belief has warrant for a person *S* only if that belief is produced in *S* by cognitive faculties functioning properly (subject to no dysfunction) in a cognitive environment that is appropriate for *S*'s kind of cognitive faculties, according to a design plan that is successfully aimed at truth. (Plantinga 2000: 156)

It is plain to see that nothing like (T) or any other modal clause figures in this analysis. Applied to design beliefs, this response would thus insist that they can in fact amount to knowledge, provided that they are produced by properly functioning DBMs in the appropriate environment according to a successfully truth-aimed design plan. Nothing in what's been said so far about the cognitive science of design beliefs suggests that they fail to meet these conditions. The fact that they don't satisfy (T) is simply irrelevant to the issue of whether they can constitute knowledge.

This hardline reply is unattractive for several reasons. Although Plantinga's analysis of warrant contains no explicitly modal clauses, modal considerations are nonetheless important to it. In his discussion of Gettier cases, to mention just one example, Plantinga wholeheartedly agrees that certain types of modal instability defeat a belief's warrant (2000, 156–161). Furthermore, accounts of knowledge that use explicitly modal conditions have become increasingly popular in recent epistemology because of their success in dealing with various

<sup>16</sup> As Bergmann (2004) argues. Cf. also Alston (1993).

<sup>17</sup> Note that this is so for most beliefs produced by these mechanisms. There are exceptions. My belief that I am not a brain in a vat presumably fails (T) since in worlds in which that belief is false (and I am a brain in a vat) I still would have believed that I am not one.

problems for older analyses of knowledge. This should count as at least some evidence that they are on the right track. Modal analyses of knowledge have seemed to many to offer an extremely natural way of explicating the almost universally shared intuition that knowledge and certain kinds of luck are incompatible (Unger 1968; Pritchard 2005).<sup>18</sup> Because of these considerations, a wholesale rejection of modal conditions on knowledge is unpromising. If we add to this the intuitive plausibility of (T) as a test for many beliefs — as testified to by the cases discussed above — the sensible thing to say is that this response is not the way to go.

A second response is to contest the claim that there is, or can be, no independent confirmation of the reliability of DBMs. As we noted before, we don't have direct independent access to facts about design in the natural world. But that doesn't mean that there could be no indirect support for design beliefs. If there is independent reason to think that God exists, that will offer indirect support to the reliability of DBMs. God is, after all, the 'first cause' of all that exists or — to put it less philosophically — the creator of heaven and earth. The existence of God thus makes it vastly more probable that there is indeed design in the natural world and that some design beliefs therefore are indeed true.

As is well known, Plantinga (2000) offers a sustained defense of the claim that if God exists, belief in God can be properly basic and indeed amount to knowledge. For this to be the case, belief in God has to be produced by a properly functioning cognitive faculty for religious knowledge — something like Calvin's *sensus divinitatis* — and satisfy the other conditions for knowledge laid out by Plantinga's general account of knowledge, which we already encountered above.

Here, we can take our cue from something Plantinga himself writes at the end of his discussion about design discourse:

If S is already a theist, S believes that these things (and indeed the whole universe) is designed. Under those conditions, Darwinian considerations will not give S a defeater for the design belief in question; her theistic belief is a defeater-deflector for the looming defeater. (2011, 261)

Although the "Darwinian considerations" he refers to here aren't the same as the results from cognitive science I presented above (rather, they're the considerations discussed in section 3), a similar point applies. For someone who is already a theist, the modal unreliability of DBMs is not a problem, since she does have independent reason to think that DBMs are (somewhat) reliable. The theist will thus insist that the iPad case discussed above — rather than the found device case — is the proper comparison for the situation with DBMs and that, therefore, it can be perfectly rational to trust the outputs of DBMs.

One might deem this pretty piddling, since design arguments have often been offered to convince atheists and agnostics of the plausibility of theism. To

---

<sup>18</sup> A sample of influential recent accounts of knowledge with prominent modal conditions includes Nozick (1981), Sosa (1999; 2007), Williamson (2000), Roush (2005), Becker (2007), and Pritchard (2005; 2012).

then be told that design arguments ought to be recast as design discourses and that the latter are only successful for those who are already theists will not set hearts racing. In reply, besides pointing out that the sober truth isn't always nearly as exciting as one would have hoped, I want to echo a further suggestion from Plantinga. Perhaps some situations to which design beliefs are the rational response are also situations in which the *sensus divinitatis* is activated (2011, 263). If that is correct, then design discourse might yet retain some of the general appeal of design arguments. Design discourse could first prompt theistic belief and then, in its wake, design beliefs. Because those design beliefs are now embedded in a context of theistic belief, they remain undefeated.

#### 8. *Concluding Remarks*

I have presented Plantinga's proposal to reinterpret arguments from design as design discourse, i.e., as means to trigger a basic belief-forming mechanism that gives rise to design beliefs, which may even amount to knowledge. One important worry for this proposal is that such design beliefs face defeaters. In particular, it might be thought that insights from cognitive science on how people form design beliefs and evolutionary explanations of these tendencies spell trouble for the idea that such beliefs may amount to knowledge. This is, after all, a common suggestion for the cases of moral beliefs and religious beliefs. For such beliefs, too, people have attempted to debunk the positive epistemic status of such beliefs by pointing out that there are evolutionary explanations for why we tend to have certain moral and/or religious beliefs. I have looked at various possible strategies to put a spin on the evidence from cognitive science that would debunk design beliefs. The upshot was that none of these strategies are ultimately successful. This doesn't show that design discourse is successful or reliable, but it does underline one of Plantinga's central theses, to wit, that there is only superficial conflict between science and religion.<sup>19</sup>

#### *References*

- Alston, William P. (1993), *The Reliability of Sense Perception*, Ithaca, NY: Cornell University Press.
- Ayala, Francisco (2008), *Darwin's Gift: To Science and Religion*, Washington, DC: Joseph Henry Press.
- Alexander, Denis (2009), *Creation or Evolution: Do We Have to Choose?*, Oxford: Monarch.

---

<sup>19</sup> Thanks to members of the Abraham Kuyper Center for Science and Religion at VU University and to Jan Willem Wieland for feedback on an earlier version of this paper. Part of the work for this paper was made possible by a grant from the Templeton World Charity Foundation. The opinions expressed in this publication are those of the author, and do not necessarily reflect the views of the Foundation.

- Baker, Deane Peter (ed.) (2007), *Alvin Plantinga*, Cambridge: Cambridge University Press.
- Barrett, Justin (2004), *Why Would Anyone Believe in God?*, Lanham, MD: AltaMira Press.
- Becker, Kelly (2007), *Epistemology Modalized*, London: Routledge.
- Behe, Michael (1996), *Darwin's Black Box*, New York: Free Press.
- Bergmann, Michael (2004), 'Epistemic Circularity: Malignant and Benign', *Philosophy and Phenomenological Research* 69(3), 708–726.
- Bergmann, Michael (2005), 'Defeaters and Higher-Level Requirements', *Philosophical Quarterly* 55(220), 419–436.
- Collins, Francis (2006), *The Language of God*, New York: Free Press.
- Collins, Robin (2009), 'The Teleological Argument: An Exploration of the Fine-Tuning of the Universe', in *The Blackwell Companion to Natural Theology*, edited by William Lane Craig and J.P. Moreland, Cambridge, MA: Wiley-Blackwell, 202–281.
- Dawkins, Richard (1986), *The Blind Watchmaker*, London: Longman.
- De Cruz, Helen, and Johan De Smedt (2010), 'Paley's iPod: The Cognitive Basis of the Design Argument Within Natural Theology', *Zygon* 45(3), 665–684.
- Dembski, William and Michael Ruse (eds.) (2001), *Debating Design*, Cambridge: Cambridge University Press.
- Dennett, Daniel (1995), *Darwin's Dangerous Idea*, London: Penguin Press.
- Evans, E. Margaret (2001), 'Cognitive and Contextual Factors in the Emergence of Diverse Belief Systems: Creation versus Evolution', *Cognitive Psychology* 42, 217–266.
- Gelman, Susan A. and Karen S. Ebeling (1998), 'Shape and Representational Status in Children's Early Naming', *Cognition* 66, 35–47.
- Gelman, Susan A. and Paul Bloom (2000), 'Young Children Are Sensitive to How an Object Was Created When Deciding What to Name It', *Cognition* 76, 91–103.
- Kelemen, Deborah (2003), 'British and American Children's Preferences for Teleo-Functional Explanations of the Natural World', *Cognition* 88, 201–221.
- Kelemen, Deborah (2004), 'Are Children 'Intuitive Theists'? Reasoning About Purpose and Design in Nature', *Psychological Science* 15, 295–301.
- Kelemen, Deborah, and Evelyn Rosset (2009), 'The Human Function Compunction: Teleological Explanation in Adults', *Cognition* 11, 138–143.
- Kelemen, Deborah, Joshua Rottman, and Rebecca Seston (2013), 'Professional Physical Scientists Display Tenacious Teleological Tendencies: Purpose-based Reasoning as a Cognitive Default', *Journal of Experimental Psychology* 142(4), 1074–1083.
- Lombrozo, Tania, Deborah Kelemen, and Deborah Zaitchik (2007), 'Inferring Design: Evidence of a Preference for Teleological Explanations in Patients with Alzheimer's Disease', *Psychological Science* 18, 999–1006.
- Manson, Neil (ed.) (2003), *God and Design: The Teleological Argument and Modern Science*, London: Routledge.



- Monton, Bradley (2009), *Seeking God in Science: An Atheist Defends Intelligent Design*, Boulder, CO: Broadview Press.
- Murray, Michael (2009), 'Scientific Explanations of Religion and the Justification of Religious Belief', in *The Believing Primate*, edited by Jeffrey Schloss and Michael Murray, New York: Oxford University Press, 168–178.
- Newman, George E., Frank C. Keil, Valerie Kuhlmeier and Karen Wynn (2010), 'Sensitivity to Design: Early Understandings of the Link Between Agents and Order', *Proceedings of the National Academy of Sciences* 107, 17140–17145
- Nozick, Robert (1981), *Philosophical Explanations*, Cambridge, MA: Harvard University Press.
- Pennock, Robert T. (ed.) (2001), *Intelligent Design Creationism and Its Critics: Philosophical, Theological, and Scientific Perspectives*, Cambridge, MA: MIT Press.
- Petrovich, Olivera (1999), 'Preschool Children's Understanding of the Dichotomy Between the Natural and the Artificial', *Psychological Reports* 84, 3–27.
- Plantinga, Alvin (1993), *Warrant and Proper Function*, New York: Oxford University Press.
- Plantinga, Alvin (2000), *Warranted Christian Belief*, New York: Oxford University Press.
- Plantinga, Alvin (2011), *Where the Conflict Really Lies: Science, Religion, and Naturalism*, New York: Oxford University Press.
- Pritchard, Duncan (2005), *Epistemic Luck*, Oxford: Oxford University Press.
- Pritchard, Duncan (2012), 'Anti-Luck Virtue Epistemology', *Journal of Philosophy* 109(3), 247–279.
- Ratzsch, Del (2003), 'Perceiving Design', in *God and Design: The Teleological Argument and Modern Science*, edited by Neil Manson, London: Routledge, 124–144.
- Roush, Sherrilyn (2005), *Tracking Truth: Knowledge, Evidence, and Science*, New York: Oxford University Press.
- Sober, Elliot (2011), 'Evolution Without Naturalism', in *Oxford Studies in Philosophy of Religion Vol. 3*, edited by Jon Kvanvig, New York: Oxford University Press, 187–221.
- Sosa, Ernest (1999), 'How to Defeat Opposition to Moore', *Philosophical Perspectives* 13, 141–153.
- Sosa, Ernest (2007), *A Virtue Epistemology: Apt Belief and Reflective Knowledge, Vol. I*, New York: Oxford University Press.
- Thurow, Joshua C. (2013), 'Does Cognitive Science Show Belief in God to Be Irrational? The Epistemic Consequences of the Cognitive Science of Religion', *International Journal for Philosophy of Religion* 74(1), 77–98.
- Unger, Peter (1968), 'An Analysis of Factual Knowledge', *Journal of Philosophy* 65(6), 157–170.
- Woudenberg, René van, and Jeroen de Ridder (2014), 'Why We Can't Know the Denials of Design Hypotheses', Unpublished manuscript.