

Where the Design Argument Goes Wrong: Auxiliary Assumptions and Unification*

Maarten Boudry and Bert Leuridan^{†‡}

Sober has reconstructed the biological design argument in the framework of likelihoodism, purporting to demonstrate that it is defective for intrinsic reasons. We argue that Sober's restriction on the introduction of auxiliary hypotheses is too restrictive, as it commits him to rejecting types of everyday reasoning that are clearly valid. Our account shows that the design argument fails, not because it is intrinsically untestable but because it clashes with the empirical evidence and fails to satisfy certain theoretical desiderata (in particular, unification). Likewise, Sober's critique of the arguments from imperfections and from evil against design is off the mark.

1. Introduction. Who gave the decisive deathblow to the argument from design on the basis of biological complexity? Both philosophers and biologists are divided on this point (Dawkins 1986; Oppy 1996; Sober 2008). Some have claimed that the biological design argument did not falter until Darwin provided a proper naturalistic explanation for adaptive complexity; others maintain that David Hume had already shattered the argument to pieces by sheer logical force several decades earlier (Hume 1779/2007). Elliott Sober has been among the philosophers who maintain that, as Hume was not in a position to offer a serious alternative explanation of

*Received November 2010; revised February 2011.

†To contact the authors, please write to Maarten Boudry: Department of Philosophy and Moral Sciences, Ghent University, Blandijnberg 2, Room 0.128, 9000 Ghent, Belgium; e-mail: maartenboudry@gmail.com.

‡We would like to thank Stefaan Blancke, Johan Braeckman, José Díez, Heather Douglas, Kareem Khalifa, James Lennox, Sebastian Lutz, Elisabeth Nemeth, Laura Perini, Herman Philipse, Peter Vickers, and John Worrall for their helpful criticisms and suggestions. We are also grateful to the anonymous reviewers of *Philosophy of Science*, whose critical comments have substantially improved this article.

Philosophy of Science, 78 (October 2011) pp. 558–578. 0031-8248/2011/7804-0006\$10.00
Copyright 2011 by the Philosophy of Science Association. All rights reserved.

adaptive complexity, it is hardly surprising that “intelligent people strongly favored the design hypothesis” (2000, 36). In his most recent book, however, Sober (2008) carefully develops what he thinks is the most charitable reconstruction of the design argument and proceeds to show why it is defective for intrinsic reasons (for earlier versions of this argument, see Sober 1999, 2002). Sober argues that the design argument can be rejected even without the need to consider alternative explanations for adaptive complexity (2008, 126): “To see why the design argument is defective, there is no need to have a view as to whether Darwin’s theory of evolution is true” (154).

We argue that Sober’s reconstruction, which is based on a discussion of auxiliary assumptions, suffers from an important problem. His requirements regarding the choice of auxiliary hypotheses and his proposed independence relations are overly restrictive, as they commit him to rejecting types of reasoning that are obviously valid. We develop an alternative and more lenient account of auxiliary assumptions, based on the explanatory virtue of unification as a way to avoid gerrymandering. In our view, if only the design argument satisfied certain theoretical requirements, it would be rendered compelling in ways that violate Sober’s restriction concerning the choice of auxiliaries. Our argument is not only relevant for philosophical discussions concerning auxiliaries, gerrymandering, and unification. Compared to Sober’s approach, it also strengthens the case against the design argument. We conclude that the design argument does not suffer from any intrinsic flaws but has simply collapsed under the weight of evidence and has been outcompeted by evolutionary theory, which is all the more damaging to the epistemic status of the design hypothesis. Theoretical immunizations by design theorists and historical examples from natural theology are discussed to support this thesis. An important corollary of our view is that Sober’s objections against the *argument from evil* and the *argument from imperfections*, which have been leveled against the design hypothesis ever since Darwin’s seminal work, are equally misguided.¹

2. Likelihoodism and the Design Argument.

2.1. Reconstructing the Design Argument. In his reconstruction of the design argument, Sober wants to arrive at “the strongest, most defensible, version of the argument” and then to show why he thinks it is “defective” (2008, 113). Sober’s reconstruction has three features we should keep in

1. For a thorough critique of Sober’s likelihood reconstruction of the cosmological design argument, see Weisberg (2005). For Sober’s response to his critics, see Sober (2009).

mind. First, it is probabilistic, not deductive. Second, it is contrastive: he does not want to evaluate the design hypothesis in isolation but only against competing hypotheses (but see sec. 4.1 for Sober's apparent departure from contrastivism). Third, he favors a 'likelihood approach' over a Bayesian approach because he refuses to assign prior probabilities to Darwin's theory of evolution, or to the existence of an intelligent designer, since these merely reflect "a subjective degree of certainty" (121). Sober applies the law of likelihood to Paley (1802), in which Paley pursued the analogy between the human eye and a pocket watch to drive home the design argument. Sober arrives at the following reconstruction, where 'ID' is the hypothesis of intelligent design and 'Chance' is the old Epicurean hypothesis of pure chance: "Observation O favors ID over Chance if and only if $\Pr(O|ID) > \Pr(O|Chance)$ " (2008, 122).

This likelihood reconstruction encounters one immediate objection. The value of $\Pr(O|ID)$ can be artificially raised to unity by tuning the hypothesis to the observations. For example, if "ID++ = there exists an omnipotent supernatural Creator for whom the creation of the bacterial flagellum is number one priority" and "O = there exists a bacterial flagellum," then the likelihood $\Pr(O|ID++)$ equals one. But why not build the observational outcome in the competing hypothesis instead? For example, if "Chance++ = a chance process produced the bacterial flagellum," then $\Pr(O|Chance++)$ likewise is one. As is clear from these examples, the mere fact that the likelihood of some contrived hypothesis equals one does not make it any more plausible.

In Sober's words, "within a likelihood framework, there is no beating a hypothesis that *entails* the observations" (2008, 131). If we allow that favorable (or unfavorable) assumptions are introduced unrestrainedly to the central hypothesis, *in casu* assumptions about the intentions and attributes of the designer, we are left with no way in which an observation O can discriminate between the competing hypotheses. The evidential significance of the observation O "will be thoroughly obscured *if we build the observational outcome into the theories we wish to test*" (132). As a result, the competing hypotheses cannot be tested against each other.

2.2. Restrictions on Auxiliary Hypotheses. If we want to avoid this problem, we somehow have to introduce restrictions on the choice of auxiliary assumptions for our central hypothesis. Sober's proposed solution is to demand "an *independent reason* for believing assumptions about goals and abilities" (2008, 144; 2002). More technically, "Hypothesis H_1 can now be tested against hypothesis H_2 if and only if there exist true auxiliary assumptions A and an observation statement O such that (i) $\Pr(O|H_1 \text{ and } A) \neq \Pr(O|H_2 \text{ and } A)$, (ii) we now are justified in believing A, and (iii) the justification we now have for believing A does not

depend on believing that H_1 is true or that H_2 is true and also does not depend on believing that O is true (or that it is false)" (152). Sober illustrates his criterion for testability with the following story:

Suppose you are on a jury. Jones is being tried for murder, but you are considering the possibility that Smith may have done the deed instead. Evidence is brought to bear: A size 12 shoe print was found in the mud outside the house where the murder was committed, as was cigar ash, and shells from a Colt .45 revolver. Do these pieces of evidence favor the hypothesis that Smith is the murderer or the hypothesis that Jones is? It is a big mistake to answer these questions by inventing assumptions. If you assume that Smith wears a size 12 shoe, smokes cigars, and owns a Colt .45 and that Jones wears a size 10 shoe, does not smoke, and does not own a gun, you can conclude that the evidence favors Smith over Jones. If you make the opposite assumptions, you can draw the opposite conclusion. . . . What is needed is independently attested information about Smith's and Jones's shoe sizes, smoking habits, and gun ownership. (145)

In relation to the design argument, this means that we cannot simply attribute intentions and motives to the designer if we do not have any independent justification for doing so. For example, from the fact that humans have eyes, we cannot conclude that the intelligent designer, if such a being exists, must have had the intention for equipping humans with eyes: "What is needed is evidence about what God would have wanted the human eye to be like, where the evidence does not require a prior commitment to the assumption that there is a God and also does not depend on looking at the eye to determine its features" (Sober 2008, 146).

In the absence of independently justified auxiliary hypotheses, we are left with a designer without attributes, and the likelihood that such a designer wanted to (and was able to) create the world we observe cannot be calculated. As a result, so argues Sober, the design hypothesis as it stands is not and has never been testable against the Epicurean view or against the evolutionary hypothesis. One cannot exclude the possibility that such independently attested auxiliary assumptions will turn up in the future, but as long as this does not happen (and Sober does not expect that it ever will), the design hypothesis remains untestable. It follows that the design argument, framed as a likelihood argument, is officially dead.

3. Introducing Auxiliaries.

3.1. Background Knowledge and Observation. Sober's solution effectively prevents the practice of building observations into one's hypothesis,

but we argue that it does much more than that and hence is too restrictive. Consider the murder scenario described by Sober, but in a somewhat different light. Does the available evidence provide support for the hypothesis that someone committed a murder in the first place? Suppose the landlord is nowhere to be found, we find blood stains and broken glass in his bedroom, and we possess all the other evidence Sober alludes to. In addition, we know that the landlord neither smokes cigars nor has size 12 shoes. A detective on the scene wants to assess the plausibility of the following rival hypotheses:

H_1 = the landlord was murdered.

H_2 = the landlord is alive and left for an unexpected walk.

H_3 = the landlord killed himself and was then dragged away.

If the detective favors the murder hypothesis, we submit that she is justified in making the additional assumption that the hypothesized murderer, whoever it was, wears a size 12 shoe, smokes cigars, and used a Colt .45. This would be a matter of sound detective work, not of accusing Smith or Jones without basis.

O = a size 12 shoe print, cigar ash, and shells from a Colt .45 revolver were found in the bedroom.

H_1 = the landlord was murdered by X .

A_1 = X wears a size 12 shoe, smokes cigars, and owns a Colt .45.

What justifies our adopting auxiliary hypothesis A_1 ? In the first place, we are informed by our background knowledge (K) on human beings wearing shoes, occasionally smoking cigars, even less occasionally murdering people, and on the Colt .45 producing specific shells. But note that K , by itself, does not warrant our adopting A_1 . Only the conjunction of K with O and H_1 does. Does the choice of A_1 'depend' on looking at O in a way that is not allowed by Sober? It seems so.² At this point, Sober's requirement commits him to rejecting types of everyday reasoning that are obviously valid, but there is a charitable way to reconstruct his argument, by fine-tuning the dependence relation as follows:

If you want to construct an auxiliary hypothesis A for testing H with respect to O , then, though your adoption of A may be informed by O , it must be so in conjunction with at least one other, independent

2. There are ways to conjoin the rival hypotheses H_2 and H_3 with suitable auxiliaries so that they too yield the observations (arguably many of these ways are contrived). At this point, however, we are merely interested in some hypothesis that is *prima facie* most plausible, when conjoined with suitable auxiliaries, and show how Sober's framework does not allow us to favor it.

reason. By contraposition, if your choice of A is solely informed by O, then you are not conducting a proper test of H.

Sober's intrinsic objection against the design argument can then be rephrased: as there is currently no independent background knowledge available about the designer, and hence we are completely in the dark as to his identity, the biological argument does not get off the ground.³ Is this weaker version of Sober's argument defensible? We think it is not: if certain conditions (to be specified below) are satisfied, the design theorist could legitimately introduce auxiliaries in a way that violates this weaker criterion as well.

3.2. Gerrymandering and Spurious Unification. The problem central to Sober's concern is the practice of gerrymandering hypotheses by inventing ad hoc auxiliaries to fit the data. This type of reasoning is pervasive in much creationist writing, and its problems were already spelled out by Darwin in his discussion of the theory of special creation: "On the ordinary view of the independent creation of each being, we can only say that so it is; that it has pleased the Creator to construct all the animals and plants in each great class on a uniform plan; but this is not a scientific explanation" (2006, 677).

As Darwin noted, the theory of special creation amounts to "restating the fact in dignified language" (2006, 336). It is designed to yield the known observations and nothing more. By contrast, the main explanatory merit of evolutionary theory lies in its power to yield a "consilience of inductions" (Whewell 1840), by bringing together a wide array of facts from different domains and explaining them as following from the same basic principles: blind variation, heritability, and selective retention (Kitcher 1985).

The design theorist might object that his hypothesis also accomplishes this kind of unification, as every observation in the natural world is subsumed under the explanation of "God's will."

$H = \text{God wants it to be the case that } O_1 . . . O_n.$

As Kitcher has pointed out, however, in such an explanatory pattern the "nonlogical vocabulary which remains is idling" (1981, 528). The pattern does not impose constraints on the sentences that can be derived by using it, and thus it is able to accommodate any observation whatsoever. In Sober's vocabulary, the only reason for adopting the assumption that God really wants a specific fact O_i to be the case depends on looking at

3. According to Sober, this also holds if we assume that the 'designer' is simply an intelligent civilization from outer space (1999, 65).

O_i and nothing else. Thus, in reality the design theorist simply posits a new divine disposition for each and every observation instead of a single unifying explanation. As Sober (2008, 181) writes, “the fact that the model postulates a single *designer* is besides the point.”

Sober’s principle that “auxiliary assumptions must be justified without assuming that O is true” effectively undermines this type of “spurious unification” (Kitcher 1981, 528), together with other ways of gerrymandering, but does it leave any room for reasonable consideration of auxiliaries? The following *reductio* he offers in support of his thesis is ineffective (Sober 2008, 145): If we assume that O is true, then so is the disjunction “either H_1 is false or O is true.” If we take this disjunction as auxiliary hypothesis A_1 , then H_1 and A_1 entails O , even if H_1 has nothing to do with O . Thus, according to Sober, we cannot allow our auxiliary hypothesis to depend on O .

However, from the fact that one intuitively illegitimate move happens to violate Sober’s rule, it does not follow that all violations run into similar problems. What Sober needs is an argument to the effect that, whenever his rule is violated, we are indeed dealing with a move that is epistemically suspect. In the next sections, we demonstrate that some violations of Sober’s rule are perfectly legitimate. In particular, we will show how the design hypothesis could be made compelling in ways that violate Sober’s restriction concerning the choice of auxiliaries.

3.3. Ruling out Uninteresting Assumptions. Consider how Sober defines his proposed independence conditions: “the auxiliary assumptions used to bring the hypotheses H_1 and H_2 into contact with the observation O must be justified without assuming H_1 or assuming H_2 , or assuming O ” (2008, 145). Does this mean that, even in the process of fleshing out a hypothesis (“to bring . . . into contact with the observation O ”), we are never justified in adopting such or such auxiliary if doing so depends on looking at O (or exclusively at O , in our charitable reconstruction)? This seems overly restrictive. If we want to bring competing hypotheses into contact with O , we want to concentrate on eligible auxiliary hypotheses, and we do not pay attention to those that are extremely unlikely to yield the data we want to explain. If the detective wants to consider the murder hypothesis, and she finds Colt .45 shells around the blood stains, she makes the additional assumption that someone murdered the victim with a Colt .45, even if, at that point, the victim has not been found and no further evidence supports her tentative hypothesis.

Similarly, suppose that William Paley, reflecting on the origin of the human eye, constructed the following design hypothesis, conjoined with two additional assumptions:

H = the human camera eye was created by an intelligent designer.
 A_1 = the designer is interested in creating camera eyes.
 A_2 = the designer is capable of designing something as complex as the camera eye.

The adoption of both A_1 and A_2 seems reasonable enough since their negation is completely uninteresting, in the sense of being very unlikely to yield the data in question:

$\sim A_1$ = the designer has no interest at all in creating camera eyes.
 $\sim A_2$ = the designer is a bungler completely incapable of producing anything as complex as the camera eye.

Evidently, the likelihood of both H and $\sim A_1$ and H and $\sim A_2$, namely, $\Pr(O|H \text{ and } \sim A_1)$ and $\Pr(O|H \text{ and } \sim A_2)$ is extremely low. If we follow Sober's approach, however, this gives us no reason for adopting A_1 and A_2 because, in the absence of background knowledge about the designer, the independence rule is violated. But how do we rule out such uninteresting auxiliaries unless we take into account the observations with which we want to bring the hypothesis into contact?⁴

At this point, Sober may protest that there is a difference between preliminarily fleshing out a hypothesis and actually putting it to the test, a difference not unlike that between context of discovery and context of justification.⁵ Although looking at the observations one sets out to explain may be admissible in the former phase of hypothesis testing, it cannot be allowed in the latter. However, there does not seem to be a straightforward way to pinpoint where the first phase ends and the second one begins. What if the detective, upon arriving at the crime scene, is confronted with a number of clues at the same time, and there is no direct way to test her hypothesis against further data? Or what if the detective gradually accumulates new data, which slowly coalesce into the murder hypothesis? Besides, if the detective is allowed to glance at the crime scene before actually putting her hypothesis to the test, so is William Paley. In that case, what prevents Paley from tentatively inferring the craftsmanship of

4. If we pause to think about it, there do not seem to be many ways of justifying the introduction of an auxiliary, except by taking the observations into account that we set out to explain. Merely having independent reasons for accepting an auxiliary is not sufficient. Take, e.g., " A_1^* = naive set theory suffers from Russell's paradox." Arguably, $\Pr(O|ID \text{ and } A_1^*) = \Pr(O|ID)$ and $\Pr(O|\text{Chance and } A_1^*) = \Pr(O|\text{Chance})$. Even if we have (very good) independent reasons for accepting A_1^* , there is no use incorporating it as an auxiliary because it has no bearing on our observations in any way.

5. We would like to thank an anonymous referee for this suggestion. Although this is a viable solution, we should note that Sober's wording ("used to bring . . . into contact") does suggest that he includes this preliminary stage as well.

the intelligent designer from observing the human eye (first stage) and then proceeding to test his hypothesis by looking at, say, the visual apparatus of other animals (second stage)? But then looking at the eye would be admissible after all, contrary to Sober's assertions, and the design hypothesis would be in the running again. It seems that Sober's case against the design argument misses the mark.

3.4. Unification. A hypothesis can derive empirical support either by accommodating known observations in particular ways or by successfully predicting new observations. Predictivists attach special epistemic status to successful predictions, but some philosophers have questioned this different assessment (Harker 2008). We will first focus on the case of accommodation, as we think that the 'mere' accommodation of known data in an appropriate way would already make the design argument convincing.

The bugbear of accommodation is the temptation of the theorist to "overfit" the data, which consists of sacrificing the simplicity of one's hypothesis in order to attain a maximal fit with the available data (Hitchcock and Sober 2004). However, even philosophers who attribute special epistemic value to prediction acknowledge that accommodation need not be problematic per se, only that prediction guards the theorist against the temptation of overfitting. "It is possible to accommodate data without overfitting them, but when one is accommodating data, the temptation to overfit is always present. By contrast, when one accurately predicts new data that were not used in formulating one's theory, there is no opportunity to overfit those data" (20).

An appropriate measure against overfitting consists of balancing simplicity (as measured, e.g., by the Akaike Information Criterion; Akaike 1973) against fit with data, so that any loss of simplicity must be offset by a sufficient gain in fit with data (not just any gain of fit; see also Leplin 1975; Forster and Sober 1994). The ideal hypothesis, if any such is allowed by the available observations, is the one that both is sufficiently simple and achieves a maximum data fit. For example, the murder hypothesis H_1 is to be preferred if and to the extent that the detective, on adopting some suitable and simple set of auxiliaries $A_1 \dots A_m$, succeeds in unifying the available circumstantial evidence $O_1 \dots O_n$ in a way that cannot be accomplished as successfully without assuming H_1 . Recall the observations: a size 12 shoe print, cigar ash, and shells from a Colt .45 revolver are found in the bedroom, we can see blood stains and broken glass on the floor, and the landlord is nowhere to be found.

H_1 = the landlord has been murdered.

A_1 = the murderer wears a size 12 shoe, smokes cigars, and used a Colt .45.

It is not difficult to invent other hypotheses with suitable auxiliaries that also entail the observations, for example, “ H_2 = the landlord left for an unexpected walk,” conjoined with the following auxiliaries: somebody just threw a stone through the window, the shells from the Colt .45 dropped out of a visitor’s pocket, the landlord just slaughtered a pig in his house before his unexpected walk, and so on. It is clear, however, that the murder hypothesis is superior because it succeeds in unifying all the available data under a simple assumption.

This is not to say that H_1 outcompetes every possible hypothesis. For example, “ H^* = the landlord went underground” can account for the data if conjoined with the auxiliary “ A^* = in order to fake his own death, the landlord has left the shells, the blood stains,” Arguably, H^* and A^* is not far more complex than H_1 and A_1 , and it is equally unifying. Therefore, H^* might be an admissible competitor for H_1 .

Before moving on to the design hypothesis, we should note that, at the end of his chapter on the intelligent design (ID) hypothesis, Sober (2008) discusses the framework of model selection theory (MST) as an alternative to likelihoodism, and in that context he himself touches upon the explanatory virtue of unification. Statistical models (propositions with adjustable parameters) are to be preferred over their rivals to the extent that they are unifying and contain few adjustable parameters.⁶ If we view the goals and abilities of the designer as the adjustable parameters of the model, such unification should be accomplished by “collect[ing] different observations together and view[ing] them as consequences of a single plan that the designer(s) has in mind” (182).

We think this is a much more promising tack to show where the design argument goes off the rails. But our approach differs from Sober’s in a number of ways. First, whereas Sober only invokes simplicity and unification when framing the design argument within MST, we invoke these notions to justify the introduction of auxiliaries that violate Sober’s testability criterion within the likelihood framework. Second, whereas Sober does not offer a real verdict based on simplicity and unification, stating only that MST shows that good fit (*in casu* of the design hypothesis) does not imply truth or predictive success, we do want to offer such a verdict. Third, Sober does not approach the MST design argument in the most charitable way: “How should this [unification] be achieved? I don’t know: this is a task for intelligent-design theorists to address” (2008, 182). That is true enough, but it is also the task of the philosopher assessing the design argument to find out under what circumstances such a unifying design model would be conceivable. If, with certain provisos, design the-

6. Strictly speaking, the goal of the Akaike Information Criterion is predictive accuracy, not truth or probable truth.

orists are free to infer the attributes of the designer from the observations they set out to explain, as the model selection approach suggests, then Sober's "devastating objection" (126) against the design argument on the level of auxiliary choice is not so devastating after all. In that case, Sober has knocked out the design hypothesis one round too early, before it can even clash with the empirical evidence. As we will see in section 3.6, this has important consequences for the theistic design hypothesis, as well as for the status of empirical arguments that have been leveled against design.

How does the explanatory virtue of unification translate to the biological design hypothesis? Consider Paley (1802), which made a deep impression on the young Darwin. Paley's central argument states that adaptive complexity in the living world bears the mark of a designing intelligence: "Arrangement, disposition of parts, subserviency of means to an end, relation of instruments to a use, imply the presence of intelligence and mind" (12). Perceptive of the explanatory virtue of unification, Paley enumerates a wide variety of examples of contrivance and usefulness in nature, and he points out the coherence of animal body plans: "In comparing the eyes of different kinds of animals, we see in their resemblances and distinctions one general plan laid down, and that plan varied with the varying exigencies to which it is to be applied" (33). Apart from noting such similarities, however, Paley seems unable to discern any overall intentional plan in the creator's work, making only vague gestures in that direction. For example, a consideration of the bountiful diversity of nature "might induce us to believe that *variety* itself . . . was a motive in the mind of the Creator, or with the agents of his will" (372). Accordingly, Paley is unable to infer much about the designer's attributes and specific intentions, except for the point that he must have been at least as powerful and wise to be able to create all things we currently observe: "The attributes of such a Being, suppose his reality to be proved, must be adequate to the magnitude, extent, and multiplicity of his operations" (474).

In his penultimate chapter, Paley (1802) attempts to demonstrate at least the goodness of the creator. Tellingly, however, he makes recourse to convoluted rationalizations to explain away the preponderance of evil in the world, notably to the argument that God's ways are inscrutable to humans (see sec. 3.6 below). In the end, Paley does not flesh out his design hypothesis any further, and failing to have done so, he places his money on the explanatory necessity of *some* designer, for even a single instance of purposeful contrivance: "were there no example in the world of contrivance except that of the eye, it would be alone sufficient to support the conclusion which we draw from it, as to the necessity of an intelligent Creator" (81).

Modern ID advocates have made little progress since Paley. To the

extent that they have made attempts at all toward unification, they have mainly accomplished one of the “spurious” sort, attributing every particular observation to God’s will and maintaining that he moves in mysterious ways (see sec. 3.6). What is interesting for our critique of Sober’s likelihood reconstruction is the fact that the design argument *might* have achieved genuine unification, if only its advocates had succeeded in subsuming a wide array of natural phenomena under the assumption of a simple and distinct creative intention on the designer’s part (or a simple set of intentions). If only a few ‘parameters’ in the design hypothesis were to provide an elegant explanation for phenomena that resist any conceivable naturalistic explanation, it seems that our worries about overfitting would be assuaged. The fact that the choice of auxiliaries about the designer’s intentions and attributes ($A_1 . . . A_m$) would depend on the observations we set out to explain ($O_1 . . . O_n$), without the support of independent background knowledge, would then be of little concern.

In what way could the design argument achieve such genuine unification? Reflecting on the vast number and variety of beetle species on earth, the biologist J. B. S. Haldane once quipped that the creator, if he exists, has an “inordinate fondness for beetles.” Assuming, for the sake of the argument, that Haldane was making a serious theological point, as it stands his design argument is not very persuasive. Suppose, however, that Haldane happened to discover that beetles have minuscule Hebrew letters written on their shields, forming edifying biblical messages. Let’s say subsequent research demonstrates that beetles all over the world display these microscopic patterns, that they are encoded in beetle DNA, and that the fossil record suggests that beetles displayed these remarkable features even before humans arose on the scene.

The scenario is rather outlandish, but it will suit our purposes. Arguably, there is no way in which Hebrew letters, as opposed to meaningless scribble, could confer any selective advantage on beetles, through either natural selection or sexual selection. Nor could the phenomenon plausibly be the result of genetic drift or the by-product of other evolutionary adaptations. In general, it is very hard to see how the explanatory repertoire of the naturalistic scientist, consisting of blind and unguided processes, could succeed in explaining anything like the existence of Hebrew beetle decorations.

In the described case, the design hypothesis, conjoined with an auxiliary hypothesis about the designer’s abilities and intentions, would allow us to explain otherwise puzzling phenomena.

H = beetles are created by an intelligent designer.

A_1 = the intelligent designer has the ability to create beetles, is in-

ordinately fond of them, and he has used their bodies to inscribe his word.

One could object that, even in such an unlikely event, all the available evidence for naturalistic evolution still stands, and one anomaly does not suffice to undermine a well-substantiated scientific theory (Oppy 1996, 534). The point is well taken but ineffective, as we could easily fancy a world in which all the phenomena of the living world would converge on the intelligent work of a creator who, judging from his works, bears a suspicious likeness to the Judeo-Christian God. For example, suppose that thousands of living organisms in this world bore an autograph in Hebrew, unique for the species to which they belong, and that all the characters together formed the words of the Old Testament. Suppose, moreover, that we would not witness any of the examples of imperfections, rudimentary organs, and botched designs that are currently viewed as betraying an evolutionary heritage (see sec. 3.6). Or if this is not sufficient, think away the fossil record, the biogeographical and anatomical evidence for evolution, the evidence from genetics and embryology, and so on. Surely there must be some point at which the evidence would tilt in favor of ID at the expense of evolutionary theory. And so it should be. Which theories we can reliably accept about the world depends for a large part on contingent matters of fact, on what the world looks like. An adherent of Sober's approach, however, would remain unmoved, even by such a fanciful scenario, because the adoption of auxiliary A_1 (the properties of the Judeo-Christian God) still depends on looking at $O_1 \dots O_n$ (without independent background knowledge).

This example illustrates that the problem with the biological design argument as it stands is not so much that it relies on observations of living organisms to provide the theorist with some clues as to the character and intentions of the alleged designer, but that it yields nothing beyond those observations. Thus, although we agree with Sober that we need some "independent" reasons, broadly construed, for adopting auxiliaries $A_1 \dots A_m$, over and above the mere observations we set out to explain, we think Sober has construed these reasons too narrowly, neglecting the role played by explanatory unification. Sober mistakenly thinks that violating his independence condition always amounts to gerrymandering, apparently because he has extrapolated from a special problem with the construction of auxiliaries to a general assessment of design reasoning.⁷

7. An additional problem of Sober's approach concerns the different nontrivial ways in which we can separate the central hypothesis from auxiliary hypotheses. For example, returning to Haldane's beetles, we could reconstruct different design arguments: "H = God created the world and all living beings separately"; "O = there are a lot of beetles"; " A_1 = God has an inordinate fondness for beetles." An alternative recon-

In fact, our approach is more faithful to Sober's commitment to contrastive hypothesis testing than Sober's own treatment of the design argument: the design argument is currently outcompeted by evolutionary theory (for a recent overview, see Coyne 2009; Dawkins 2009), but if only design theorists would come up with evidence that defies all explanatory efforts in a naturalistic framework and that is elegantly explained on some suitable design hypothesis (in the sense discussed above), they would certainly deserve our attention (Boudry, Blancke, and Braeckman 2010). It will be clear that this assessment is all the more damaging to the intelligent hypothesis (see sec. 4.2).

3.5. Prediction. In the murder investigation we discussed, the detective need not predict, say, the exact location where the murderer hid his weapon. If the weapon is found by accident, fingerprints and a plausible motive of the suspect may count as sufficient incriminating evidence to convince a judge. Likewise, although the design hypothesis need not make spectacular predictions of novel phenomena, this would of course be a sure way of boosting the plausibility of the design hypothesis dramatically. As we noted in the previous section, predictivists attach special epistemic status to successful predictions. Predictivism comes in many flavors, and some of these flavors have faced some important criticisms, most notably from Sober himself (Hitchcock and Sober 2004).⁸ Hitchcock and Sober distinguish between global and local predictivism. The first "maintains that a theory which successfully predicts some observation will always be superior to one that accommodates the same observation" (3). For the latter, prediction is only sometimes superior to accommodation. Hitchcock and Sober's sympathies lie with the local variant (see sec. 3.4), as they show that there are cases in which accommodation is better than prediction. For the sake of the argument, however, suppose we are confronted with a strong predictivist who would be unimpressed by the ability of the design hypothesis to unify and accommodate known data. Is there any

struction would be to break up H further, into a core hypothesis and a number of auxiliary hypotheses, e.g., "H* = an intelligent being X created the world"; "A₂ = X created all living beings in the world separately"; "A₃ = X is omnipotent, benevolent, and omniscient (and God is the only person with these attributes)." Where does the design argument go wrong, according to Sober? Depending on how we slice up the cake, different propositions will count as auxiliary hypotheses. If we take up the first reconstruction, Sober will find only the additional assumption A₁ about God's fondness for beetles problematic (because it depends on O), whereas in the second reconstruction, the very attribution of omnipotence and benevolence to X, and the proposition about X's modus operandi, will be disallowed by Sober (because we do not have independent reasons for accepting A₂ and A₃).

8. One of the big divides is between "temporal novelty" and "use novelty." For a critique of creationism from a use novelty perspective, see Worrall (2002).

‘possible world’ in which the design hypothesis can achieve predictive success in addition to explanatory unification?

Suppose that many different organisms bore an autograph in Hebrew, unique for the species to which they belong. Suppose also that these autographs formed a very large part of the Old Testament, except for some missing quotes. Then the hypothesis H and the auxiliary A_1 of the previous section may be used to predict that there exist organisms we have as yet not discovered or not studied carefully enough, which bear the requisite inscriptions (maybe some verses from the book of Jonah on the fin of a new whale species). If we are able to predict what are the missing inscriptions, this furnishes us with an extra reason to accept H and A_1 , in addition to their presumed unificatory power.

It is not entirely clear, however, how Sober’s restrictions apply to predictions of new data, as opposed to the mere accommodation of known observations. If the design argument would allow us to make successful predictions of phenomena that have a very low probability on any non-contrived naturalistic hypothesis we can think of, would Sober still refuse to accept it? In the (novel) prediction case, the observation O that we use to test our competing hypotheses cannot enter into our considerations for choosing auxiliaries $A_1 \dots A_m$ because, by definition, O has not been observed yet. In what sense is the “independence” of $A_1 \dots A_m$ to be understood? Is it acceptable if our justification of $A_1 \dots A_m$ depends on other observations that are already known? If so, why does Sober not leave room for such cases of predictive success in setting up his intrinsic argument against design? In any case, our argument does not hinge on Sober’s approach disallowing some forms of successful prediction, as we have already demonstrated that it precludes valid forms of explanatory unification.

3.6. Imperfections and Evils. Many philosophers and scientists, starting with Darwin himself (for recent examples, see Coyne 2009, 86–91; Avise 2010a, 2010b), have argued that the clumsy and botched works of nature provide evidence against the design hypothesis. According to Sober, both the argument from imperfections and the argument from evil (Sober 2004) fall victim to the same objection that he leveled against the design argument itself, namely, that they make unwarranted assumptions regarding the character and intentions of the alleged designer.⁹ For example, discussing Stephen Jay Gould’s famous argument about the clumsy design of the panda’s pseudthumb (Gould 1980), Sober charges Gould with

9. Sober (2008, 164–67) rehearses the same line of reasoning, but he writes that he is not so sure anymore whether this puts the biological design argument on a par with the argument from evil.

simply “inventing assumptions” about the designer to reach a preestablished conclusion (Sober 2008, 128).

Sober’s criticism is off the mark for both a specific and a more general reason. First, Sober fails to see that these arguments are put forward in the particular context of the widespread belief in a benevolent and omnipotent creator with a purposeful creation plan. As soon as one accepts these traditional assumptions about the designer, as most theists do, including ID theorists (Forrest and Gross 2004), the pervasiveness of botched design and especially the existence of needless suffering is most damning (Hume 1779/2007; Mackie 1955). It goes without saying that, if one relinquishes some of the traditional attributes of God, the argument from evil no longer has any force. But the same does not apply to many instances of the argument from imperfections, which brings us to the second and more general problem with Sober’s argument.

At this point, one may argue that, although Gould was right to challenge theists with the panda’s thumb, his argument is still irrelevant with regard to the design hypothesis *in abstracto*. But we would go one step further: even if a design theorist is not committed to any particular religious doctrine about the designer’s attributes, the existence of puzzling imperfections and rudimentary organs, together with the countless instances of ineffective and wasteful processes, should worry him nonetheless. Such senseless and botched structures present a challenge not just to the traditional theological account (as in the case of the argument from evil) but to any attempt at subsuming the phenomena of the living world under a coherent design plan and, thus, to any attempt at genuine unification. This is not the proper place to enumerate examples or to give a full account of the argument from bad design, but let us briefly point out its evidential impact and logical structure. According to Sober, hypothesis testing is an inherently contrastive activity. We agree. The argument from bad design discriminates between evolution and design because the pervasiveness of such biological imperfections is much more plausible on an evolutionary understanding than on any noncontrived design hypothesis. In particular, the argument forces the design theorist either to invent a particular intention on the part of the designer for each new observation or to state that the designer must have wanted the world to look as though it evolved. The thrust of the bad design argument is one of disintegration: the sensible auxiliaries of the design hypothesis are ruled out, and only the contrived ones remain (Kitcher 1993, 18–25).

The surest indication that the unification of biological phenomena under the design hypothesis, conjoined with some suitable auxiliaries, is an all but impossible task is the fact that those who are eager to make a scientific case for design have never taken up the challenge to do so. Indeed, William Paley himself made only vague gestures in that direction.

Instead of fleshing out their design hypothesis, ID theorists have insisted that the designer is inscrutable and his intentions unfathomable (see, e.g., Johnson 1991, 67; Behe 2006, 223).¹⁰ They have even accused their critics of making unwarranted assumptions about the intelligent designer (Nelson 1996). For example, Michael Behe wrote: “Another problem with the argument from imperfection is that it critically depends on a psychoanalysis of the unidentified designer. Yet the reasons that a designer would or would not do anything are virtually impossible to know unless the designer tells you specifically what those reasons are” (2006, 223).

Surprisingly, but in conformity with his view on auxiliary assumptions, Sober grants that this is a “good reply by creationists” (2007, 4). However, we submit that Behe’s response is an all too convenient way of insulating the design argument against empirical objections without adding any empirical substance to the theory (Pennock 1999, 249). Behe’s insistence on the inscrutability of the designer is not a sign of sensitivity to a pressing epistemological problem that his critics have overlooked, but it is an epistemological retreat that is symptomatic of a degenerated research program (Boudry and Braeckman 2011). As Philip Kitcher noted, “As the evidence accumulates, creationists increasingly must take refuge in responses Darwin saw as unsatisfactory evasions, appealing to the thought that these properties of life are unfathomable mysteries” (2007, 58).

The evasive arguments of modern creationists indicate that there is no noncontrived way to flesh out the design hypothesis that will stand up to the facts, no matter what auxiliary hypotheses one adopts. Taking into account that the living world, and especially the peculiar examples of ‘bad design’, looks very much like the kind of world we would expect if there were no design at all but only mindless natural processes at work, the biological design hypothesis is effectively dead.

4. Conclusion.

4.1. Likelihoods and Contrastivism. Sober has correctly identified the main problem with the likelihood reconstruction of design reasoning, namely, that one cannot freely invent auxiliary assumptions to raise the likelihood of one’s favorite hypothesis. But his solution does not hold water. To demand that auxiliary hypotheses be justified “independently” of the available data one sets out to explain is overly restrictive and

10. For political reasons, the ID movement prefers to brush aside theological quarrels and fight for a common cause. The ‘minimal’ design hypothesis is interesting because ID advocates think it allows them to circumvent the establishment clause against the teaching of religion (Forrest and Gross 2004).

commits one to rejecting forms of obviously valid reasoning.¹¹ Even on a charitable reconstruction of his independence condition, Sober's intrinsic objection against the (present) testability of the design argument fails, as he has mistakenly identified the creationist practice of gerrymandering as an inevitable trap into which all observation-based introduction of auxiliaries must fall.

While Sober precludes the design hypothesis from entering the boxing ring, we do extend such an invitation. If the opponent is evolutionary theory, though, we would not stake our money on the flyweight called design. Advocates of ID are perfectly free to construct observation-based auxiliary hypotheses about the intentions and attributes of their designer, provided that these assumptions are elegant and unifying and are not just tailored to individual observations (which they typically are). In fact, *pace* Sober's likelihood approach, this is what a reasonable critic would demand from them (Pennock 1999, 199–201; Dawes 2007, 78–79). As long as design theorists fail to flesh out their hypothesis, we are left with an unnamed and unknown designer, and we can do no more than restating the facts in dignified language. Not surprisingly, when confronted with the argument from imperfection, design theorists have dodged the issue altogether, insisting that the whole affair is unfathomable.

Despite his commitment to contrastivism, Sober (2008) has arrived at the conclusion that the biological design argument fails even before it enters into competition with rival hypotheses. If Sober's goal was to arrive at the "strongest, most defensible, version of the [design] argument" (113), however, he should have allowed it to enter into competition with rival hypotheses and then fault it for its lack of unification and predictive success. This is not to say that Sober's verdict is in conflict with his metatheoretical views, because any contrastivist is free to propose additional criteria for hypotheses to be met before one even gets to the point of worrying about contrastive testing. That having been said, our proposal seems more in the spirit of contrastivism, as it does allow the design hypothesis to enter the boxing ring of hypothesis testing. It is also more in line with Sober's earlier view that, before Darwin came up with a serious alternative explanation for adaptive complexity, it was not surprising that "intelligent people strongly favored the design hypothesis" and that "Darwin entirely altered the dialectical landscape of this problem" (36). Remarkably, Sober (2008, 125) rejects that very claim in almost exactly the

11. An advocate of Bayesianism may argue that Sober makes trouble for himself by strictly adhering to the likelihood approach and refusing to talk about prior probabilities. For a Bayesian, the contrived hypotheses simply must have lower priors than the weaker and more general design hypothesis ("the world was created by a designer of some kind").

same wording, arguing that Paley's design argument was long dead even before evolutionary theory arrived on the scene.¹²

4.2. *Taking the Design Argument Seriously (and Then Rejecting It)*. Ever since the design argument was formulated, there have been philosophical attempts to demonstrate that it is guilty of some fundamental flaw of reasoning and that we do not need an alternative explanation to see why this is so. Spinoza and Hume were among the first to argue against the program of natural theology (for more recent examples of this approach, see Oppy [1996], Pigliucci [2002], and Scott [2004]). We think that the design argument was difficult to resist before the advent of Darwin's theory, even though Hume had already pointed out some of the damaging problems it faces; note, however, that this historical assessment does not necessarily follow from the argument developed in this article. Even if one accepts the claim that the design inference is not intrinsically defective, one is still free to maintain that the empirical evidence, as it was available to natural theologians before Darwin, never favored it (for a critical discussion of the evidential warrant of the design argument before Darwin, see Oppy [1996] and Gliboff [2000]).

Given the philosophical consensus view that the biological design argument is a failure, is it really important to quarrel over where exactly it goes wrong? We think it is. The fate that befell the design argument illustrates a number of important philosophical issues regarding the choice of auxiliary hypotheses, the problem of gerrymandering, and the explanatory virtue of unification. Moreover, different diagnoses of the design argument are wedded to different assessments of its epistemic status. One unexpected consequence of typical a priori or fundamental objections to the design argument is that, ironically, they are less damaging to the design hypothesis than a posteriori objections (Boudry et al. 2010). If we accept Sober's critique of auxiliary assumptions, not only is the design argument stillborn even before any empirical evidence can be brought to bear on it, but the empirical arguments against design will not get off the ground either (Sober 2008, 126–28). If advocates of design are not allowed to make unjustified assumptions about the designer's attributes, then neither are their critics. Hence, Sober's symmetric critique unwittingly suggests that the critics are equally as unjustified in rejecting ID as the advocates

12. Sober writes that, according to a common opinion among biologists, "Paley reasoned correctly . . . but that the dialectical landscape shifted profoundly when a third hypothesis [Darwin's theory of evolution] was formulated" (2008, 125). On the next page, he rejects this position and claims that Paley's design argument has always been flawed.

are in defending it. We think this is mistaken, and it is a conclusion that Sober would want to avoid.

REFERENCES

- Akaike, Hirotugu. 1973. "Information Theory and an Extension of the Maximum Likelihood." In *Proceedings of the 2nd International Symposium on Information Theory*, ed. B. N. Petrov and F. Csaki, 267–81. Budapest: Akademiai Kiado.
- Avise, John C. 2010a. "Footprints of Nonsentient Design inside the Human Genome." *Proceedings of the National Academy of Sciences* 107 (Suppl. 2): 8969–76.
- . 2010b. *Inside the Human Genome: A Case for Non-intelligent Design*. Oxford: Oxford University Press.
- Behe, Michael J. 2006. *Darwin's Black Box: The Biochemical Challenge to Evolution*. 10th ed. New York: Simon & Schuster.
- Boudry, Maarten, Stefaan Blancke, and Johan Braeckman. 2010. "How Not to Attack Intelligent Design Creationism: Philosophical Misconceptions about Methodological Naturalism." *Foundations of Science* 15 (3): 227–44.
- Boudry, Maarten, and Johan Braeckman. 2011. "Immunizing Strategies and Epistemic Defense Mechanisms." *Philosophia* 39 (1): 145–61.
- Coyne, Jerry A. 2009. *Why Evolution Is True*. New York: Viking.
- Darwin, Charles. 2006. *The Origin of Species: A Variorum Text*. Philadelphia: University of Pennsylvania Press.
- Dawes, G. W. 2007. "What Is Wrong with Intelligent Design?" *International Journal for Philosophy of Religion* 61 (2): 69–81.
- Dawkins, Richard. 1986. *The Blind Watchmaker*. Harlow: Longman.
- . 2009. *The Greatest Show on Earth: The Evidence for Evolution*. London: Bantam.
- Forrest, Barbara C., and Paul R. Gross. 2004. *Creationism's Trojan Horse: The Wedge of Intelligent Design*. Oxford: Oxford University Press.
- Forster, Malcolm R., and E. Sober. 1994. "How to Tell When Simpler, More Unified, or Less Ad-Hoc Theories Will Provide More Accurate Predictions." *British Journal for the Philosophy of Science* 45 (1): 1–35.
- Gliboff, Sander. 2000. "Paley's Design Argument as an Inference to the Best Explanation; or, Dawkins' Dilemma." *Studies in History and Philosophy of Science C* 31 (4): 579–97.
- Gould, Stephen Jay. 1980. *The Panda's Thumb: More Reflections in Natural History*. 1st ed. New York: Norton.
- Harker, David. 2008. "On the Predilections for Predictions." *British Journal for the Philosophy of Science* 59 (3): 429–53.
- Hitchcock, Christopher, and Elliott Sober. 2004. "Prediction versus Accommodation and the Risk of Overfitting." *British Journal for the Philosophy of Science* 55 (1): 1–34.
- Hume, David. 1779/2007. *Dialogues Concerning Natural Religion and Other Writings*. Cambridge Texts in the History of Philosophy. Cambridge: Cambridge University Press.
- Johnson, Phillip E. 1991. *Darwin on Trial*. Washington, DC: Regnery Gateway.
- Kitcher, Philip. 1981. "Explanatory Unification." *Philosophy of Science* 48 (4): 507–31.
- . 1985. "Darwin's Achievement." In *Reason and Rationality in Natural Science*, ed. Nicholas Rescher, 127–89. Lanham, MD: University Press of America.
- . 1993. *The Advancement of Science: Science without Legend, Objectivity without Illusions*. New York: Oxford University Press.
- . 2007. *Living with Darwin: Evolution, Design, and the Future of Faith: Philosophy in Action*. Oxford: Oxford University Press.
- Leplin, Jarrett. 1975. "The Concept of an Ad Hoc Hypothesis." *Studies in History and Philosophy of Science* 5 (4): 309–45.
- Mackie, John L. 1955. "Evil and Omnipotence." *Mind* 64 (254): 200–212.
- Nelson, Paul A. 1996. "The Role of Theology in Current Evolutionary Reasoning." *Biology and Philosophy* 11 (4): 493–517.

- Oppy, Graham. 1996. "Hume and the Argument for Biological Design." *Biology and Philosophy* 11 (4): 519–34.
- Paley, William. 1802. *Natural Theology*. London: Gould & Lincoln.
- Pennock, Robert T. 1999. *Tower of Babel: The Evidence against the New Creationism*. Cambridge, MA: MIT Press.
- Pigliucci, Massimo. 2002. *Denying Evolution: Creationism, Scientism, and the Nature of Science*. Sunderland, MA: Sinauer.
- Scott, Eugenie Carol. 2004. *Evolution vs. Creationism: An Introduction*. Berkeley: University of California Press.
- Sober, Elliott. 1999. "Testability." *Proceedings and Addresses of the American Philosophical Association* 73 (2): 47–76.
- . 2000. *Philosophy of Biology*. 2nd ed. Boulder, CO: Westview.
- . 2002. "Intelligent Design and Probability Reasoning (Developing an Philosophical Epistemology of Irreducible Complexity)." *International Journal for Philosophy of Religion* 52 (2): 65–80.
- . 2004. "The Design Argument." In *The Blackwell Guide to Philosophy of Religion*, ed. W. Mann, 117–47. Oxford: Blackwell.
- . 2007. "What Is Wrong with Intelligent Design?" *Quarterly Review of Biology* 82 (1): 3–8.
- . 2008. *Evidence and Evolution: The Logic Behind the Science*. Cambridge: Cambridge University Press.
- . 2009. "Absence of Evidence and Evidence of Absence: Evidential Transitivity in Connection with Fossils, Fishing, Fine-Tuning, and Firing Squads." *Philosophical Studies* 143 (1): 63–90.
- Weisberg, Jonathan. 2005. "Firing Squads and Fine-Tuning: Sober on the Design Argument." *British Journal for the Philosophy of Science* 56 (4): 809–21.
- Whewell, William. 1840. *The Philosophy of the Inductive Sciences: Founded upon Their History*. London: Parker.
- Worrall, John. 2002. "New Evidence for Old." In *In the Scope of Logic, Methodology and Philosophy of Science*. Vol. 1 of *11th International Conference of Logic, Methodology and Philosophy of Science*, ed. J. Gardenfors, K. Wolenski, and K. Kijania-Placek, 191–209. Dordrecht: Kluwer.