On Likelihoodism and Intelligent Design

Sebastian Lutz*

Draft: 2011–02–14

Abstract

Two common and plausible claims in the philosophy of science are that (i) a theory that makes no predictions is not testable and (ii) one cannot confirm a theory by criticizing a competing one absent further assumptions about their relation. Elliott Sober has developed these claims within likelihoodism, which defines the testability and confirmation of a theory only in contrast to another, and has argued that the claims hold for intelligent design (ID) when contrasted with evolutionary theory (ET). I show that Sober's arguments rely on a contentious hidden premise, and that within likelihoodism, both claims are false for ID and ET under his assumptions and one very weak further assumption about ID and ET. I then show that, given Sober's assumptions, the claims are true for a non-contrastive criterion of testability close to the Bayesian one and the relevance criterion of confirmation.

§1. Introduction. Likelihoodism as defended, for instance, by Sober (1990; 2008) and Royall (1997) defines the testing (confirmation and disconfirmation) of a theory only in contrast to another theory. On its own, likelihoodism therefore requires a fundamental revision of all concepts that rely on confirmation or testing and are commonly not defined contrastively, such as the notion of testability (§3). Likelihoodism therefore also requires the modification of claims that rely on these concepts, including the common and plausible claims that (i) a theory that does not make predictions is not testable and (ii) a theory cannot be confirmed by criticizing a competing theory, absent further assumptions about the theories' relation (reviewed in §2). Over the last decade, Sober (1999; 2007; 2008) has argued convincingly that intelligent design (ID), the theory that all complex adaptations of organisms are created by

*Theoretical Philosophy Unit, Utrecht University. sebastian.lutz@gmx.net. Many thanks to Thomas Müller, Herman Philipse, Elliott Sober, and Alana Yu for helpful comments and discussions.

a designer, makes no predictions.[1] Therefore, he argues, claims (i) and (ii) are true within likelihoodism when ID is contrasted with evolutionary theory (ET) (§4). I will show (§5) that Sober's arguments rest on a contentious hidden premise and that under one very weak assumption about ID and ET, both claims are false in likelihoodism. Since both claims are at the core of Sober's criticism of ID and are generally plausible in their own right, this provides a strong reason to abandon the contrastive definitions of testability and confirmation in favor of two non-contrastive ones, under which the claims are true.

§2. Testability and confirmation in Bayesianism. Claims (i) and (ii) are easily proved within Bayesianism, a very common conception of confirmation that Sober (2008, §1.2) discusses but ultimately rejects. Here, the confirmation of a theory H by an observation O relative to suitable auxiliary assumptions A is defined as follows (Howson and Urbach 1993, §7a; Sober 2008, 15):

Definition (I). Observation O confirms H relative to A if and only if

$$\Pr(H \mid O \wedge A) > \Pr(H \mid A)\,. \tag{1}$$

For H to be actually confirmed, O has to be true. Here and in the following the auxiliary assumptions must be included because "hypotheses rarely make observational predictions on their own" (Sober 1999, 54). H is then tested by O if and only if $\Pr(H \mid O \wedge A) \neq \Pr(H \mid A)$.

It is clear that claim (ii) holds, since no competing theory H' occurs in definition (I). For a criticism of H' to confirm H, one would need at least the further assumption that for some observation O, $\Pr(H' \mid O \wedge A) < \Pr(H' \mid A)$ only if $\Pr(H \mid O \wedge A) < \Pr(H \mid A)$. This would be the case, for instance, if the auxiliary assumptions A were such that H and H' are the only plausible competing theories.

Sober (1999, 48) states for any relation R:

> If a set of observations provides a test of a proposition because it bears relation R to that proposition, then a proposition is testable when it is possible for there to be a set of observations that bears relation R to the proposition. Testing is to testability as dissolving is to solubility.

Since definition (I) determines what it is to test a theory, it therefore also determines a criterion of testability (cf. Sober 2008, 150)[2]:

[1] In fact, Sober (2007, 3) defines ID as the theory that *the* complex adaptations of organisms are created by a designer. Implicitly relying on Russell's explication of the definite article, he then claims that ID does predict the existence of complex adaptations. Since this sole prediction occurs for any theory whatsoever that purports to give a cause for "the" adaptations, and since the prediction disappears when the definite article is exchanged for a universal quantifier or explicated following Strawson, I consider it an artifact. In the discussion that follows, it would only necessitate that claim (i) be rephrased to exclude this particular prediction.

[2] Sober (2008, 144f) does not mention auxiliary assumptions in his discussion (and rejection) of the definition, but is very explicit about their relevance.

Definition (II). H can be tested relative to A if and only if there is a possible observation O such that

$$\Pr(H \mid O \wedge A) \neq \Pr(H \mid A)\,. \tag{2}$$

Sober now states that a theory makes predictions "either by deductively entailing that an observation will occur or by conferring a probability on an observational outcome" (Sober 2008, 130). Formally, Sober (2008, 151) defines for deductive predictions relative to A that proposition P "has observational implications if and only if there exist[s an] observation statement O, such that [. . .] P & A entails O, but A by itself does not entail O". Sober does not give a formalization for probabilistic predictions; I suggest

Definition (III). H makes probabilistic predictions relative to A if and only if there is a possible observation O such that

$$\Pr(O \mid H \wedge A) \neq \Pr(O \mid A)\,. \tag{3}$$

According to this definition, a theory H makes probabilistic predictions if and only if there exist a value p and an observation statement O such that $H \wedge A$ confers p on O, but A by itself does not confer p on O. This nicely captures Sober's informal formulation given above. It is furthermore easily shown that according to definition (III), H does not make any probabilistic predictions if and only if it makes the same probabilistic predictions as a tautology (which makes none).[3] Finally, this definition contains Sober's definition of deductive predictions as a special case (Lutz 2011b, claim 8).

Claim (i) now follows from Bayes's theorem,

$$\frac{\Pr(H \mid O \wedge A)}{\Pr(H \mid A)} = \frac{\Pr(O \mid H \wedge A)}{\Pr(O \mid A)}\,, \tag{4}$$

since one can easily prove that a theory is testable according to definition (II) if and only if it makes probabilistic predictions according to definition (III).

§3. Contrastive testability and confirmation. Sober (2008, 24–30, 150) eschews both Bayesian confirmation (definition (I)) and Bayesian testability (definition (II)) because they rely on the assignment of probabilities to whole theories. Instead, he suggests to define confirmation contrastively (Sober 2008, 32):

Definition (IV). Observation O favors hypothesis H_1 over hypothesis H_2 relative to A if and only if

$$\Pr(O \mid H_1 \wedge A) > \Pr(O \mid H_2 \wedge A)\,. \tag{5}$$

[3] Two theories H_1 and H_2 make the same probabilistic predictions given A if and only if for all observations O, $\Pr(O \mid H_1 \wedge A) = \Pr(O \mid H_2 \wedge A)$. It is clear that tautologies do not make any probabilistic predictions, since adding a tautology to any set of sentences with a set of possible probability distributions will not further restrict the latter set.

The terms in the inequality are called the *likelihoods* of H_1 and H_2 (for O), from which this approach derives its name 'likelihoodism'. Through the relations between confirmation, testing, and testability, definition (IV) determines testability (Sober 2008, 152):

Definition (V). H_1 can be tested against H_2 relative to A if and only if there is a possible observation O such that

$$\Pr(O\,|\,H_1 \wedge A) \neq \Pr(O\,|\,H_2 \wedge A)\,. \tag{6}$$

For the purposes of definition (V), Sober (2010, 2f) considers the inequality (6) true if and only if both likelihoods are defined and different (cf. Lutz 2011a, §3.1). Since inequality (5) entails inequality (6) and definition (IV) determines definition (V), inequality (5) is also true only if both likelihoods are defined.

According to Sober (2008, 37), likelihoodism does not require a fundamental revision of the common notions of confirmation and testability. Rather, likelihoodism "emerges as a statistical philosophy distinct from Bayesianism only when [it] is not possible" to "assign probabilities to hypotheses [. . .] by appeal to empirical information". However, this alleged continuity with Bayesianism has been questioned both for contrastive confirmation (Fitelson 2007) and contrastive testability (Lutz 2011b, §4) on formal grounds.

§4. Testability and confirmation of ID against ET. The core of Sober's application of the likelihood approach to the debate about ID (1999, 62ff; 2007, 6; 2008, §2.12, §2.15) is a convincing case for the claim that ID does not make any predictions because it is impossible to know the designer's intentions. From this he concludes that ID "cannot be tested against evolutionary theory" (Sober 1999, 64). That ID is not testable is also a common position outside of likelihoodism (cf. Gould 1983, 256), and an instance of claim (i).

To show that within likelihoodism, claim (ii) applies to ID when it is contrasted with ET, Sober (2007, 7) argues as follows:

> Defenders of ID often claim to test their position [. . .] by criticizing the theory of evolution. Behe (1996) contends that evolutionary processes cannot produce "irreducibly complex" adaptations; since we observe such traits, evolutionary theory is refuted, leaving ID as the only position standing. [T]his argument does nothing to test ID. For ID to be testable, *it* must make predictions. The fact that a different theory makes a prediction says nothing about whether ID is testable.

Going from bottom to top, Sober's argument amounts to the following: Since ID does not make predictions, it is not testable. Therefore, it is not tested by observations of irreducibly complex adaptations (this is a trivial proof given definitions (V) and (IV)). Hence ID is not confirmed by them. Note also that Sober here assumes that ET makes predictions. Elsewhere, Sober (1999, 66f) points out that there may be other theories besides ID which predict traits that are improbable according to ET, and continues:

> [T]he defect in this argument that I'm now pointing to is different. [...] The worst-case scenario for Darwinism is that the theory, with appropriate auxiliary assumptions, entails that what we observe was very improbable. However, this, by itself, isn't enough to reject Darwinism and opt for the hypothesis of intelligent design. We need to know how probable it is that the features would exist, if they were the result of intelligent design. *Both* hypotheses must make predictions if the observations are to help us choose between them.

Sober here directly infers that because ID does not make predictions, a trait that is improbable according to ET does not test, and therefore also does not confirm, ID. That a low likelihood of ET cannot confirm ID is also a common position outside of likelihoodism (cf. Pennock 2011, 188), and an instance of claim (ii). Note that outside of likelihoodism, this instance of claim (ii) does not rely on any specific feature of ID itself. To establish the claim for ID in contrast to ET, on the other hand, Sober's argument relies on the premise that ID does not make predictions.

§5. The failure of likelihoodism. I will now show that Sober's application of the likelihood approach to the debate about ID and ET not only fails, but also provides support for two alternative, non-contrastive conceptions of testability and confirmation.

Sober's arguments for claims (i) and (ii) about ID and ET have a hidden premise. By assumption, ID does not makes predictions, so that $\Pr(O \mid \text{ID} \wedge A) = \Pr(O \mid A)$ whenever both probabilities are defined. Furthermore, it follows from claim (i) that inequality (6) is false for every O, and thus $\Pr(O \mid \text{ET} \wedge A) = \Pr(O \mid \text{ID} \wedge A)$ whenever both probabilities are defined. Therefore, *for any O such that* $\Pr(O \mid \text{ET} \wedge A) \neq \Pr(O \mid A)$ *and both probabilities are defined,* $\Pr(O \mid \text{ID} \wedge A)$ *must be undefined*. This same hidden premise also follows from the premise that ID makes no predictions and claim (ii), that a low likelihood of ET cannot confirm ID. For if $\Pr(O \mid \text{ET} \wedge A) < \Pr(O \mid A)$, then $\Pr(\neg O \mid \text{ET} \wedge A) > \Pr(\neg O \mid A)$, and thus no matter the value of ET's likelihood, it is possible that ID is confirmed in contrast to ET unless the hidden premise is fulfilled. The premise is thus necessary for both of Sober's arguments to be valid. It is easy to see that it is also sufficient.

However, contrary to the hidden premise, it is plausible that for a theory H that makes no predictions, $\Pr(O \mid H \wedge A)$ is defined whenever $\Pr(O \mid A)$ is defined. For if $\Pr(O \mid A)$ is defined, then, given A, one must expect a specific regularity of occurrences of O. But if $\Pr(O \mid H \wedge A)$ is undefined, one must expect a breakdown of this regularity, and this expectation is plausibly a prediction. An example would be the prediction that under specific circumstances, some law fails that was assumed to hold universally.

That $\Pr(O \mid \text{ID} \wedge A)$ is at least *sometimes* defined when $\Pr(O \mid A)$ is defined follows from an example given by Sober (1999, n. 24) himself:

> Let H_1 = God created the eye, E = Jones is pregnant, A = Jones is sexually active, and H_2 = Jones used birth control. It is possible to test H_1 against H_2;

> given independently attested background assumptions A, E favors H_1 over H_2. The reason is that $\Pr(E \mid A) = \Pr(E \mid A \,\&\, H_1) > \Pr(E \mid A \,\&\, H_2)$.

Since for Sober (1999, 62, 65), H_1 here stands for ID, $\Pr(O \mid \text{ID} \wedge A)$ is defined for some observation O, for example that Jones is pregnant. It is thus not clear why, for all O with $\Pr(O \mid \text{ET} \wedge A) \neq \Pr(O \mid A)$, ID should render the probability for O undefined.

ET makes predictions, and, more specifically, there are many O^* such that $\Pr(O^* \mid \text{ET} \wedge A) \neq \Pr(O^* \mid A)$ and both probabilities are defined. Given the preceding considerations, it is very plausible that *at least for one* such O^*, $\Pr(O^* \mid \text{ID} \wedge A)$ is defined. This very weak assumption is also independently plausible: Take the statement that the eye of some organism has a specific feature. Based on our background knowledge about the ratio of the occurrence of this feature in other organisms' eyes, we might be able to assign a specific probability p to the occurrence of this feature, and based on our background knowledge and ET, we might be able to assign a different probability. But based on our background knowledge, we can also infer that *if* that eye was created by a designer, this designer had the intention and ability to create an eye with that feature with probability p. In conjunction with ID, the probability for the occurrence of the feature then does not become undefined, but rather keeps the value p.

Under this very weak assumption, claim (i) is false for ID and ET within likelihoodism. For then $\Pr(O^* \mid \text{ET} \wedge A) \neq \Pr(O^* \mid A) = \Pr(O^* \mid \text{ID} \wedge A)$, where all probabilities are defined, so that ID can be tested against ET. It is also easy to prove that the observation O^* favors ID over ET according to definition (IV) if and only if $\Pr(O^* \mid \text{ET} \wedge A) < \Pr(O^* \mid A)$. Thus the real worst-case scenario for ET is that it assigns a true observation statement O^* a lower probability than the auxiliary assumptions alone. And if Behe were to prove this for some true O^*, he would thereby show that ID is contrastively confirmed. Therefore claim (ii) is also false for ID and ET within likelihoodism.

This argument can be repeated for any two theories that fulfill the very weak assumption: a tautologous theory contrasted with quantum physics, for instance, or the nonsense theory 'Foo is bar' contrasted with plate tectonics. But the important conclusion is *not* that ID and analogous theories are contrastively testable and could be contrastively confirmed by discovering some low likelihood of ET or analogous theories. For Sober's criticisms of ID are the common and plausible ones: ID does not make predictions and therefore cannot be tested nor confirmed. The above considerations only show that Sober's expression of these criticisms in terms of contrastive testability and contrastive confirmation is inconsistent with his premises and a very weak assumption about ID and ET.

It is a severe criticism to charge that a theory does not make predictions, because it means that the theory only asserts observation statements or probabilities for observation statements that are already expected independently of the theory. But to make this point, definition (V) is not needed; definition (III) already captures Sober's criticism. Definition (III) is also a good candidate for a criterion of testability: Arguably unlike contrastive testability, it is continuous with Bayesian testability (because claim (i) holds for Bayesianism). And since it does not require assigning probabilities to whole theories, it is immune from Sober's criticism of Bayesianism.

Because of the relation between testability and testing, the use of definition (III) as criterion of testability entails that an observation O tests H relative to A if and only if $\Pr(O \mid H \wedge A) \neq \Pr(O \mid A)$. This allows for two definitions of confirmation (one with a 'less-than', one with a 'greater-than' sign), one of which is, arguably unlike contrastive confirmation, continuous with Bayesian confirmation:

Definition (VI). Observation O confirms H relative to A if and only if

$$\Pr(O \mid H \wedge A) > \Pr(O \mid A) . \tag{7}$$

This definition is equivalent to the Bayesian definition (I) whenever the theory can be assigned a probability, but, like Sober's definition (IV), does not presume that this is possible. It has also frequently been defended as a criterion of confirmation (cf. Mackie 1969).

With definition (III) as a criterion of testability and with definition (VI), claims (i) and (ii) are indeed true if ID does not make predictions. That the two definitions stay close to the common concepts of testability and confirmation is an additional advantage. For the contrastive concepts demand such fundamental revision in the conceptualization of common claims that even proponents of likelihoodism can be led into inconsistency.

References

Behe, M. J. (1996). *Darwin's Black Box: The Biochemical Challenge to Evolution*. Free Press, New York.

Fitelson, B. (2007). Likelihoodism, Bayesianism, and relational confirmation. *Synthese*, 156:473–489.

Gould, S. J. (1983). *Hen's Teeth and Horse's Toes*. W. W. Norton & Company, New York, NY, and London.

Howson, C. and Urbach, P. (1993). *Scientific Reasoning: The Bayesian Approach*. Open Court, Chicago and La Salle, IL. Second edition. Second printing 1996.

Lutz, S. (2011a). On a contrastive criterion of testability I: Defining contrastive testability. Preprint: `http://philsci-archive.pitt.edu/8574/`. Forthcoming.

Lutz, S. (2011b). On a contrastive criterion of testability II: The material inadequacy of contrastive testability. Preprint: `http://philsci-archive.pitt.edu/8575/`. Forthcoming.

Mackie, J. L. (1969). The relevance criterion of confirmation. *The British Journal for the Philosophy of Science*, pages 27–40.

Pennock, R. T. (2011). Can't philosophers tell the difference between science and religion?: Demarcation revisited. *Synthese*, 178(2):177–206. Special issue: *Evolution and its Rivals*.

Royall, R. M. (1997). *Statistical Evidence: A Likelihood Paradigm*, volume 71 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, New York.

Sober, E. (1990). Contrastive empiricism. In Savage, C. W., editor, *Scientific Theories*, volume 14 of *Minnesota Studies in the Philosophy of Science*, pages 392–410. University of Minnesota Press, Minneapolis, MN.

Sober, E. (1999). Testability. *Proceedings and Addresses of the American Philosophical Association*, 73(2):47–76.

Sober, E. (2007). What is wrong with intelligent design? *The Quarterly Review of Biology*, 82(1):3–8.

Sober, E. (2008). *Evidence and Evolution: The Logic Behind the Science*. Cambridge University Press, Cambridge.

Sober, E. (2010). Comments on Sebastian Lutz's "On Sober's criterion of contrastive testability". Unpublished typescript.