Bayesianism and IBE: the case of individual vs. group selection

January 18, 2016

Abstract

Different views have been proposed about how Inference to the Best Explanation (IBE) and Bayesianism might be compatible with one another. One is a hybrid view, according to which explanatory considerations play a role in driving the Bayesian machinery. Another is an “emergent compatibilist” view, in which an independently motivated Bayesian model of IBE is provided, so that explanatory considerations emerge from the Bayesian machinery rather than driving it. Specific scientific arguments can serve as test cases for these general views. Here I argue that the case of Williams’ argument against group selection, discussed by Elliott Sober, is better understood with the emergent compatibilist picture, than with the hybrid view. This analysis of the case challenges Elliott Sober’s claim that the epistemic significance of appeals to the explanatory virtue of parsimony is highly case-specific. Instead I suggest a more unified picture of IBE and its connection to Bayesianism.

Keywords: Bayesianism, group selection, IBE, emergent compatibilism, parsimony.

1 Introduction

There is a general question about how to relate two paradigmatic ways of talking about scientific inference. On the one hand, there is the Bayesian approach, according to which scientific inference involves determining which theories are the most probable in the light of the evidence. On the other hand, many episodes of theory evaluation in the history of science appear to be also aptly characterised as cases of Inference to the Best Explanation (IBE). What is the relationship then between IBE and Bayesianism? One possible answer is that the two are incompatible: IBE and Bayesian updating constitute two quite different rules for responding to the evidence (Van Fraassen 1989). However, it has been more common to think that the two approaches must be compatible. How then do the two approaches go together exactly?

To date, a common approach has been to look for places in the Bayesian framework where explanatory considerations could be feasibly brought in to play a role. Various philosophers have then recommended what might be called ‘explanationist Bayesianism’, according to which (depending on the version) it is either de facto the case, or reasonable, or even essential that explanatory considerations play critical roles in the Bayesian machinery (Lipton 2004; Weisberg 2009; Poston 2014). For example, one common suggestion is that explanatory considerations ‘guide’ or ‘constrain’ the assignment of prior probabilities.

An alternative way for IBE and Bayesianism to be compatible was proposed in Henderson 2014. The idea here is that, when the structure of scientific theories is taken properly into account, a Bayesian model of the preference for better explanations can be provided, making only assumptions which are quite natural and independent of any explanatory concerns. The key is to recognise that scientific theory evaluation occurs at multiple levels of generality or abstraction. A general-level or framework theory is often regarded as providing a better explanation to the extent that it explains the phenomena without too much reliance of fine-tuning of auxiliary hypotheses. For natural choices of priors, the Bayesian method of calculating likelihoods for general frameworks of theories inherently penalizes such reliance on fine-tuning. Explanatory considerations then need not constrain or guide the Bayesian machinery, particularly the choice of priors. Rather inference to the best explanation emerges automatically from the Bayesian model. I have called this view ‘emergent compatibilism’, in contrast with the ‘constraint-based compatibilism’ provided by explanationist Bayesian approaches.

In this paper, I will reexamine a case of IBE, which has been treated along explanationist Bayesian lines by Elliott Sober (Sober 1990). The case is the influential attack by the biologist George C. Williams (Williams 1966) on group selection hypotheses in evolutionary biology. Williams argues that the hypothesis that natural selection operates only at the level of individual organisms provides a better explanation of altruistic behaviour than a hypothesis also invoking group selection. This is because, on his view, invoking only individual level selection provides a more parsimonious account. Sober argues that the parsimony considerations in Williams’ argument make it reasonable to assign a higher prior to the individual selection hypothesis.

I will argue that this case is actually much better modelled according to the emergent compatibilist approach. The explanatory considerations behind Williams’ appeal to parsimony are driven by a concern to avoid fine-tuning the theory, and as such are naturally taken into account in the Bayesian likelihoods. There is no need then for explanatory considerations to govern the assignment of priors.

The reassessment of the case of Williams' argument has more general implications. Sober uses this example as part of a wider argument in favour of what we might call a 'local' view of the epistemic significance of parsimony (Sober 1990). In contrast to this view, I suggest that Williams' argument shares basic features with other cases of IBE. The Bayesian treatment of the Williams' case is also similar to the treatment of other cases of IBE, in which explanatory considerations are reflected in the Bayesian likelihood. Thus, consideration of these cases appears to point to a more general characterisation of IBE and its evidential relevance, which contrasts with Sober's localism.

2 IBE and Bayesianism

There is a long-standing puzzle over how explanatory and confirmatory aspects of scientific theories are related. On the one hand, it seems that scientific theory evaluation should be driven primarily by how well confirmed various rival theories are in relation to the evidence. Philosophers of science have proposed a number of theories about what such empirical support or confirmation consists in. Prominent among these is the Bayesian approach. In the Bayesian approach, agents' attitudes towards hypotheses in a hypothesis space $\mathcal{H}$ are taken to be degrees of belief which can be represented by a probability distribution (or density) over $\mathcal{H}$. The Bayesian starts from a prior probability distribution $p(T)$ representing initial degrees of belief in theories T in $\mathcal{H}$ and as evidence D is obtained, updates the prior by conditionalisation to a 'posterior' distribution $p(T|D)$. The posterior distribution is the conditional probability $p(T|D)$, given by Bayes' rule $p(T|D) = \frac{p(D|T)p(T)}{p(D)}$. In this equation, $p(D|T)$ is the 'likelihood' for the theory T and $p(T)$ is the 'prior'.

On the other hand, scientific arguments for one theory over another often appeal more or less explicitly to explanatory advantages of that theory. For example, Darwin argues very explicitly in *The Origin of Species* that the theory of natural selection provides a better explanation of 'large classes of facts' than special creationism (Darwin 1859 (1962). Copernicus argues that his heliocentric theory possesses greater 'harmony' than its Ptolemaic rival, which may be seen as an appeal to a kind of explanatory virtue (Copernicus 1543).

The fact that both Bayesianism and IBE appear to be broadly applicable to cases in the history of science suggests that they are compatible with one another. But then the puzzle is how to characterise their relationship. How does IBE fit in with the evaluation of theories in terms of their empirical support? Is IBE just another way of talking about empirical support? Or is it an additional component which needs to be included in a full account of theory evaluation?

As we saw in the introduction, a dominant approach to reconciling the two paradigms has been to recommend that Bayesianism and IBE should be amalgamated into an explanatorily-driven form of Bayesianism. Minimally, this position requires finding roles that explanatory considerations could play in the Bayesian framework. A stronger version requires defending the idea that explanatory considerations are somehow critical or essential for the operation of Bayesianism.

One of the key suggestions about the role for explanatory considerations is that they act as a constraint on the prior probabilities[1]. The task is then to justify giving higher priors to more explanatory hypotheses. One way to tackle this task is by seeking a general reason for the connection. Ted Poston, for example, argues that if more explanatory hypotheses are not given higher priors, inductive skepticism follows (Poston 2014). Others have advocated that simpler hypotheses should be given higher priors (eg. Jeffreys 1998), though there has also been serious opposition to this view (eg. Kelly 2007). The basic problem, critics urge, is that we have no reason to think the world is more likely to be simple than complex, so no grounds for making that assumption in the prior.

Another approach is to adopt a more local strategy. The idea is that perhaps even without a general connection between parsimony and priors, in certain specific cases, we can justify giving higher priors to more parsimonious hypotheses. In this vein, Sober argues that Williams' appeal to parsimony in his arguments against group selection hypotheses provides such a local justification for assigning them a lower prior.

However, there may be no need to defend assigning higher priors to more explanatory hypotheses either on a global or a local level. In Henderson

---

[1] There have been other suggestions too. For example that explanatory considerations help to determine likelihoods, relevant data and hypotheses for consideration in the hypothesis space (Lipton 2004), or that explanatory considerations aid in application of the principle of indifference (Huemer 2009, Poston 2014).

2014, I have proposed a different kind of Bayesian model of IBE. The key observation here is that scientific theory evaluation can be thought of as occurring at different levels. The overall aim of theory evaluation is to find the best specific theory with all its parameters filled in, which can explain the data. But it is often helpful to think of accomplishing this aim in two stages: on the one hand, finding the right theory at the general level, that is the right general principles or ‘framework’, and on the other finding the right parameter values or specific auxiliary hypotheses which allow this framework to fit or explain the data. To give a very simple example, suppose we are interested in the relationship between two variables, x and y. One question is what kind of general functional relationship exist between those variables. One ‘framework’ level theory might be that the relationship between the two is direct proportionality – i.e. $y = ax$, for some constant a. Now suppose a can take integer values from 1 to 10. Then we have 10 fitted theories produced by the framework $y=ax$ corresponding to the 10 different auxiliary hypotheses about the value for the constant a. Many theories contain open variables which require fixing before they make specific predictions. For instance, Newton’s second law of motion says there is a force, F, which is directly proportional to the product of the mass and the acceleration of a body. In a fitted theory for this framework, we fill in a particular force for the case in hand, for example a force obeying Hooke’s law if the system is a spring.

My suggestion is that a large part of what makes for better explanation is a feature of the relationship between the specific and framework levels of theory evaluation, in particular how much tuning of specific parameters or auxiliaries, is required in order for the higher level framework to produce a good empirical fit to the data (Henderson 2014). One of the key characteristics of a theory that provides a better explanation is that it can account for the phenomena without requiring that the auxiliary hypotheses be highly fine-tuned. The theory will provide a better explanation if it explains from its core principles, rather than relying on such fine-tuning.

When we build a Bayesian model, it is helpful to capture this distinction between frameworks and specific (or fitted) theories. Rather than the standard approach of placing all candidate specific theories in one big hypothesis space, we can use a hierarchical Bayesian model, in which the hypothesis space is structured into levels (Henderson et al. 2010). Frameworks can compete with one another at a high level of the model, whereas the fitted theories occupy a lower level. Bayesian updating then is carried out at each level of the model.

In order to evaluate the frameworks at the high level, the Bayesian updates a prior probability $p(F_i)$ over the frameworks to a posterior $p(F_i|D)$. The posterior depends on both the prior for the model $p(F_i)$ and on the likelihood $p(D|F_i)$. The likelihood for the framework F_i, $p(D|F_i)$ is an average over the likelihoods for the specific hypotheses h in the framework, weighted by priors for those hypotheses with respect to the framework $p(h|F_i)$. As long as the prior probability is spread roughly evenly over the hypotheses of the model, the average likelihood penalises less simple, or more fine-tuned explanations. That is because in such cases we have to average over a relatively large number of ill-fitting specific hypotheses which have very low likelihoods. The principles at work here are those that underpin the Bayesian approach to 'model selection', which is known to favour simpler hypotheses (MacKay 2003).[2]

Notice that in this type of Bayesian account, there is no need to postulate that explanatory considerations constrain the priors. There are some assumptions about how the prior probabilities are spread over the specific hypotheses in a framework, but these need not be motivated by anything to do with explanatoriness. Rather they appear to be quite 'natural' assumptions that one would make on independent grounds. Thus, this type of model provides an independently motivated Bayesian account of why better explanations are preferred in scientific inference.

Williams' argument against group selection hypotheses provides an interesting case of an inference to the best explanation. Elliott Sober has treated this case along explanationist Bayesian lines, adopting the local strategy of arguing that there are case-specific reasons to take the explanatory considerations, particularly parsimony, as a guide to the prior probabilities. I will now argue that the case is better treated along the lines of the alternative Bayesian model I have sketched above, where the explanatory considerations naturally emerge from the calculation of the likelihoods in a hierarchical Bayesian model. Thus, the example is one where explanatory considerations emerge from the Bayesian model, rather than being explicitly introduced into it. To see this, let us first review the example.

---

[2] Sometimes, this is seen by making approximations to the Bayesian likelihood, such as the Bayesian Information Criterion or BIC. The BIC contains two terms corresponding to the best-fit that can be achieved by the theory and a penalty for complexity, which increases with the number of free parameters in the theory.

3 Williams' argument against group selection

The debate over units of selection goes back to the very origin of the theory of natural selection. The theory of natural selection in evolutionary biology can be put in somewhat abstract terms as follows. If there are biological entities which reproduce with some variation in traits, such that different variants differ in fitness, and the traits are heritable, then those variants which are more fit are more likely to survive and reproduce. This statement is neutral about which biological entities are the 'units of selection'. In Darwin's original theory of natural selection, the units of selection were individual organisms. The idea then was that organisms generally inherit their traits, but they reproduce with some variation. Organisms which are better adapted to their environment are more 'fit' and out-breed organisms which are less well adapted. Thus over time, and generations, there will be a shift in the population towards organisms with better adapted traits.

However, the existence of biologically altruistic behavioural traits presents a puzzle for the theory based only on selection of individual organisms. Biologically altruistic traits are traits which increase the fitness of other organisms, whilst reducing the fitness of the altruistic organism itself. There are many examples of apparent altruism in the natural world. One is the 'wagon-training' of musk oxen[3]. This is a defensive formation that the oxen adopt when they are attacked by wolves. The larger stronger animals stand on the outside protecting the young and weaker members of the group. At first sight the theory of individual selection seems unable to explain such behaviour. Wouldn't a large musk ox who protects the weaker members of the herd do worse than a selfish ox who makes a run for it, or even stands behind the weaker members, making it more likely that they will be attacked instead of him? In general, why would an organism possessing altruistic traits have any selective advantage at the level of individual organisms, since possession of these traits is clearly detrimental to the organism's own ability to survive and reproduce?

Since Darwin, people have considered whether it might help to extend the theory of natural selection to allow not only selection between individuals, but also selection between groups. The idea, at least up to the 1960s, was that groups might themselves serve as units of selection. That is, they might reproduce, perhaps by some process of splitting, inheriting traits, and some groups might then out-compete other groups in a process of selection. So groups of altruistic musk oxen would be selected over groups

---

[3] See Sober 1990, p. 80.

of selfish oxen. Thus the reason for the existence of the trait in the male musk ox is its benefit to the group rather than to the ox himself.[4]

So we have two competing theories about the units of selection, each trying to explain the existence of biological altruism. One allows for individual organisms to serve as units of selection – call this theory IS. The other allows for selection to occur both at the individual and at the group level – call this theory GS. Prior to the 1960s, as Sober points out, the group selection theory was unpopular among population geneticists, but widely assumed to be true by field naturalists. The tide turned in 1966, when the biologist George C. Williams published a book criticizing the idea of group selection (Williams 1966).

Part of what Williams does in his book is to argue that, despite initial appearances, individual level selection can explain many cases of apparent biological altruism. Some cases, he says, can be accounted for as 'statistical effects' of individual level selection. For example, in the case of the musk oxen, what might really be going on is that the threat felt by an ox depends on its size relative to a predator. There is some threshold of predator size which determines whether the ox responds with 'counterthreat or with flight'. For predators of a certain size, larger oxen will be more inclined to stand their ground than to flee and the result is that they will end up in a more exposed position, seemingly protecting weaker members of the herd. Williams suggests that this case is one of a number of others which may also be explained in terms of different adaptive responses of individuals according to their own strengths.

In addition, Williams argues that whilst group selection is a possible evolutionary process, explanations which invoke group selection are relatively 'onerous'. Group selection hypotheses, he thinks, should only be invoked when 'simpler' (genic or individual) forms of natural selection failed to give an explanation. Williams can be taken to be making an argument that theories invoking only individual selection provide a better explanation of the biological phenomena than theories with group selection.

But what exactly is the reason that Williams sees group selection hypotheses as lacking in parsimony? Sober offers the following plausible interpretation. In a group selection explanation, the altruistic groups do better and so produce more offspring. This explains how we can have more

---

[4] Notice that this way of treating group selection should be taken in historical context. Since Williams' book, the idea of group selection has been revived in a different form, largely due to work of D. S. Wilson and Elliott Sober (Sober and Wilson 1998).

altruists. The problem is that there isn't just group selection going on. There is also still selection between organisms in a group. And the effect of that will always be to favour the selfish. Groups will experience 'subversion from within', meaning they will tend to be overtaken by selfish organisms. What this means is that if group selection is to be effective it must occur at a fast enough rate to overcome the effects of individual selection. This imposes a number of constraints on acceptable rates of group reproduction and other parameters. Thus, as Sober says, "The rational kernel of Williams' parsimony argument is that the evolution of altruism by group selection requires a number of restrictive assumptions about population structure". (Sober 1990, p. 83).

3.1 An explanationist Bayesian account

Sober argues that, in a Bayesian framework, the considerations involved in Williams' appeal to parsimony impact the prior and not the likelihood of the group selection theory.

> Williams suggests that the hypothesis of group selection, if true, would explain the observations, and that the same is true for the hypothesis of individual selection that he invents. This means ... that the two hypotheses have identical likelihoods. If so, the hypotheses will differ in overall plausibility only if they have different priors. (Sober 1990, p. 144)

Sober then provides an account of what is going on in Williams' argument which is supposed to justify the assignment of a lower prior to the group selection hypothesis, not in terms of a general argument, but for this particular case. Sober suggests that the reason that parsimony impacts on the plausibility of the hypotheses is because the foregoing argument about the restrictive assumptions on population structures amounts to a "biological judgment about the relative frequency of certain population structures in nature" (Sober 1990, p. 146). The idea is that since the population structures supporting the group selection explanation are relatively rare, we may regard the group selection hypothesis as less plausible, and hence give it a lower prior.

However, making the connection between parsimony and plausibility via an appeal to relative frequency is not so convincing, since it is not clear how we should think of this relative frequency – in particular, it is not clear what the relevant population is for the population structures. In the next section, I will suggest that in fact Williams' argument is much more naturally captured by the hierarchical Bayesian model, which does not require us to give even a local justification for a connection between parsimony and plausibility.

3.2 An emergent compatiblist account

The types of explanatory considerations that Sober has identified in Williams' argument are exactly those which can be easily modeled in the hierarchical terms described above. In this example, there are two competing frameworks: the individual selection framework in which natural selection is only allowed to operate on individual organisms, and the group selection framework in which both individual-level and group-level selection are allowed. In each case, filling in details of various parameters, such as selection rates, produces particular individual selection or group selection based fitted models.

When we ask whether the individual selection theory or the group selection theory provides the better explanation of biological phenomena, we are asking which of the two general frameworks IS or GS provides the better explanation. What Williams takes to be less explanatory about GS is that it needs to be combined with various very specific auxiliary hypotheses in order to account for the phenomena.

If it is true, as Williams alleges, that the group selection framework needs to be highly fine-tuned, this will be reflected in a lower Bayesian likelihood for that theory, since the theory will be automatically penalised for fine-tuning by the averaging process. The Bayesian then favours the more explanatory theory without assigning any higher prior probability to the individual selection theory. Indeed it is quite possible for the priors for the competing frameworks to be equal: $p(IS) = p(GS)$.

This case thus exemplifies the 'emergent compatibilist' position I argued for in Henderson 2014. According to emergent compatibilism, there is no need for explanatory considerations to play an autonomous role in the Bayesian machinery. Rather, given independently plausible choices of priors, the Bayesian favours the more explanatory hypotheses. We are then able to explain, using a Bayesian model, why explanatory considerations play a role in the inference, rather than stipulating that they do.

The explanatory considerations in Williams' argument are a feature of the relationship between framework and the auxiliaries required to produce a fully fitted theory. Therefore, the explanatory considerations are not naturally captured in Bayesian models which do not preserve the distinction between frameworks and fitted theories. In Sober's Bayesian treatment, the theories in question are fully fitted theories with parameters adjusted to account for the observations. In that account, $p(IS^*) > p(GS^*)$, and $p(D|IS^*) = p(D|GS^*)$ where * denotes that the theory is fitted.

There is no conflict in formal terms between the hierarchical model I have advocated and Sober's explanationist Bayesian model. Sober's model works with fitted hypotheses and assigns a higher prior to the individual selection hypotheses. Indeed, we can always consider the support for specific hypotheses h by marginalising out the frameworks

$$p(h|D) = \sum_i p(h|F_i, D)p(F_i|D) \quad (1)$$

If we just consider the updating of $p(h)$ to $p(h|D)$ for the specific hypotheses h, it will be as if we did weight the prior in favour of the more explanatory hypotheses. However, although it is always possible to model episodes only in terms of specific hypotheses, by so doing, one misses important aspects of the structure of the reasoning which help to justify why the prior is higher for more explanatory specific hypotheses. Considering the role of the frameworks provides an independently motivated explanation of *why* explanatory considerations play a role in inference. Equation 1 makes it clear how the support for specific hypotheses h from D depends on two components:

i) how well each framework F_i is supported by the data (as expressed in $p(F_i|D)$), and

ii) how well the specific h is supported by D, with respect to the framework F_i (as expressed in $p(h|F_i,D)$).

The first component, as we saw above, naturally reflects the explanationist concern with fine-tuning since it involves an averaged likelihood. By contrast, in Bayesian reconstructions which work with fitted models, explanatory considerations can be grafted onto the Bayesian machine by having them constrain the priors, but it remains rather opaque why they, rather than some other considerations, should be what determines the priors. On the alternative account that I am recommending, there is no need for explanatory considerations to constrain the priors explicitly. As we saw in the group selection example, the priors for the competing frameworks could even be equal.

Some assumptions about priors do play a role of course. In the calculation of the average likelihood, the likelihoods for specific hypotheses are weighted by a prior $p(h|F_i)$. Different assignments of these priors will result in different values for the average likelihood $p(D|F_i)$. However, a natural choice is a prior which is uniform over the different specific hypotheses, at least in situations where there is no background information favouring any particular hypothesis or range of hypotheses. For an objective Bayesian, this choice might be motivated by a consideration such as the principle of indifference. The important point is that, although the results do depend on assignments of priors to the specific hypotheses, there are generally reasons for these assignments, which can be motivated independently of the desire to ensure compatibility with IBE.

4 On the generality of IBE

Part of the purpose of giving a Bayesian model of IBE is to elucidate the epistemic significance of explanatory considerations by connecting them to an established account of evidential support. We have discussed two different ways of modelling IBE in Bayesian terms: the explanationist Bayesian approach of grafting explanatory considerations onto the Bayesian machinery, and the hierarchical Bayesian approach in which explanatory considerations emerge. The first approach makes Bayesianism dependent upon explanationist guidance, whereas the second treats Bayesianism and IBE as independent, but compatible, characterisations of scientific inference.

The two approaches also have quite different implications for how “local” we take IBE to be. The explanationist Bayesian model is compatible with, and indeed encourages the view that the epistemic significance of explanatory considerations such as parsimony is highly case-specific. For example, Sober has advocated the view that the epistemic significance of parsimony is highly “local”.

> When a scientist uses the idea [of parsimony], it has meaning only because it is embedded in a very specific context of inquiry. Only because of a set of background assumptions does parsimony connect with plausibility in a particular research problem. What makes parsimony reasonable in one context therefore may have nothing in common with why it matters in another. The philosopher’s mistake is to think that there is a single global principle that spans diverse subject matters. (Sober 1990, p. 140)

Sober argues that not only are the considerations which justify constraints on the prior specific to the case in cases like Williams’, but there are also other cases where the considerations behind parsimony are quite different and are better seen as reflected in the Bayesian likelihood rather than the prior. He gives the example of the appeal to parsimony by cladists in phylogenetic inference. Overall then, Sober’s conclusion is that there is nothing much in common between the cases of parsimony, and no common representation in Bayesian terms.

However, Williams' argument only appears to be a very case-specific use of parsimony if we fail to identify and model the key structural feature that it has in common with other cases of IBE. In fact, the core explanatory considerations in Williams' arguments can be analysed in terms of the relationship between different levels of theory evaluation. Specifically, in Williams' argument, what is taken to be less explanatory about the theory involving group selection is that it needs to be combined with various very specific auxiliary hypotheses in order to account for the phenomena. This kind of concern is a very common feature of other cases of IBE.

In Henderson 2014, I discussed another example. This was the case of the explanation of retrograde motion by the Ptolemaic and Copernican theories. Planets generally traverse the sky in an eastward direction from night to night. However, from time to time, they have periods of 'retrograde motion' in which they go into reverse and move westward. The Ptolemaic theory, which has the Earth at the centre of the planetary orbits, explained retrograde motion by placing each planet on an epicycle, mounted on a deferent that carries the planet in its circular orbit around the Earth. When the planet moves backwards on its epicycle with respect to the motion of the deferent, it would appear from Earth to be in retrograde motion. By contrast, according to the Copernican theory, the Sun is at the centre of the planetary orbits, including Earth's. Planets are then to be expected to be in retrograde motion whenever the Earth overtakes the planet in its orbit (or vice versa). Thus, the Copernican theory explains the phenomenon without the need to rely on the auxiliary hypothesis of an epicycle, or to account for the details of the motions by fine-tuning the rates of the epicycle orbits. This is the key reason that the Copernican theory provided a better explanation of retrograde motion than the Ptolemaic theory.

Other comparisons between theories can be understood along similar lines. As another example, consider the rivalry between the Newtonian corpuscular theory of light and the wave theory of light in the 19th century. The Newtonian theory of light treated light as a stream of corpuscles. It could explain diffraction phenomena, but it invoked different auxiliary hypotheses for different types of diffraction. The pattern of light observed on a screen as the result of shining light on a hair was explained in terms of an ether surrounding the hair whose variation in density produced the pattern on the screen. The observation of rings between two pieces of glass was explained by postulating that by passing through a medium such as glass, corpuscles of light acquire a disposition to either reflect or transmit through a subsequent surface. Newton called these dispositions ‘fits of easy transmission and easy reflection’. The corpuscular explanation of thin slit diffraction relied on yet further auxiliary hypotheses. By contrast, the wave theory of light could use the same basic principles to explain all these different phenomena of diffraction. In that sense, wave theorists could reasonably claim that it provided a better explanation of these observations than the corpuscular theory.

The concern to avoid fine-tuning also explains why scientists often claim an explanatory advantage for their theory based on its ability to explain from its basic principles, rather than requiring a number of arrangements of auxiliary hypotheses. For example, Lavoisier points out that his new oxygen theory can explain various phenomena such as combustion and calcination of metals, saying

> I have deduced all the explanations from a simple principle, that pure or vital air is composed of a principle particular to it, which forms its base, and which I have named the oxygen principle, combined with the matter of ore and heat. Once this principle was admitted, the main difficulties of chemistry appeared to dissipate and vanish, and all the phenomena were explained with an astonishing simplicity.[5]

Lavoisier claims that his theory provides simpler explanations of the different phenomena because they are all made in terms of his basic oxygen principle, and do not need to invoke particular assumptions on which the rival phlogiston theory required, such as that the phlogiston given off in combustion has negative weight (see Thagard 1978, pp. 77-78).

In all these cases, scientists are taking theories to be less explanatory if they rely too heavily on fine-tuning auxiliary hypotheses. And in all these cases, a Bayesian account is available in which the frameworks which are overly fine-tuned receive lower Bayesian likelihoods, because of the way these framework likelihoods are computed as averages of specific theory likelihoods. The basic – and generally applicable – method of Bayesian model selection lies behind this approach. In Henderson 2014, I explain how the Copernican vs Ptolemaic example may be handled in these terms. The

---

[5] Thagard’s translation (Thagard 1978). French original is: Jai déduit toutes les explications dun principe simple, cest que l’air pur, l’air vital, est composé dun principe particulier qui lui est propre, qui en forme la base, et que jai nommé principe oxygine, combiné avec la matière du feu et de la chaleur. Ce principe une fois admis, les principales difficultés de la chimie ont paru sévanouir et se dissiper, et tous les phénomènes se sont expliqués avec une étonnante simplicité (Lavoisier 1783).

phylogenetic case that Sober discusses can also be modeled in a similar way. In that case, the competing frameworks are different tree structures that represent the branching structure of the evolutionary history. The auxiliary hypotheses are the particular assumptions about the probabilistic weighting of different branches. To compute the overall support for a particular tree, we compute the average likelihood over all the possible branch weightings. As in the case of Williams' argument, the main concern is to find the tree for which the explanation of the observed traits is due primarily to the structure of the tree itself, and not to particular choices of the branch probabilities.[6]

In cases such as these, it is not that there are are no 'subject-matter-specific' assumptions or judgments at work. In particular, the identification of the framework and its parametrisation are important background assumptions in any given case. However, there is a structural commonality to the explanatory considerations involved, in terms of the relationship between different levels of theory evaluation, which makes possible a more unified connection to Bayesianism than Sober has suggested.

5 Conclusion

In this paper, we revisited Williams' argument against group selection. I argued that, contrary to the treatment by Elliott Sober, this case is not best seen as one in which the scientist is recommending that the more explanatory (more parsimonious) theory gives a reason for assigning it a higher prior probability. Rather, the core consideration in Williams' argument is a concern that group selection hypotheses rely too heavily on fine-tuning of auxiliary hypotheses in providing their explanation of biological phenomena. Since the explanatory considerations are a feature of the relationship between levels of theory evaluation, their evidential significance is obscured in a Bayesian model like Sober's which works only with a hypothesis space of specific hypotheses. However, their significance can be illuminated in a Bayesian model which distinguishes between the framework and specific theories. In such a model, we have a 'model selection' process governing the comparison of frameworks, and a 'model-fitting' process determining the right parameter values for a given framework. Model selection, both in Bayesian and non-Bayesian settings, takes into account how fine-tuned the framework is, and hence accounts naturally for the explanatory considerations in IBE. In Bayesian model

---

[6] This is discussed in detail in Sober 1988. There are a number of subtleties and controversies over the right way to do phylogenetic inference which need not detain us here.

selection, this means that less explanatory hypotheses will typically (subject to reasonable, but independently motivated, assumptions about priors) have a lower likelihood than hypotheses which provide a better explanation.

Sober has argued against the idea that there is a ‘single global principle’ behind the explanatory virtue of parsimony which transcends different subject matters. Part of the reason is that he presents examples in which, he alleges, parsimony plays quite different roles with respect to evidential support. In some cases, he suggests, parsimony is reflected in Bayesian likelihoods. In others, such as Williams’ argument against group selection, he argues that the prior probability is determined by the parsimony considerations. In my view, the kind of explanatory considerations found in Williams’ argument are a feature of the relationship between levels of theory evaluation and can also be identified in various other cases of IBE. They can also be modeled in the Bayesian framework in a similar way. All this suggests that a more unified picture of IBE and its evidential significance may be available. But this conjecture does still need to be tested by going beyond the handful of cases so far considered. More work needs to be done to see if other cases of IBE can be analysed in a similar fashion, when they are examined in detail.

References

N. Copernicus. *De Revolutionibus Orbium Caelestium*. 1543.

C. Darwin. *The Origin of Species by Means of Natural Selection*. Collier, New York, 6th edition, 1859 (1962).

L. Henderson, N. D. Goodman, J. B. Tenenbaum & J. F. Woodward. The Structure and Dynamics of Scientific Theories: A Hierarchical Bayesian Perspective. *Philosophy of Science* 77 (2): 172-200, 2010.

L. Henderson. Bayesianism and inference to the best explanation. *British Journal for the Philosophy of Science*, 65:687–715, 2014.

M. Huemer. Explanationist aid for the theory of inductive logic. *Brit. J. Phil. Sci.*, 60:345–375, 2009.

H. Jeffreys. *Theory of Probability*. Oxford University Press, Oxford, 3rd edition, 1998. First edition 1939.

K. Kelly. Ockham’s razor, empirical complexity and truth finding efficiency. *Theoretical Computer Science*, pages 270–289, 2007.

A. Lavoisier. Reflexions sur le phlogistique. *Memoires de l’Academie des Sciences*, pages 505–538, 1783.

P. Lipton. *Inference to the Best Explanation*. Routledge, London, UK., 2nd edition, 2004.

D. J. C. MacKay. *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, Cambridge, 2003.

T. Poston. *Reason and Explanation: a defense of explanatory coherentism*. Palgrave Macmillan, 2014.

E. Sober. *Reconstructing the past: parsimony, evolution and inference*. MIT Press, 1988.

E. Sober. Let’s razor Ockham’s razor. In D. Knowles, editor, *Explanation and its Limits*, pages 73–93. Cambridge University Press, UK, 1990.

E. Sober. What is the problem of simplicity? In *Simplicity, inference and modelling*, pages 13–32. Cambridge University Press, 2002.

E. Sober and D. S. Wilson. *Unto others: the evolution and psychology of unselfish behavior*. Harvard University Press, 1998.

P. R. Thagard. The best explanation: Criteria for theory choice. *J. Phil.*, 75:76–92, 1978.

B. C. Van Fraassen. *Laws and Symmetry*. Clarendon Press, Oxford, UK, 1989.

J. Weisberg. Locating IBE in the Bayesian framework. *Synthese*, 167:125–144, 2009.

G. C. Williams. *Adaptation and Natural Selection: A Critique of some Current Evolutionary Thought*. Princeton University Press, New Jersey, 1966.