

This is a manuscript accepted for publication in *Philosophy of Science*.

This version may be subject to change during the production process.

Philosophy in Science – Some Personal Reflections

Elliott Sober

Univ of Wisconsin – Philosophy, Madison UNITED STATES

Abstract: The task of *Philosophy In Science (PinS)* is to use philosophical tools to help solve scientific problems. This paper describes how I stumbled into this line of work and then addressed several topics in philosophy of biology – units of selection, cladistic parsimony, robustness and trade-offs in model building, adaptationism, and evidence for common ancestry – often in collaboration with scientists. I conclude by offering advice for would-be PinS practitioners.

The idea that *Philosophy in Science (PinS)* is a distinctive way of doing philosophy (Laplaine *et al.* 2019; Pradeu *et al.* 2021) that some but not all philosophers of science engage in presupposes a distinction between philosophical problems and scientific problems. Under the influence of Quine (1953), I was skeptical of that idea at the start of my career. It sounded to me like an untenable dualism. However, I gradually moved away from the Quinean picture, in part because of doubts about Quine's epistemological holism and doubts about his behaviorist commitments (Sober 2000). But more importantly, my change of mind was based on the realization that the dominant goal in science is for scientific theories to be evaluated by using observational evidence, whereas the dominant goal in philosophy is for philosophical theories to be evaluated by inventing arguments (not by finding new observations). The fact that this pattern has exceptions opens the door to PinS. Conceptual analysis and attention to arguments (the philosopher's bailiwick) can do good work in science. And scientific observations, along with the scientific theories that those observations justify, can do good work in philosophy, thus giving rise to SinP (= science in philosophy).

In what follows, I'll briefly describe several of the research projects I've undertaken that fall under the PinS heading. After that, I'll provide some practical advice about how philosophers of science might go about PinS research projects.

Units of Selection

The first scientific controversy in which I intervened was about group selection in evolutionary biology. At Bill Wimsatt's suggestion, I read George C. Williams's (1966) book *Adaptation and Natural Selection*. It was full of arguments against hypotheses of group selection – that a trait has evolved because it helped groups to successfully compete with other groups. I was struck by the fact that Williams's arguments were a mixed bag. At one point in the book Williams argues that the hypothesis group selection can be tested against the hypothesis of individual selection by looking at sex ratio (the frequencies of males and females in a species). The group selection hypothesis predicts that the sex ratio should be female-biased when there is lots of food and male-biased when food is scarce. This is because female bias (wherein a population has the smallest number of males consistent with all females reproducing) maximizes group productivity, whereas male bias is a way of reducing a population's productivity (so that it avoids overexploiting scarce food resources). The individual selection hypothesis, on the other hand, predicts that the sex ratio should be even. This is because if the sex ratio were uneven, individual selection would favor the over-production of the minority sex. Williams says that the data we have on sex ratios indicates that sex ratios are always even, so he declared a victory for individual selection. Williams was unaware that female-biased sex ratios occur in numerous species, but the argument he made was otherwise a very good piece of reasoning.

The oddity about Williams's book is that he offered other arguments that were of a decidedly different character. They are more or less *a priori*. For example, he claimed that (1) group selection hypotheses are unparsimonious, and that (2) only genes can be units of selection (organisms can't, and neither can groups) because a gene can exist through numerous generations whereas an organism or a group usually has a much shorter lifespan. The first of these arguments makes sense if you compare "Trait T evolved because of group and individual selection" with "Trait T evolved solely because of individual selection," but the philosophical question then arises of why parsimony is a guide to truth. The argument also raises the question of why one shouldn't compare "trait T evolved by individual and group selection" with "trait T evolved solely because of group selection." As for argument (2), the type/token distinction helps show that the argument is flawed. Gene *tokens* are evanescent, but gene *types* can be exemplified over long stretches of time; the same can be said of organism and group tokens and organism and group types.

Williams's book, and the work of W.D Hamilton (1964) and John Maynard Smith (1976) inspired Richard Dawkins to write his popularization of the anti-group selection position, *The Selfish Gene* (Dawkins 1976). Williams, Maynard Smith, and Dawkins were bitterly opposed to the idea of group selection. They thought that hypotheses of group selection aren't just factually incorrect; they are, in addition, instances of sloppy thinking, affronts to clear-headed science. To my surprise, there were other biologists who took a very different position. Lewontin (1970) took group selection seriously as a form of natural selection that is conceptually coherent, and sometimes empirically correct. Wilson (1975) did the same, and developed the idea of trait-groups to flesh out how group selection models should be built. And Michael Wade (1978)

published a beautiful essay arguing that extant arguments against the effectiveness of group selection were loaded with assumptions that *a priori* bias the case against group selection.

Seeing that there were two sides to this question, not just one, and seeing further that the anti-group selectionists made a host of questionable arguments, I started collaborating with Lewontin and with Wilson. In both cases, conceptual clarification was a central part of our project. What defines a group in the sense relevant to this problem? What does group selection mean? How is the concept of group fitness to be understood? Why are unconditional arguments against group selection flawed? I spent my first sabbatical in Lewontin's lab, which was a life-changing experience. He and I worked on a paper together (Sober and Lewontin 1982). Wilson and I worked together (mostly via email), writing several papers (Sober and Wilson 1989, 1994a, 1994b, 2011), and eventually writing a book on the subject (Sober and Wilson 1998).

2 Cladistic Parsimony

While I was in Lewontin's lab (1980-1981), I learned about cladistic parsimony from John Beatty. Cladistic parsimony is a methodology used in evolutionary biology to infer phylogenetic trees. A war was going on then – between cladists and those who thought that phylogenetic inference was a problem in statistics whose solution was to be sought by using the concept of likelihood.¹ Cladists sometimes defended their position by invoking Popper's idea of falsification; they argued that the most parsimonious tree is the one that is least falsified by one's observations. At other times, they argued that more parsimonious trees have greater explanatory

¹ It is standard in statistics to distinguish the probability of a hypothesis H from its likelihood. Where O is an observation statement, the former is the probability of H given O ($\Pr(H|O)$) while the latter is the probability of O given H ($\Pr(O|H)$). Likelihoodists are anti-Bayesians because they want to avoid talking about the probabilities of hypotheses.

power, and that this was a reason to believe those most parsimonious trees. Likelihoodists like Felsenstein (1978) fought back, arguing that parsimony is statistically inconsistent – that using parsimony to infer a phylogenetic tree can guarantee that your inference will converge on a false hypothesis the more data you obtain. Felsenstein further suggested that this problematic situation does not arise in the case of likelihood inference. Here was a controversy among scientists about the evaluation of inference procedures – a worthy topic for a philosopher of science!

I had written my Ph.D. dissertation about parsimony (though I was unaware of its role in phylogenetic inference) under the influence of Popper's insistence that thinking about the probabilities of theories in science is a dead end. After I finished the dissertation, I felt ready to move on to different research projects, but learning about the parsimony wars in evolutionary biology got my attention. The use of likelihoods interested me. Reading Hacking (1965) and Edwards (1971) was a real eye-opener. I attended Hennig Society meetings (Willi Hennig was the founding father of cladistic taxonomy) and was sympathetic with their point of view, but doubts gradually started to creep in. I thought that cladists' use of Popperian falsification was a mistake, since phylogenetic tree hypotheses do not have deductive implications about what traits one should find in extant species. And I gained some sympathy for the likelihoodist approach, though I was troubled by the fact that using likelihoods requires that one make assumptions about the evolutionary processes at work in lineages. Cladists had hoped to do without such assumptions, and I sympathized with their desire to do so. The result of this thinking was my 1988 book *Reconstructing the Past*. After Malcolm Forster and I wrote our first paper about the Akaike Information Criterion (Forster and Sober 1994), I saw how AIC can be used in phylogenetic inference, allowing one to consider several process models, not just one, in the

evaluation of multiple phylogenetic trees. I published an article on this in *Systematic Biology* (Sober 2004) and then discussed it in my book *Ockham's Razors* (Sober 2015).

I mentioned in the previous section how the familiar and unremarkable type/token distinction revealed a flaw in one of Williams's arguments against group selection. Something similar happened in my work on phylogenetic inference. As philosophers, we are familiar with the idea that proposition X presupposes proposition P only if X entails P. The presuppositions of X are propositions that need to be true if X is to be true; the fact that proposition S suffices for X's being true doesn't show that X presupposes S. Presuppositions are necessary conditions, not sufficient conditions. Felsenstein (1973) showed that a sufficient condition for the most parsimonious phylogenetic tree (given a data set) to be the tree that has the highest likelihood is that evolutionary changes in branches should be very improbable. Many biologists interpreted that result to mean that parsimony presupposes that changes are improbable. Tuffley and Steel (1997) then showed that a very different model entails that the most parsimonious tree is the tree with the highest likelihood; their model assumes that each trait evolves by drift, but the probabilities of changes on branches can be large. Combining Felsenstein's result with that of Tuffley and Steel, I argued that cladistic parsimony does *not* presuppose that change is rare. The two models each provide a sufficient condition for parsimony and likelihood to be ordinarily equivalent, which means that neither model is a necessary condition for that equivalence.

3 Robustness and Trade-offs in Model Building

While I was in Lewontin's lab, I met Steven Orzack and we got to talking about Richard Levins's (1966) much discussed paper "The Strategy of Model Building in Population Biology." Levins was a theoretical ecologist, and he wrote this paper for his peers, but it could easily have

appeared in a philosophy of science journal. Levins claimed that there are three properties of a scientific model that are in conflict with each other – realism, generality, and precision. He says that you can't maximize all three, but many of his readers thought he was saying that increasing any two of them would require in diminution of the third. He also claimed that scientists are able to discover truths about nature by considering numerous models that are all false; the way to do this is by finding “robust theorems,” which are propositions that are entailed by each of the models considered. Orzack and I came to think that Levins never gave clear definitions of the three model characteristics, nor did he provide a good argument for his conflict thesis. A similar puzzle struck us about his robustness thesis – how can deducing proposition R from each of several false models be evidence for R's truth? We published these skeptical thoughts in *Quarterly Review of Biology* (Orzack and Sober 1993).

4 Adaptationism

Orzack and I also worked on adaptationism, an *ism* that was energetically debated in the 1970's and 1980's, in large measure because of its connection with the sociobiology controversy. The question is how “important” natural selection was in the evolution of the phenotypic traits of the organisms we now observe. In a paper that appeared in the *American Naturalist*, Orzack and I drew a three-way distinction about the role that natural selection might play in the evolution of a trait T in a lineage:

(U) Natural selection played some role in the evolution of T.

(I) Natural selection played an important role in the evolution of T.

(O) Natural selection was the only important cause of the evolution of T.²

These propositions are listed in order of their logical strength. We argued that critics of adaptationism often endorse U and I as claims about most of the phenotypes found in nature; what they reject is proposition O as a claim about most such phenotypes. O is the distinctively adaptationist claim. Proposition I, or its generalization to most phenotypes, is endorsed by Gould and Lewontin (1979) in their famous paper “The Spandrels of San Marco.” If these arch anti-adaptationists endorse proposition I, the shooting match must be about proposition O.

Adaptationism as just described is a claim about nature, but there is a sister thesis that is purely methodological. It says that in considering why a trait evolved, it is essential to consider explanations that invoke natural selection. Orzack and I discuss this thesis of *methodological* adaptationism in our paper by focusing on optimality models. Anti-adaptationists often claimed that optimality models are not only false – they are pointless. Orzack and I argue that optimality models are needed. If you are going to reject proposition O as a claim about trait T, you need a model that describes what trait T would be like if selection were the only important cause of its evolution.

In another paper, Orzack and I developed a method for testing whether the evolution of a single trait in each of several different lineages was influenced by phylogenetic inertia, also known as “ancestral influence” (Orzack and Sober 2001). This happens when a descendant has a higher probability of exhibiting trait T, owing to the fact that its ancestor also had trait T. This can happen even if T would be better for the descendant than T’. Phylogenetic inertia is a

² This is a slight modification of what Orzack and Sober (1994) say. It’s the one I used in Sober (1993).

possible cause that can render strong adaptationist claims false. We called this inference procedure “the method of controlled comparisons.”

5 Evidence for Common Ancestry

Mike Steel and I published a series of papers on the epistemology of common ancestry (Sober and Steel 2002, 2011, 2014, 2015, 2017). They concern how one should test hypotheses of common ancestry, which traits provide the strongest evidence for common ancestry, and the sense in which the evolutionary process destroys information. Some of these papers appeared in science journals, others in philosophy of science journals. I think there is no good reason why these different papers should have found these different homes. The topics we addressed strike evolutionary biologists as obviously relevant to science, and they also wear their philosophical credentials on their sleeves. Mike and I were pleased when Theobald (2010) picked up on our suggestion about how hypotheses of common ancestry should be tested.

6 Practical Advice

As you’ve seen in the preceding sections, the work I’ve done that counts as PinS has often involved intervening in scientific controversies, and it has often involved collaboration with scientists. I recommend both of these strategies for philosophers of science who want to do PinS work.

Philosophers of science often want scientists to pay attention to their ideas, but they know from experience that it is often hard to get scientists to do this. Scientists are busy, and scientists often think, rightly or wrongly, that philosophy can’t do them any good in their scientific work. To break through this barrier, one strategy is to find disagreements among scientists that you

judge to have a philosophical component. If you intervene in this scientific disagreement, criticizing one side and supporting the other, this raises the probability that scientists will read your work; they can't summarily dismiss it as irrelevant to what they care about. And if you criticize both sides, you may get attention from both. The point is that diving into a pre-existing scientific controversy means that the scientists in question already have a stake in the issues you're writing about. You may be welcomed as a friend by one side and labeled a foe by the other. Or perhaps your critical conclusions will lead both sides to hate you. And, of course, you may simply be ignored; indeed, that seems to be the fate of most papers, whether or not they are instances of PinS research.

Notice that my advice to do PinS by finding scientific controversies that have a philosophical dimension is orthogonal to the philosophical task of discovering presuppositions that all scientific theories are obliged to make. If scientists don't disagree about whether X is true, it probably won't interest them to hear from philosophers who provide arguments, even good arguments, for thinking that X is true. I say "probably" here, not *for sure*. And if you argue that a universal assumption in science is mistaken, the odds are that scientists will dismiss your claim out of hand without bothering to look at the details of your ingenious argument.

If you want to follow this interventionist strategy, you could do it solo, but you might consider working together with a scientist who has a stake in the scientific controversy. How to do this? Read the scientific literature on a current controversy, develop some ideas about how philosophy is relevant to it, and contact a scientist whose work seems to intersect with the work you're doing. See if you can strike up a conversation on this. Don't send the scientist a long paper out of the blue. If the scientific controversy involves defenders of X arguing with defenders of Y, it may be best to send your email to someone whose bottom line you agree with,

at least approximately, and maybe start off by discussing with that person what is wrong about the other side. However, you've got to do something more than just agree with scientists to get their attention. You've got to produce something new – for example, a new pro-X argument, or a new criticism of the Y side.

Although I have carried out several PinS collaborations via email, I want to emphasize the value of spending extended periods of time together with scientists. My sabbatical in Lewontin's lab was a life-changer for me.

What tools should PinS researchers have in their pockets? I've mentioned the type/token distinction and the concept of presupposition, but these are just examples that happened to come in handy once upon a time. There are many other ideas that are familiar to philosophers that might also be helpful. But there is a broader remark that is as obvious as it is important: you've got to understand the science you are taking as your subject. If the science makes use of probability, you need to understand enough about probability to follow what is going on. Taking a course in pure mathematics is probably not the best thing to do here, nor are most philosophy of probability courses what you need; it would be better to attend a methods course in the science in question. Methods courses, however, are often cookbook in character. They often don't delve much into the justifications of various methods. This can be frustrating for a philosopher, but it gives you hands-on experience in the techniques that scientists actually use. As a philosopher, this may lead you to wonder about the justification of those standard methods – a topic that often won't interest practicing scientists very much, but that doesn't mean that they shouldn't interest philosophers!

One of the benefits of working on PinS projects is that it helps you develop ideas that fall under the heading of SinP. PinSers use philosophy to solve scientific problems whereas SinPers

use ideas from science to do philosophy. To show how PinS research can generate ideas that fall in the SinP category, I want to mention a few examples of how my work on cladistic parsimony led me to some more general thoughts about the principle of parsimony. Philosophers sometimes complain that it isn't clear what parsimony means, the implication being that until a unique meaning is identified, the question of justification cannot be considered. I think this is backwards. Cladistic parsimony makes it very clear what parsimony means in a specific biological context. We can consider what its justification is. If parsimony means something different in some other area or problem, we can consider its epistemic credentials there too. And if parsimony has several candidate meanings in a given context, we can try out each, evaluating whether it is epistemically relevant. Another example concerns David Lewis's (1973) claim that qualitative parsimony is epistemically relevant but quantitative parsimony is not. For Lewis, what matters is how many *kinds* of objects a theory postulates, not how many *token* objects it postulates. Counterexamples to Lewis's claim are obvious when you consider phylogenetic inference. In inferring whether two token organisms or groups of organisms have a common token ancestor, the common ancestry hypothesis is more parsimonious than the hypothesis of separate ancestry. Indeed, there is an argument inspired by Reichenbach's (1956) principle of the common cause that shows that the common ancestry hypothesis has the higher likelihood (Sober 1988, 2015). This shows that Lewis's interpretation of the principle of parsimony is too narrow. And still another philosophical pronouncement falls by the way once you look at science. This is the idea that the principle of parsimony has no justification. What is true is that it has no universal and unconditional justification. However, given assumptions that make sense in a given research context, justifications are often available. In addition to inferences about common ancestry in biology, there is the task of estimating a model's predictive accuracy in

statistics; in the framework of AIC, the number of adjustable parameters a model has is relevant to that estimate (Forster and Sober 1994).

My final comment is a warning. I've found that papers I've worked on that fall in the PinS category, whether or not they were co-authored with biologists, have more chance of getting read by scientists if they appear in science journals. This is not surprising. Note, however, that if you publish a paper in a science journal, your colleagues in philosophy who are not philosophers of science may dismiss it, thinking that what you've done is science, not philosophy. This might hurt your career. However, the risk will be diminished if you combine your PinS work with SinP work; the latter will make your philosophical *bona fides* clear, even if the former raises suspicions.

Acknowledgments: I thank Paul Griffiths, Maël Lemoine, Deborah Mayo, Thomas Pradeu, and Randolf Nesse for useful discussion of PinS at our PSA-2021 symposium on that topic.

References

Dawkins, R. (1976). *The Selfish Gene*. Oxford University Press.

Edwards, A. (1971). *Likelihood*. Cambridge University Press.

Felsenstein, J. (1973). "Maximum Likelihood and Minimum-Step Methods for Estimating Evolutionary Trees from Data on Discrete Characters." *Systematic Zoology* 22: 240-249.

Felsenstein, J. (1978). "Cases in which Parsimony and Compatibility Methods can be Positively Misleading." *Systematic Biology* 27: 401-410.

- Forster, M. and Sober, E. (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: 1-36.
- Gould, S. J., and Lewontin, R.C. (1979). "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme." *Proceedings of the Royal Society of London B: Biological Sciences* 205 (1161): 581–98.
- Hacking, I. (1965). *The Logic of Statistical Inference*. Cambridge University Press.
- Hamilton, W.D. (1964). "The Genetical Theory of Social Behavior, I and II." *Journal of Theoretical Biology* 7: 1-52.
- Laplane, L., Mantovani, P., Adolphs, R., Chang, H., Mantovani, A. McFall-Ngai, M., Rovelli, C., Sober, E., Pradeu, T. (2019) "Why Science Needs Philosophy." *Proceedings of the National Academies of Science USA* 116(10): 3948-3952.
- Levins, R. (1966). "The Strategy of Model Building in Population Biology." *American Scientist* 54: 421-431.
- Lewis, D. (1973). *Counterfactuals*. Basil Blackwell.
- Lewontin, R. (1970). "The Units of Selection." *Annual Review of Ecology and Systematics* 1: 1-18.
- Maynard Smith, J. (1976). "Group Selection." *Quarterly Review of Biology* 51: 277-283.
- Orzack, S. and Sober, E. (1993). "A Critical Examination of Richard Levin's 'The Strategy of Model Building in Population Biology'." *Quarterly Review of*

Biology 68: 533-546.

Orzack, S. and Sober, E. (1994). "Optimality Models and the Test of Adaptationism."

American Naturalist 143: 361-380.

Orzack, S. and Sober, E. (2001). "Adaptation, Phylogenetic Inertia, and the Method of

Controlled Comparisons." In S. Orzack and E. Sober (eds.), *Adaptationism and*

Optimality, Cambridge University Press, pp. 45-63.

Pradeu, T., Lemoine, M., Khelifaoui, M. and Gingras, Y. (2021). "Philosophy in Science: Can

Philosophers of Science Permeate through Science and Produce Scientific Knowledge?"

The British Journal for the Philosophy of Science, May. <https://doi.org/10.1086/715518>.

Quine, W. (1953). "Two Dogmas of Empiricism." In *From a Logical Point of View*. Harvard

University Press, pp. 20-46.

Reichenbach, H. (1956). *The Direction of Time*. University of California Press.

Sober, E. (1988). *Reconstructing the Past – Parsimony, Evolution, and Inference*. Cambridge

University Press.

Sober, E. (2000). "Quine's Two Dogmas." *Proceedings of the Aristotelian Society* 74: 237-280.

Sober, E. (2004). "The Contest between Likelihood and Parsimony." *Systematic Biology* 53: 6-

16.

Sober, E. (2015). *Ockham's Razors*. Cambridge University Press

Sober, E. and Lewontin, R. (1982). "Artifact, Cause and Genic Selection." *Philosophy of Science*

49(2): 157–180.

- Sober, E. and Steel, M. (2002). “Testing the Hypothesis of Common Ancestry.” *Journal of Theoretical Biology* 218: 395-408.
- Sober, E. and Steel, M. (2011). “Entropy Increase and Information Loss in Markov Models of Evolution.” *Biology and Philosophy* 26: 223-250.
- Sober, E. and Steel, M. (2014). “Time and Knowability in Evolutionary Processes.” *Philosophy of Science* 81: 537- 557.
- Sober, E. and Steel, M. (2015). “How Probable is Common Ancestry According to Different Evolutionary Processes?” *Journal of Theoretical Biology* 373: 111-116.
- Sober, E. and Steel, M. (2017). “Similarities as Evidence for Common Ancestry – A Likelihood Epistemology.” *British Journal for the Philosophy of Science* 68: 617–638.
- Sober, E. and Wilson, D. (1989). “Reviving the Superorganism.” *Journal of Theoretical Biology* 136: 337-56.
- Sober, E. (1993). *Philosophy of Biology*. Westview Press; 2nd edition 2000, Taylor and Francis.
- Sober, E. and Wilson, D. (1994a). “A Critical Review of Philosophical Discussion of the Units of Selection Problem.” *Philosophy of Science* 61: 534-555.
- Sober, E. and Wilson, D. (1994b). “Reintroducing Group Selection to the Human Behavioral Sciences.” *Behavior and Brain Sciences* 17: 585-654.
- Sober, E. and Wilson, D. (1998). *Unto Others – the Evolution and Psychology of Unselfish Behavior*. Harvard University Press.

- Sober, E. and Wilson, D. (2011). "Adaptation and Natural Selection Revisited." *Journal of Evolutionary Biology* 24: 462–468.
- Theobald, D. (2010). "A Formal Test of the Theory of Universal Common Ancestry." *Nature* 465(7295): 219-222.
- Tuffley, C. and Steel, M. (1997). "Links Between Maximum Likelihood and Maximum Parsimony under a Simple Model of Site Substitution." *Bulletin of Mathematical Biology* 59: 581-607.
- Wade, M. (1978). "A Critical Review of Models of Group Selection." *Quarterly Review of Biology* 53: 101-114.
- Williams, G. C. (1966). *Adaptation and Natural Selection*. Princeton University Press.
- Wilson, D.S. (1975). "A Theory of Group Selection." *Proceedings of the National Academies of Science USA*. 72: 143-46