Pseudo-solutions to the demarcation problem

Abstract
Although many philosophers have given up on the idea that there are any necessary and sufficient conditions for demarcating science from pseudo-science, some philosophers have recently championed the idea that 'science' might be thought of as in terms of Wittgensteinian family resemblance. In this paper I argue that this proposal is inherently problematic and also does not do justice to the problem of demarcation, which is inherently a normative problem. Instead, I argue for a solution according to which the meaning of terms such as 'science' is determined by a paradigm or basic predicate the properties of which must be fully specifiable (at least in principle). I argue that testability is not one of these properties; it is a red herring and should not figure at all in the correct list of demarcation criteria.

1 Introduction

What distinguishes science from non-science or pseudo-science? This is the venerable 'demarcation' problem in the philosophy of science. While vividly discussed in the 1960s and 70s by the likes of Karl Popper and Thomas S. Kuhn (1970), the problem has all but vanished from most philosophers' current concerns, which is much more specialised on matters concerning the particular sciences.[1] In many ways, this is regrettable. The demarcation problem is arguably the issue with the widest public interest and appeal. It is here where philosophers of science can have a real impact on scientific and public disputes. Also, it seems almost embarrassing that philosophers of *science* have so much to say about the

[1] There are some notable recent exceptions. See e.g. Pigliucci and Boudry (2013).

foundational issues in physics, biology, chemistry, and the social sciences, but so little about science per se.

The demarcation problem is of tremendous societal importance: to understand whether or not something is a science will influence our decisions to trust or mistrust that source of information and to act accordingly. For example, to understand that climate modelers use scientific methods to generate their conclusions will influence one's decision to try to reduce one's carbon footprint or to support politicians that promise to invest in renewable energies. One of course need not oneself have the competence to evaluate the status of those methods. Instead, one can defer to the experts. Still, one accepts implicitly that the method used by climate scientists are scientific; why else trust *their* words rather than the climate change deniers'? But what are the criteria that underlie our judgement of a method being scientific? In other words, what are the criteria by which we deem a method a trustworthy, scientific one, which is capable of producing reliable knowledge about the world?

Any particular solution to the demarcation problem will also have comprehensive consequences for the way we educate our children. Particularly in the U.S.A., educators have to deal with people who believe that creationism should be taught alongside the theory of evolution or even instead of it. Educators denying this honour to creationism – at least in part – is based on the judgement that creationism does not satisfy the same scientific standards as the theory of evolution. No matter how intuitive this judgment, though, the criteria that underlie it need to be explicated for them to have any normative force – as indeed they have been in a number of famous court trials.[2]

[2] In several court cases in the USA, in which it was debated whether creationism may be taught at high schools. One of these, namely *McLean v. Arkansas Board of Education* (1981/2) gained special prominence amongst philosophers due to the court's consultation of the philosopher of biology Michael Ruse as expert witness. Laudan (1982) published a critical reply to the judgement based on Ruse's report.

The question of what is scientific and what isn't does not only play a role between those who are proficient in science and those who are not. It in fact also has a role to play in the higher echelons of science itself. The perhaps most prominent recent example is String Theory and the idea of the multiverse, whose scientific status has been discussed within the scientific community itself (Smolin 2007, Woit 2011, Ellis and Silk 2014).

This paper is structured as follows. In Section 2 I put the target of the demarcation problem into sharper focus. In Section 3 I consider Laudan's influential claim that the demarcation problem is a pseudo-problem with no solution. Laudan's claim that necessary and sufficient conditions for science can be had, has given rise to the view that science might best be understood in terms of Wittgenstein's family resemblance. I criticize this view in Section 4. In Section 5, I argue that the meaning of 'science' is best understood in terms of a paradigm or a basic predicate, whose properties must in principle fully specifiable. In Section 6 I argue that falsifiability should not figure on the set of correct demarcation criteria. I conclude the paper in Section 7.

2 The contrast: non-science or pseudo-science?

Before we can start to discuss any proposed solution to the demarcation problem, we should get clear on the contrast class we intend to pick out. One contrast class of science is non-science. A big class indeed! However, usually non-science here is understood more narrowly as (at least seemingly) methodologically radically different intellectual activities such as metaphysics, religion, and literature (Popper 2002, 39). But even this narrower class is somewhat ill-defined, as many people believe that mathematics is not a science either (because of its non-empirical methods) (ibid.). Pseudo-science, on the other hand, seems better delineated: it refers to all those intellectual activities whose practitioners believe that they are able to produce insights on a par with scientific ones, when actually that is not the case. Clearly, climate change denialism, creationism (sometimes also called 'intelligent design'), Chinese medicine, homeopathy, and astrology all satisfy this characterization of pseudo-science.

Although some commentators have characterized pseudo-science as an activity that *pretends* to be scientific (Ladyman 2013, Hansson 2017), I think we need to be careful. Pretense would presuppose an intention to deceive, which I think usually is not the case for pseudoscience; usually practitioners of pseudo-science believe in earnest that their methods can compete with those of science and in fact view the methods of science disparagingly. They thus wouldn't have any motivation to mimic the methods of science. Instead, pseudoscience might be better characterized as intellectual activity which seeks to generate understanding and knowledge about the world without using the appropriate means for doing so. The question of what the appropriate means are for generating knowledge about the world is at the heart of the demarcation problem.

3 Laudan's deflationism

Laudan (1983) once argued that a successful solution to the demarcation problem would require specifying necessary and sufficient conditions. His reasons were as follows. If a demarcation criterion provides only necessary conditions N for scientificity, then we can say on the basis of N that something like homeopathy is not a science when it does *not* satisfy N, but we cannot say whether something like physics is a science when it *does* satisfy N. On the other hand, if we are given only sufficient conditions S for scientificity, then we can say that something like physics is a science when it meets S, but we cannot say that something like homeopathy is not scientific when it does not satisfy S. Since Laudan did not see any plausible candidate solutions to the demarcation problem that would successfully specify necessary and sufficient conditions, he concluded that the demarcation problem was "spurious" and a "pseudo-problem" (ibid, 124). He recommended that we should therefore dispense with terms like 'unscientific' entirely (ibid., 125).

Instead of trying to look for necessary and sufficient conditions (in vain), Laudan suggested, we can contend ourselves with sorting good from bad theories via *empirical tests*. For example, we may identify creationism as a false theory by carrying out empirical tests of its prediction of the young age of the earth (e.g. via carbon dating). Such a deflationist strategy, however, cannot be fully satisfactory, as it would not allow us to draw the line that deserves to be drawn between creationism and evolution: these two theories simply do not have the same scientific status. Creationism fares worse than evolution not only because it is a theory that makes more false claims about the world than the theory of evolution. That would also be true for classical Mendelianism as compared to molecular genetics. No, there seem to be elements in creationism that have no place in a scientific theory. For example, no serious contender for a scientific theory should postulate supernatural powers for solving the problems in its domain. But Laudan's deflationist strategy would not allow us to raise such issues; it would limit us to empirical testing. Such a restriction seems particularly problematic in educational contexts: we should not want to grant creationism the appearance of a legitimate (but false) scientific theory. In such contexts, Laudan's deflationism lacks the normative force of accounts that do not try to dissolve the demarcation problem.

4 Demarcation and Wittgensteinian family resemblance

Recently, various authors have argued that – contra Laudan – the demarcation problem can be tackled without presupposing any necessary and sufficient condition for what it is for something to be a science. A very popular solution several authors have proposed is to treat the sciences as a class of activities held together by Wittgensteinian family resemblance (Dupré 1993, Irzik and Nola 2011, Hoyningen-Huene 2013, Pigliucci 2013). In what follows, I argue that the concept of family resemblance is incoherent and, just like deflationism, lacks normative force.

Wittgenstein (1953) realised that many of the general terms or kind terms which we use in our language denote sets of entities which are only loosely bound together by similarity; there is nothing that they all have in common. For example, we denote a large variety of activities "game": card games, board games, ball games, computer games, schoolyard games, hide-and-seek, and even mind games. It seems impossible to come up with a definition of the term 'game' that would unify all these activities and spell out necessary and sufficient conditions. Instead, Wittgenstein concluded, all we can say in such instances is that there is a "complicated network of similarities overlapping and criss-crossing".

There is a famous problem with the concept of family resemblance, known as the problem of 'wide-open texture' and was first raised by Richman (1962): if, as the idea of family resemblance has it, things denoted by a kind term are individuated on the basis of similarity, and since, as Goodman (1972) pointed out, 'anything is similar to anything else', then on what basis can we "*refuse* to apply the term to anything"? Boxing and street fighting are similar in many ways (in both activities fists are thrown in order to harm the opponent), yet only one of them is a sport (Pompa 1967). So if it is similarity that binds together sets of things in family resemblance, on what basis would we refuse to subsume street fighting under the term 'sport'?

The problem does not only concern the negative delimitation of kinds; it also concerns their positive delimitation. My brother and I are similar: we are both tall and have protruding ears. Does that make us members of the same family? No. My brother and I are members of the same family in virtue of our hereditary relations. Likewise, there are many similarities between the sciences which are not suitable for justifying membership in the science kind. For example, physics and chemistry are similar in that most of their practitioners are white men. Yet we would rightly consider this similarity as outright irrelevant to the question of what it is that makes them members in the science kind. There has got to be something else that picks out the *right* similarities for us. Arguably, it is the kind identity itself that determines which similarity relations are relevant to the individuation of the kind and which ones aren't (Pompa 1967). That is, in the case of biological families, which are defined in terms of hereditary relations, it is those hereditary relations which determine that a resemblance (e.g., my and my brother's protruding ears) is a resemblance within a particular family rather than between members of different families (e.g., my and Barack Obama's protruding ears). In other words, it seems we must define the kind identity *before* we can single out certain similarities that can justify kind membership.[3] But then, what is the work that family resemblance does for us? Given that these issues at the core of the idea of family resemblance, it seems highly dubious that it could form the basis of a successful solution of the demarcation problem. In particular, it is hard to see how family resemblance could have any normative force as in: one ought to use scientific methods, not pseudoscientific ones when investigating the world.

5 The paradigm / basic predicate solution

Simon (1969) insists, contrary to Wittgenstein, that 'when we do use a word to apply to things that share no common essential feature, we are using it in different senses' (412). In other words, Simon considers it a necessary condition that a set of entities have something in common for us to be justified in using a general term with the *same meaning* for that set. That is, the meaning of the term 'game' would be a different one when we apply it to such different activities as chess and schoolyard ball games (the former involves a strict set of rules, the latter doesn't). But, of course, those different meanings of the word game must somehow be related in order for them to be recognizable as being subsumed by the same term (Simon also speaks of a 'family of meanings').

---

[3] Pigliucci (2013), in his use of the family resemblance concept, suggests that theoretical understanding and empirical knowledge might be two dimensions along which the sciences and the pseudosciences might be grouped. Whereas the pseudosciences are at the lower end of those scales, the mature sciences are at the top of it. It was suggested to me that Pigliucci therefore does provide relevance criteria for similarity relationships for the family resemblance of science. But what is needed to address the demarcation problem are *methodological* criteria. On Pigliucci's proposal a perfectly scientific but very young endeavour which has hitherto yielded very little understanding and empirical knowledge, would be ruled unscientific.

Simon proposes that there is a paradigm example of each general term or kind term which establishes commonality of different meanings of a single such term.[4] Say there is an activity *a* which is a G1 and an activity *b* which is a G2, and *a* and *b* share no common features; then *a* and *b* can still be recognized as instances of G (which includes both G1 and G2) because there is a paradigm example of G, namely *g*, which has a feature in common with both *a* and *b*. The features of the paradigm *g* therefore delineate the set of all activities that can be a G. However, the features of *g* do not set out necessary and sufficient conditions for all members of G, since there clearly are members, such as *a* and *b*, which do not share any features (see Figure 1).

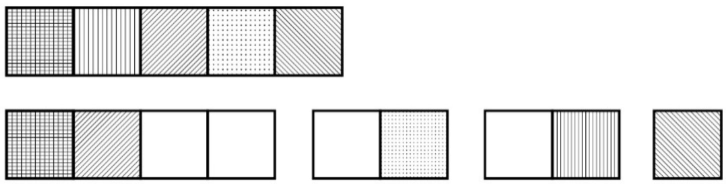

Figure 1: The paradigm / basic predicate view. The upper set of boxes represents a kind paradigm or basic predicate. The lower sets of boxes represent instances of the kind which either have very little or even nothing in common and nevertheless each share at least one feature with the paradigm/basic predicate.

In the case of science, Simon's paradigm idea could translate as follows. Suppose activity *a* is an instance of biology, activity *b* is an instance of psychology, and activity *c* is an instance of chemistry. Even though biology, chemistry, and psychology are very different, and even though there might not even be one thing they all share, they would still be recognizable as science because they all have at least one property they share with the paradigm science (e.g., physics) or with the basic predicate. As Simon puts it,

> [w]e do not in fact insist that a discipline cannot properly be called a science, or a procedure be called scientific, if it lacks any one of the features of being experimental, mathematical, systematic, and painstaking; and yet we

[4] I take it that Simon's view of a kind-defining paradigm has affinities to Eleanor Rosch's famous psychological theory of human categorization, namely 'prototype theory'.

> demand all of these characteristics (and perhaps other besides) of a paradigm, such as physics. (Simon 1969)

For Simon, the paradigm does not actually have to exist; it may just be 'conceived'. This proposal, made already by Pompa (1967), is known as the 'basic predicate solution' (cf. Bellaimey 1990). We may thus refer to this solution as the paradigm/basic predicate solution or the PBP solution, for short. An advantage of this solution is that it is not tied to any particular science: even our best example of science might not incorporate all the features that we would want to associate with science.

It is worth noticing that even though there are no necessary and sufficient conditions for a kind membership according to the PBP solution, there is still a set of properties that is necessary and sufficient for the PBP *itself*. Since the PBP individuates the kind, the solution may thus be called an 'essentialism in disguise'. Accordingly, a change in the PBP will also change the set of instances that are picked out by it, which in turn should amount to a change in the meaning of the relevant kind term. This is not too implausible, I think. When zoologists decided to incorporate the taxon monotremes into the natural kind mammal, the meaning of the latter arguably changed, as it now includes a feature which was ruled non-mammalian before that – namely, the feature of egg-laying. The same is true for zoologists' decision to exclude whales from the kind fish (Kuhn 1990, LaPorte 2009). In the case of science, there have been radical inventions, such as use of statistics in experimental tests in the late nineteenth century, which, one might argue, have altered what we mean when we use the term 'science'.

It is time for us to take stock now. A central attraction of the family resemblance concept is that it allows for 'fuzzy' boundaries: there are no criteria that would categorically rule some activity in or out of a term's extension. This seems appropriate for the term 'science', as science seems too diverse for it to be plausible that there could be any necessary and sufficient criteria for scientificity that would capture *all* activities which we regard as scientific. However, as we just saw, the family resemblance concept is beset with problems. Not only is there the well-known problem of wide-open texture, but it is also questionable whether similarities can really *justify* kind membership without some prior relevance determination through the kind in question. But without such justifications, it is hard to see how a solution to the demarcation problem based on the concept of family resemblance could have any normative force.

The PBP solution, as we have seen, offers something of a compromise between family resemblance and essentialism: although, just like on the family resemblance solution, there is not one feature that all sciences must share, there *are* necessary and sufficient conditions for the PBP, which delineates the kind. More specifically, this means that it is necessary that a science possess at least one feature that characterizes the PBP. And possession of *at least* one of the features of the PBP is sufficient for an activity to be a science.

Whether or not we'll ever manage to identify the right demarcation criteria is an open question. Indeed, I have said nothing here about how this might be done practically. My goal was more basic: to argue that the sciences must have something in common for us to *reliably* and *correctly* apply the term 'science' to physics, chemistry, and biology, and reliably and correctly deny application of the term to homeopathy, astrology, and creationism. What we can say at this point, though, I think, is that some criteria are probably not going to be the right demarcation criteria. I want to use the remainder of this paper to argue that testability very likely is not a demarcation criterion, even when combined with other criteria.

6 Falsifiability: a red herring

Popper (1959a, 1963/1978) famously advanced the thesis that scientific theories are falsifiable whereas pseudo-scientific theories are not. More specifically, it must be possible to prove theories wrong on the basis of empirical evidence for them to deserve the label 'scientific'. In contrast, Popper attributed to pseudo-scientific theories the capacity to accommodate almost any state of affairs. Popper's favorite example to illustrate this point was the contrast between psychoanalysis and Einstein's general theory of relativity. In Popper's presentation, A. Adler's notion of the 'inferiority complex' (long since part of popular language) could accommodate the contrary world-scenarios of a man scarifying his life in his attempt to save a child from drowning and of a man pushing a child into the water with the intent of killing it by appealing to the essentially same explanation. Both scenarios the psychoanalyst could explain by the man seeking to overcome his inferiority complex: in the first case, the psychoanalyst might say that the man wanted to show that he was brave enough to risk his own life, and in the latter case, that the man wanted to prove that he was brave enough to kill (Popper 1963/1978, 35).[5] In contrast to such apparent flexibility, Einstein's general theory of relativity made a very precise prediction about the bending of star light by the gravitational field of our sun. When Einstein derived this prediction in 1915, there was a real possibility that this prediction would not pan out. Four years later, A. Eddington and his colleagues confirmed Einstein's prediction with their observations at solar eclipses in Northern Brazil and on an island off the coast of Central Africa.[6]

Falsifiability remains very popular as a demarcation criterion amongst scientists and philosophers alike. As for example Godfrey-Smith (2009) writes in his excellent and widely-read introduction to the philosophy of science:

> What is Popper's single most important and enduring contribution to philosophy of science? I'd say it is his use of the idea of 'riskiness' [of being proven wrong] to describe the kind of contact that scientific theories have with observation. Popper was right to concentrate on the ideas of exposure and risk in his description of science. Science tries to formulate and handle ideas in such a way they are exposed to falsification and modification via observation. […] Popper's analysis of how this exposure works does not work too well, but the basic idea is good. (70-71)

[5] Popper's other example was Freud's Oedipus complex (Popper 1963/1978, 35).
[6] This confirmation was not as clear-cut as it is usually presented. See Earman and Glymour (1980) and Schindler (2013).

Of course, Popper's falsificationism, which outright rejects any use of inductive inference, has few if any proponents in the philosophy of science. Likewise, few scientists who endorse falsifiability would commit themselves to the idea that theories cannot be confirmed (as implied by the denial of induction). Still, many scientists and philosophers would probably agree that a theory cannot receive any confirmation if it cannot in principle be shown to be wrong on the basis of empirical evidence. Whether we call this characteristic falsifiability or testability is secondary, so long as it is understood that an endorsement of falsifiability does not necessarily entail a commitment to a falsificationism that rejects inductive inference.

First off, as is well known amongst philosophers of science, the idea that we can demarcate science from pseudoscience via the falsifiability criterion suffers from a crucial defect. Since any theory requires auxiliary assumptions in order to be empirically testable, we can never determine on the basis of logic alone whether a failed test points to the falsity of the theory or the auxiliary assumptions. A test in the laboratory, for example, will require assumptions about the functioning of the apparatus, the purity of the sample, the laboratory conditions such as lighting and temperature, etc. This is of course the well-known Duhem-Quine thesis. Popper himself was aware of it:

> no conclusive disproof of a theory can ever be produced; for it is always possible to say that the experimental results are not reliable, or that the discrepancies which are asserted to exist between the experimental results and the theory are only apparent and that they will disappear with the advance of our understanding. (Popper 1963/1978, 50, cf. p. 56)

Since Popper knew that a "purely logical analysis of science" (as he put it himself), was not sufficient for an accurate characterisation and understanding of science, he argued that such analysis must be complemented by a study of *methods*, i.e., a study of "our manner of dealing with scientific systems: by what we do with them and what we do to them" (ibid.). In particular, Popper believed, a good method of science must forbid auxiliary hypotheses which are introduced in an *ad hoc* fashion to save a theory from refutation, whereby he believed that ad hoc hypotheses decrease the degree of falsifiability of the overall theoretical system (ibid., 42 and 83). Indeed, he went as far as stipulating that "if our [theoretical] system is threatened we will *never* save it by any kind of conventionalist stratagem [i.e., ad hoc moves]" (61; section 20; added emphasis). It is therefore understandable why Lakatos (1978) once characterized Popper's falsificationism as "introducing new, non-justifiationist criteria for appraising scientific theories *based on anti-ad hocness*" (ibid., 39; added emphasis).

In what follows I will review a widely appreciated fact amongst scholars working on the demarcation problem (Pigliucci and Boudry 2013, Hansson 2017), namely that falsifiability is neither necessary nor sufficient for demarcation. Contrary to those scholars, however, I do not think that this means merely that falsifiability cannot be the *only* demarcation criterion; I believe it is simply a red herring. Furthermore, I will argue that Popper's condemnation of ad hoc hypotheses is unjustified. Although they may be undesirable for theoretical reasons, ad hoc hypotheses are an indispensable part of the scientific method. At the very least, ad hoc hypotheses cannot be understood in terms of the lack of falsifiability, as Popper believed. Again, falsifiability should not figure in the correct set of demarcation criteria.

6.1 Falsifiable pseudoscience, unfalsifiable science

Popper's main examples of unfalsifiable theories come from psychoanalysis. However, A. Grünbaum argued in several publications that Popper's examples were entirely "contrived" and that the actual theories used by psychoanalysts (at the time) were testable indeed (e.g. Grünbaum 1979). Furthermore psychoanalysts *regularly* gave up hypotheses when they didn't conform with the evidence. If this is right and if psychoanalysis really *were* a pseudo-science (which it may), the falsifiability would not be a criterion by which we could wield it out – contrary to Popper had hoped for. But then, falsifiability wouldn't be sufficient for identifying something as a science.

Another, less controversially pseudo-scientific pursuit is astrology. Yet, it has been shown even *empirically* that it is falsifiable. In an article that appeared in *Nature*, Carlson (1985) showed that the 'fundamental thesis of natal astrology', namely the idea that the positions of heavenly objects at one's date of birth determine one's character and destiny, could indeed be tested (and moreover, and not surprisingly, turned out to be false). More specifically, Carlson conducted a double-blind test with two parts. In the first part, subjects were given a choice of three astrological personality interpretations of their 'natal charts' provided by astrologers (i.e., certain constellations of astrological objects at the time of the subject's birth) and were asked to rank and score each of these interpretations. In the second part, a group of astrologers (nominated by an influential astrological society) were asked to choose and score two out of three descriptions of personalities (based on a personality test), which they thought were likely to be associated with a randomly drawn natal chart; only one of those descriptions was of course determined on the basis of each subject's natal chart (via an astrological computer programme under the supervision of high-rank astrologers). Although the first part of the experiment was somewhat inconclusive, the result of the second part was very significant and showed that astrologers were not better than chance in picking the right personality descriptions (1 out of 3), whereas astrologers had predicted they would get at least half of the choices right. One may perhaps question here the *specific* prediction of astrologers (it seems to be based on a mere guess rather than follow from any theory), but it is clearly the case that astrologers *failed* to deliver on what they claimed astrology is capable of, namely predicting somebody's personality from their natal charts.

Creationism, another of the standard examples of a pseudo-science, has also been argued to be falsifiable. For example, creationism predicts that animal and human fossils ought to be found in not too distinct geological strata, as God created humans and animals on the same day (Laudan 1982). Overall, then, falsifiability is not sufficient for telling science apart from pseudo-science. In fact, there are reasons to doubt that it is even necessary for scientificity.

Popper himself held for a long time that the theory of evolution—clearly a scientific—is not testable, because he thought that it was essentially tautological ("the survival of the fittest" for him just meant "the survival of those who survive").[7] Instead, Popper conceived of the theory of evolution as a "metaphysical research programme" that could generate testable models (Popper 1974). Later Popper changed his mind. He pointed out that not *all* traits of organisms are explained by natural selection; some are recognized to be the outcome of *genetic drift*, i.e., random genetic variation, or sexual selection (as e.g. the tail of the peacock). The theory of evolution is thus not a universal theory of biological traits. In fact, Popper went as far as saying that it is "not only testable, but it turns out to be not strictly universally true" (Popper 1978, 346). His reasoning seems to have been that because the theory of evolution is false for traits that arise from genetic drift, it is testable. And that seems to be Popper's *only* reason for deeming evolution testable. In other words, the testability of evolution is bought at the price of it being false. Clearly, though, this is too high a price to pay—the theory of evolution is as an empirically successful a theory as any.[8]

Ruse (1977) has argued against Popper that evolution is falsifiable after all. According to Ruse, in a response to a thought experiment proposed by Popper (1974), claims that a possible falsification of evolution would be the hypothetical discovery of life on another planet which was "tailor-made for speciation". That is, we suppose that on that planet there are geographically isolated areas with different climatic and nutritious conditions, that there are possible, but rare ways for

[7] But see Gould (1976)} who argued against the idea that evolution is tautological that fit variants within a population are not just variants that survive; they are variants that survive because they inherit (genetically caused) features from their predecessors that allowed them to survive. See also Ruse (1977) for a critical discussion of Popper's claim.

[8] A reasonable stance would be that evolution is a theory with a certain scope. Traits that are caused by genetic drift only are not within its scope. That does not mean, however, that the theory of evolution is false with regards to the phenomena outside its scope; those phenomena are simply irrelevant for its assessment. Furthermore, it may well be the case that inside the scope of phenomena it is supposed to apply to, the theory of evolution is true. It is not false in instances that lie outside its scope.

organisms to migrate from one to another isolated area, that there is genetic variation in the populations present, and finally, that the life on that planet has existed for an evolutionarily relevant amount of time. Now, contrary to our planet, where there is species variety through natural selection, on that planet there has never been *any speciation whatsoever*. This would pose a problem to the theory of evolution. Ruse claims that this makes the theory of evolution falsifiable at least in principle. One may object, though, that the theory of evolution is a theory about life on earth; it says nothing about how (some other form of) life might evolve in other corners of the universe. So the thought experiment Ruse considers arguably is not even be relevant.

Another point Ruse makes is that evolution is *systematic* in that it predicts that traits that are adaptive in one environment, will be adaptive in similar environments. For example, the theory of evolution predicts the existence of analogous traits between different species with different evolutionary ancestors, such as the wings in bats and birds: they both answer similar kinds of selection pressures. But of course, Popper may have had in mind something more specific: the theory of evolution does not predict *any particular* analogy or homology. Whether or not any two species exhibit those is fairly contingent. There is thus none of the risk-taking that Popper thought characterized scientific theories like Einstein's general theory of relativity (Popper 1974). In sum, evolutionary theory, despite being clearly a scientific theory, comes out as highly problematic, if not altogether unscientific on Popper's demarcation criterion.

There is at least one other theory embraced by a sizable scientific community which is widely considered to be untestable (at least in its current form) by friends and foes alike, namely string theory. The origins of string theory go back to the early 1970s. Originally, it was devised as a theory of the strong interactions, but soon developed into a theory of "everything", i.e., a theory of all four fundamental forces of nature (gravity plus the strong, weak, and electromagnetic interactions). At the core of the theory is the assumption that the fundamental physical entities are not particles but strings and membranes (higher dimensional extensions of one-dimensional strings). There are many twists and turns that this development took throughout the ensuing decades with the current endpoints of superstring theory (string theory plus supersymmetry, which posits symmetries between bosons and fermions)[9] and M-theory (the proposed limit of currently five different superstring theories). There is still a plethora of theoretical challenges: such as the many extra dimensions the theory postulates (currently six additional to the four dimensions of spacetime), the incredibly high number of vacuum states (10^{500}), and the lack of a single, fully coherent theory. And yet, string theory is still the most promising theory for achieving the unification of all four fundamental forces and the quantification of gravity.[10]

The fact that string theory has still not produced any testable predictions beyond the predictions made by the theories it unifies (and the predictions of supersymmetry) has been widely criticized and regarded as reason to no longer pursue string theory (Cartwright and Frigg 2007, Smolin 2007, Woit 2011, Ellis and Silk 2014). Others have suggested that string theory research has changed the standards of theory-evaluation in science and that the lack of better alternatives speaks to string theory's likely truth for (Dawid 2013, Dawid et al. 2014). Still others have pointed out that testability is actually viewed as an important desideratum even within the string theory community and is strived for (Johansson and Matsubara 2011, Camilleri and Ritson 2015). Nevertheless, the lack of testability does not render string theory meaningless or unscientific as a research programme (ibid.).

---

[9] For each member of those two groups, there would have to be a "superpartner". E.g., for the electron (a fermion), there would have to be a "selectron" and for the Higgs (a boson), there would have to be a Higgsino (following the official naming conventions). Supersymmetry is "modular" to string theory, that is, supersymmetry could be true without string theory being true. Yet, it would certainly be positive for string theory if supersymmetry could be confirmed (as current versions of string theory presuppose it).

[10] Dawid (2013) argues that the fact that string theory was originally developed as a theory of the strong forces only and that the unification of gravity with the other forces was *unexpected* should count in favour of string theory.

On the contrary, given that string theory is being pursued by a very significant part of the community of theoretical high energy physicists[11], any good naturalist will have to accommodate it within a theory of science. This of course does not mean that string theory will turn out true or that physicists *are* on the right track. Only that their pursuit should not be deemed unscientific.

Overall, we can thus conclude from these examples that testability seems to be highly problematic as a demarcation criterion: it fails to identify clear pseudo-sciences as such (e.g., astrology) and it identifies theories and practices as pseudo-scientific when they are not (e.g. string theory and possibly evolution). Although this is a point appreciated by many philosophers working on the demarcation problem ever since Laudan's (1983) influential paper, falsifiability is still regularly listed as one amongst other important criteria of scientificity (Pigliucci and Boudry 2013, Hansson 2017). Yet, I believe there is little reason for upholding testability as a demarcation criterion at all. I simply think it's a red herring. The problem does not lie with its categoricalness either. It simply wouldn't make much sense, I think, to say that quantum mechanics is more scientific than the theory of evolution, because quantum mechanics is arguably more testable than the theory of evolution. On the contrary, the theory of evolution just seems *as* scientific as quantum mechanics, no matter its arguably limited testability.

6.2 Ad hocness

Recall that Popper distinguished between pseudo-scientific theories that are not testable and those that are testable, but whose proponents avoid falsification by introducing ad hoc hypotheses.[12] For example, in response to the famous ether drift null result by Michelson and Morley in 1887, ether theorists such as A. H. Lorentz postulated a contraction of the measuring apparatus in the direction of motion of the

---

[11] Smolin (2007) and Woit (2011) have complained that string theory is so dominant that alternative approaches are not getting enough attention.

[12] This idea enjoys popularity amongst philosophers currently working on the demarcation problem. See e.g. Boudry (2013).

earth through the ether by an amount that would balance out any measurable either drift. The Lorentz-FitzGerald contraction hypothesis, as it came to be known, thus basically neutralized the negative evidence against the ether theory. According to Popper (who uses this example in his *Logic of Scientific Discoveries* as an unfalsifiable hypothesis (Popper 1959a, 62), ad hoc hypotheses generally decrease the falsifiability of the theories they amend. More specifically, "degrees of ad hocness are related (inversely) to degrees of testability and significance" (Popper 1959b). Popper also believed that ad hoc hypotheses "cannot be tested independently" (Popper 1976, 986), i.e., they cannot be tested by experiments that are independent of those for which the ad hoc hypotheses were introduced.

But there are doubts about the Popperian picture of ad hoc hypotheses. First of all, one may question the relationship between testability and ad hocness identified by Popper. As Barnes (2008, 11), following Bamford (1993, 349-50), has pointed out, it would for example not seem plausible to think that the hypothesis "bread nourishes" is more testable than the hypothesis "all bread nourishes except that grown in a particular region of France" that was devised upon the discovery of the relevant fact. In fact, those two sentences seem equally testable with the only difference being that bread does or does not nourish in a particular region in France. Indeed, the hypothesis is even independently testable: one can go and find out (empirically) about the particular reasons for why the bread in that particular region in France does not nourish, for example by investigating the wheat used in the process, etc.

With regards to the Lorentz-FitzGerald contraction hypothesis, it has been argued that, contra Popper, this hypothesis was indeed independently testable (Grünbaum 1959). More specifically, the contraction hypothesis entailed a measurable (but significantly lower) ether drift result. There even was an experiment, namely the so-called Kennedy-Thorndike experiment, which falsified this prediction of the contraction hypothesis.[13] Popper conceded in print to Grünbaum that the contraction hypothesis was indeed independently testable, even though he did not give up on the idea that ad hoc hypotheses are somehow deficient with regards to independent testability.[14]

There are many other accounts of ad hocness and independent testability is no longer seen as a sign of ad hocness by some philosophers of science (Schindler 2018).[15] We should also note that it would do great harm to science if scientists would follow Popper's demand *never* to save their theories from refutation by ad hoc hypotheses. Ad hoc hypotheses have in fact demonstrably contributed to the progress of science. One of the most well-known examples is the discovery of Neptune in 1846. The existence of neighboring Neptune came to be suspected after it turned out that the orbit of Uranus did not behave as predicted by Newtonian mechanics. Rather than refuting Newtonian mechanics on that basis, astronomers 'saved' by the Neptune hypothesis and were ultimately rewarded for it.[16]

There are many other examples were ad hoc hypotheses have contributed to scientific progress. One of the more entertaining ones is Einstein's discovery of the cosmological constant. Originally, Einstein in 1917 introduced it into his equations of general relativity in an ad hoc fashion, just in order to be able to retain the idea of a static (non-expanding) universe. With Hubble's discovery of the expansion of the universe, however, Einstein abandoned the constant in 1929 and allegedly even called it his "biggest blunder" (Gamow 1970). Even though most cosmological

---

[13] Whereas the Michelson-Morley experiment showed that the speed of light is independent of the orientation of the experimental apparatus, the Kennedy-Thorndike experiment showed that it was independent also of the velocity of the apparatus. For more details see e.g. (Janssen 2002).
[14] Popper in fact went as far as saying that even Einstein's special theory of relativity should be considered ad hoc to some degree and that the contraction hypothesis was only less independently testable (and therefore more ad hoc) (Popper 1966).
[15] The most popular view of ad hocness these days probably is that ad hoc hypotheses have no independent support. That is, it is conceded that ad hoc hypotheses might have independent testability, but there is no empirical support for those independent predictions.
[16] Popper himself discusses this example, but classifies the Neptune hypothesis as a *testable* auxiliary conjecture, and sticks to the idea that ad hoc hypotheses are not independently testable (Popper in Schilpp 1974, 986-7).

models up until the early 1990s made do without the cosmological constant, with the discovery in the late 1990s that the universe expansion actually accelerates, cosmologists reintroduced the constant. It thus today describes the energy density in the vacuum of space, also known as dark energy.

Even though it is undisputable that ad hoc maneuvers have contributed to the progress of science, they do remain theoretically undesirable devices for many scientists. To say that a hypothesis was introduced in an ad hoc fashion still carries mostly negative connotations; nobody would pride themselves in his or her own theory being dependent on many ad hoc assumptions! But we may well appreciate both of these things: that ad hoc hypotheses are theoretically undesirable *and* that they have contributed to scientific progress. This is so because of the many goals scientists pursue in their research, empirical adequacy and coherent explanation are two important ones. Ideally, scientists would like to have theories that both coherently explain the world and that are empirically adequate. But often, things are not as neat as that. Often, scientists do have to forego explanatory coherence in order to achieve empirical adequacy. At the same time, scientists usually do not rest content once they have achieved empirical adequacy via ad hoc hypotheses, but rather seek to re-establish explanatory coherence. This is illustrated nicely in the following case.

In order to derive the frequency distribution and temperature dependence of black body radiation which he had been working on, Max Planck in 1900 introduced to physics the quantization of energy: he assumed that light could be emitted and absorbed only in amounts of energy equal to the product of the frequencies of light and integral multiples of what is now known as Planck constant. Planck's hypothesis is a paradigmatic ad hoc hypothesis: there were no reasons for introducing it *other than* to derive his black body radiation law. We all know, though, what revolution Planck's discovery sparked!

I think we can conclude from all this that the use of ad hoc hypotheses, though theoretically undesirable, can contribute to the progress of science. Although the defining characteristics of ad hoc hypotheses remain disputed, it is widely accepted (eventually even by Popper himself), that ad hoc hypotheses are testable. Most importantly for our concerns, the use of ad hoc hypotheses cannot be a criterion for distinguishing pseudo-science from science.

7 Conclusion

In this paper, I argued that the popular idea of science as a family resemblance is inherently problematic; it cannot be used to address the demarcation problem. Instead I proposed that the meaning of 'science' is best understood in terms of a paradigm or a basic predicate, whose properties must be fully specifiable for our talk about science vs. pseudoscience to be coherent. The still widely popular falsifiability criterion, I argued, should not figure as one of those properties. Demarcation, whatever form it will eventually take, will have to proceed without it.

References

Bamford, G. 1993. Popper's Explications of Ad Hocness: Circularity, Empirical Content, and Scientific Practice. *The British Journal for the Philosophy of Science*, **44** (2): 335-355.

Barnes, E.C. 2008. *The paradox of predictivism*. Cambridge: Cambridge University Press.

Bellaimey, J.E. 1990. Family Resemblances and the Problem of the Under-Determination of Extension. *Philosophical Investigations*, **13** (1): 31-43.

Boudry, M. 2013. Loki's Wager and Laudan's Error. In *Philosophy of Pseudoscience: Reconsidering the Demarcation Problem*, Massimo Pigliucci and Maarten Boudry (eds.), Chicago: University of Chicago Press, 79-101.

Camilleri, K. and S. Ritson. 2015. The role of heuristic appraisal in conflicting assessments of string theory. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, **51**: 44-56.

Carlson, S. 1985. A double-blind test of astrology. *Nature*, **318** (6045): 419-425.

Cartwright, N. and R. Frigg. 2007. String theory under scrutiny. *Physics World*, **20** (9): 14.

Dawid, R. 2013. *String theory and the scientific method*. Cambridge: Cambridge University Press.

Dawid, R., S. Hartmann, and J. Sprenger. 2014. The no alternatives argument. *The British Journal for the Philosophy of Science*: axt045.

Dupré, J. 1993. *The disorder of things: Metaphysical foundations of the disunity of science*. Harvard: Harvard University Press.

Earman, J. and C. Glymour. 1980. Relativity and eclipses: The British eclipse expeditions of 1919 and their predecessors. *Historical Studies in the Physical Sciences*, **11** (1): 49-85.

Ellis, G. and J. Silk. 2014. Scientific method: Defend the integrity of physics. *Nature*, **516** (18 December 2014): 321-323.

Gamow, G. 1970. *My world line*. New York: Viking.

Godfrey-Smith, P. 2009. *Theory and reality: An introduction to the philosophy of science*: University of Chicago Press.

Goodman, N. 1972. Seven Strictures on Similarity. In *Problems and Projects*, Indianapolis: Bobs-Merril.

Gould, S.J. 1976. This view of life: Darwin's untimely burial. *Natural History*, **85** (8): 24p.

Grünbaum, A. 1959. The falsifiability of the Lorentz-Fitzgerald contraction hypothesis. *The British journal for the philosophy of science*, **10** (37): 48-50.

———. 1979. Is Freudian psychoanalytic theory pseudo-scientific by Karl Popper's criterion of demarcation? *American Philosophical Quarterly*, **16** (2): 131-141.

Hansson, S.O. 2017. Science and pseudo-science. *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, https://plato.stanford.edu/archives/sum2017/entries/pseudo-science/.

Hoyningen-Huene, P. 2013. *Systematicity: The nature of science*. Oxford: Oxford University Press.

Irzik, G. and R. Nola. 2011. A family resemblance approach to the nature of science for science education. *Science & Education*, **20** (7-8): 591-607.

Janssen, M. 2002. Reconsidering a scientific revolution: The case of Einstein versus Lorentz. *Physics in Perspective*, **4** (4): 421-446.

Johansson, L.G. and K. Matsubara. 2011. String theory and general methodology: A mutual evaluation. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics*, **42** (3): 199-210.

Kuhn, T.S. 1970. Logic of discovery or psychology of research. In *Criticism and the Growth of Knowledge, Proceedings of the International Colloquium in the Philosophy of Science*, I. Lakatos and A. Musgrave (eds.), Cambridge: Cambridge University Press, 1-24.

———. 1990. Dubbing and redubbing: The vulnerability of rigid designation. *Minnesota studies in the philosophy of science*, **14**: 298-318.

Ladyman, J. 2013. Toward a Demarcation of Science from Pseudoscience. In *Philosophy of Pseudoscience: Reconsidering the Demarcation Problem*, Massimo Pigliucci and Maarten Boudry (eds.), Chicago: Chicago University Press.

LaPorte, J. 2009. *Natural kinds and conceptual change*. Cambridge: Cambridge University Press.

Laudan, L. 1982. Commentary: Science at the bar—causes for concern. *Science, Technology & Human Values*, **7** (4): 16-19.

———. 1983. The demise of the demarcation problem. In *Physics, Philosophy and Psychoanalysis: Essays in Honour of A. Grünbaum*, Robert S. Cohen and Larry Laudan (eds.), Dordrecht: Reidel, 111-127.

Pigliucci, M. 2013. The demarcation problem: a (belated) response to Laudan. In *Philosophy of Pseudoscience: Reconsidering the Demarcation Problem*, Massimo Pigliucci and Maarten Boudry (eds.), Chicago: University of Chicago Press, 9-29.

Pigliucci, M. and M. Boudry. 2013. *Philosophy of Pseudoscience: reconsidering the demarcation problem*. Chicago: University of Chicago Press.

Pompa, L. 1967. Family Resemblance. *The Philosophical Quarterly*, **17** (66): 63-69.

Popper, K.R. 1959a. *The logic of scientific discovery*. London: Routledge.

———. 1959b. Testability and 'ad-hocness' of the contraction hypothesis. *British Journal for the Philosophy of Science*, **10** (37): 50.

———. 1963/1978. *Conjectures and Refutations: The Growth of Scientific Knowledge*. fourth edition ed. London: Butler & Tanner Limited.

———. 1966. A Note on the Difference between the Lorentz-Fitzgerald Contraction and the Einstein Contraction. *The British journal for the philosophy of science*, **16** (64): 332-333.

———. 1974. Darwinism as a metaphysical research programme. In *The Philosophy of Karl Popper*, P. A. Schlipp (ed.), LaSalle, Ill.: Open Court, 167-174.

———. 1976. *Unended Quest: An Intellectual Autobiography*. London: Routledge.

———. 1978. Natural selection and the emergence of mind. *Dialectica*, **32** (3-4): 339-355.

———. 2002. *The logic of scientific discovery*: Routledge.

Richman, R.J. 1962. "Something Common". *The Journal of Philosophy*, **59** (26): 821-830.

Ruse, M. 1977. Karl Popper's Philosophy of Biology. *Philosophy of Science*, **44** (4): 638-661.

Schilpp, P.A. 1974. *The Philosophy of Karl Popper*: Open Court.

Schindler, S. 2013. Theory-laden experimentation. *Studies in History and Philosophy of Science Part A*, **44** (1): 89-101.

———. 2018. A coherentist conception of ad hoc hypotheses. *Studies in History and Philosophy of Science*, **67** (February): 54–6.

Simon, M.A. 1969. When is a resemblance a family resemblance? *Mind*, **78** (311): 408-416.

Smolin, L. 2007. *The Trouble With Physics: The Rise of String Theory, The Fall of a Science, and What Comes Next*. Boston (M.A.): Houghton Mifflin Harcourt.

Wittgenstein, L. 1953. *Philosophical investigations*. Translated by G.E.M. Anscombe. Oxford: Blackwell.

Woit, P. 2011. *Not Even Wrong: The Failure of String Theory and the Continuing Challenge to Unify the Laws of Physics*. London: Random House.