

A conceptual problem for Stanford's New Induction

Um problema conceitual na Nova Indução de Stanford

Bruno Malavolta e Silva

<http://orcid.org/0000-0003-3200-3700>

Universidade Federal do Rio Grande do Sul, Programa de Pós-Graduação em Filosofia, Porto Alegre/RS. Email: malavolta3@gmail.com.

ABSTRACT

The problem of unconceived alternatives (or the New Induction) states that, since scientists have recurrently failed to conceive relevant theoretical alternatives for some domains of science, current scientists are probably also failing to do so. Therefore, there may be theories which still exceed the grasp of scientists' imagination, and one should not endorse a realist stance towards current science. In this paper, I raise a conceptual worry for the formulation of this problem: what does it mean to say that scientists failed to conceive a *relevant* theory? What aggravates the problem is that no simple notion of relevance makes the New Induction as strong as it initially seems. I consider the three more obvious interpretations of relevance: relevance as objective probability; relevance as epistemic probability assessed by current scientists; and relevance as epistemic probability assessed by past scientists. I argue that assuming any of these three notions implies difficulties for the New Induction, hence their proponents shouldn't take the notion of relevance for granted. A more precise definition of relevance is essential to understand what are the difficulties surrounding the problem of unconceived alternatives as an epistemic worry. Until now, such notion is missing.

KeyWords: Scientific realism, unconceived alternatives, Kyle Stanford, New Induction, Pessimistic Induction.

RESUMO

O problema das alternativas não concebidas (ou a Nova Indução Pessimista) afirma que, dado que cientistas do passado falharam recorrentemente em conceber alternativas teóricas relevantes para certos



domínios da ciência, então cientistas atuais provavelmente falham do mesmo modo. Portanto, podem haver teorias que ainda estão além da imaginação dos cientistas, e não devemos endossar uma atitude realista perante a ciência atual. Neste artigo, desenvolvo um enigma conceitual para a formulação desse problema: o que significa dizer que os cientistas falharam em conceber uma teoria *relevante*? O enigma é agravado, porquanto nenhuma noção simples de relevância é capaz de manter a Nova Indução tão plausível quanto ela parece inicialmente. Considero três interpretações mais óbvias de relevância: relevância como probabilidade objetiva; relevância como probabilidade epistêmica avaliada pelos cientistas atuais; e relevância como probabilidade epistêmica avaliada pelos cientistas do passado. Argumento que assumir qualquer dessas três interpretações resulta em dificuldades para a Nova Indução, e por isso seus proponentes não devem tomar a noção de relevância como dada. Uma definição mais precisa de relevância é essencial para compreender quais dificuldades circundam o problema das alternativas não concebidas enquanto um problema epistêmico. Até então, tal definição não existe.

Palavras-chave: Realismo científico, alternativas não concebidas, Kyle Stanford, Nova Indução, Meta-Indução Pessimista.

Introduction: The problem of relevant unconceived alternatives

Scientific realists sustain that theories from mature sciences are approximately true about an independent reality, including some unobservable aspects of it. In the current state of discussion, one of the main challenges for scientific realists is Kyle Stanford's Problem from Unconceived Alternatives, also called New Induction from the history of science. In *Exceeding Our Grasp*, Stanford (2006) argues that we have good reasons to believe that, at the most fundamental domains of scientific inquiry, the correct descriptions of reality are beyond the imagination of current scientists. Therefore, we shouldn't believe in our current theories at those fundamental domains. Instead, we should expect those theories to be eventually replaced by theoretical alternatives currently unconceived by us. Stanford's argument can be formulated as following:

- (1) At domain *D*, scientists have recurrently failed to conceive some relevant theoretical alternative when accepting a theory.
- (2) ∴ Therefore, scientists probably don't have a good capacity to conceive all the relevant theoretical alternatives for *D*.
- (3) Reliably performing an eliminative inference requires a reliable capacity to conceive all the relevant theoretical alternatives for the inquired domain.
- (4) ∴ Scientists probably don't have the capacity to perform reliable eliminative inferences about *D*.

Let's start by looking at claim (3). Stanford's problem challenges the reliability of eliminative inferences at certain specific contexts of inquiry. An eliminative inference is an inferential process which starts by gathering the set of minimally plausible hypothesis for a given domain, proceed by testing and eliminating them until there is only one left, and then infer the truth of the remaining survivor (Lipton, 2004). Stanford's problem is inspired by a worry about eliminative inferences raised by Pierre Duhem: if we fail to conceive a relevant hypothesis, then we will not assess it nor seek to eliminate it; hence our eliminative inference will be unreliable, because we may have failed to conceive precisely the true hypothesis, which would lead us into inferring "the best of a bad lot" of false hypotheses (the expression is from Van Fraassen, 1980).

Of course, given that eliminative inferences are very entrenched in our cognitive life, both scientifically and quotidian (Douven, 2011), we can assume that this worry is not a real threat for us in ordinary contexts. Normally, when we are confident about our ability to perform an eliminative inference, we

have a sufficiently good ability to conceive all the relevant theoretical alternatives, and thus our quotidian eliminative inferences work in a good amount of the time. Stanford concedes this point (2006, Chapter 2). But the threat from unconceived alternatives stops being a purely conceptual worry and becomes a solid epistemic threat when we have positive evidence to believe that we are bad at conceiving theories for some specific context. For example, if I don't have any background knowledge about how computers work, then it would be futile for me to try to do an eliminative inference about why my computer stopped working; I don't have what it takes even to conceive the relevant possible causes.

So, in order for the threat of unconceived alternatives to become a real problem, we must have reasons to doubt scientists' imaginative abilities. Here comes the New Induction from the history of science, expressed in (1) to defend (2). Stanford argues that, if we are to reflectively judge the capability of scientists to conceive theories about some matter, then we should seek evidence in history of science, and check the scientists' previous record of conceiving theories for that domain. And when we check this, we'll find that, for the most remote and theoretical domains of scientific inquiry (such as the nature of light and her maximum speed, the origins of life, or what happened at the first moments after the Big Bang), scientists exhibit a pattern of repeated failure: historically, scientists recurrently failed to conceive some relevant theoretical alternative for the domain – precisely the one that posteriorly replaced the initially endorsed theory. To empirically show this historical pattern of failures, Stanford provides a list of scientific domains exhibiting it: (Stanford's list mimics Larry Laudan's notorious list of empirically successful theories without referents, see Laudan 1981):

from elemental to early corpuscularian chemistry to Stahl's phlogiston theory to Lavoisier's oxygen chemistry to Daltonian atomic and contemporary chemistry;
from various versions of preformationism to epigenetic theories of embryology
from the caloric theory of heat to later and ultimately contemporary thermodynamic theories
from effluvial theories of electricity and magnetism to theories of the electromagnetic ether and contemporary electromagnetism
from humoral imbalance to miasmatic to contagion and ultimately germ theories of disease
from eighteenth century corpuscular theories of light to nineteenth century wave theories to the contemporary quantum mechanical conception
from Darwin's pangenesis theory of inheritance to Weismann's germ-plasm theory to Mendelian and then contemporary molecular genetics
from Cuvier's theory of functionally integrated and necessarily static biological species and from Lamarck's autogenesis to Darwin's evolutionary theory (Stanford, 2006, p. 20–21).

This list gives us an inductive ground to claim that, as matter of fact, scientists probably don't have a good capacity to conceive all the relevant alternatives at each of these domains. In this sense, the New Induction turns the problem of unconceived alternatives into a concrete epistemic worry instead of a cartesian skeptical challenge, supporting (2) with (1).

Stanford's argument provoked all sorts of reactions. Among the main concerns, some objected that the New Induction fails because current theories enjoy a much higher degree of confirmation than theories from past science, and therefore we shouldn't judge the fate of current science by its distant past (Forber, 2008; Godfrey-Smith, 2008; Devitt, 2011). Stanford replied that current scientific institutions got more conservative and that actual scientists have less freedom to develop alternatives than they had in the past, so that the problem for unconceived alternatives remains a threat (Stanford, 2015; 2019). A second attack on the New Induction claims that it is neutralized by the historical continuity between theories, since this shows that unconceived alternatives did not threaten the stability of past science, and hence will not threaten the stability of future science. If unconceived alternatives exist, they are probably conservative with current science (Enfield, 2005; Magnus, 2006; Chakravartty, 2008). But others replied

that Stanford's problem is not reducible to the traditional antirealist arguments related to the historical discontinuity of science, so that it still adds to the debate (Magnus, 2010; Egg, 2016). This certainly does not exhaust the discussion surrounding Stanford's argument, and the debate is alive and kicking.

In this paper I raise a conceptual worry which has been neglected by the discussion so far, and can potentially affect the plausibility of the New Induction. The worry is what I'll call the problem of *theoretical relevance*: what does it mean to say that scientists failed to conceive *relevant* theoretical alternatives? In order to argue that eliminative inferences are unwarranted for some domains, Stanford starts from the premise that past scientists have recurrently failed to conceive relevant theories. But then the argument requires a minimal criterion of *relevance* for making this claim sustainable. If the problem of unconceived alternatives was merely a problem of irrelevant unconceived alternatives, then it would not threaten the reliability of eliminative inferences, since they can still be reliable even without us conceiving every irrelevant hypothesis. So, assuming some notion of relevance is crucial to formulate Stanford's argument. But what exactly is theoretical relevance?

It's tempting to dismiss the question with superficial answers. One might say that relevant theories are those that must be investigated by science in order to know or understand a domain. This kind of answer is not entirely empty, but it merely postpones the essential problem behind our initial question. It invites us to ask again, what makes a theory to be one that must be investigated by science?

The problem of relevance also appears in the discussions of Fred Dretske's epistemology of relevant alternatives. Dretske proposes "to think of knowledge as an evidential state in which *all the relevant alternatives* (to what is known) *are eliminated*" (2000, p. 52). This certainly harmonizes with the proposal of eliminative inferences, where we know a conclusion by reflecting on the relevant alternatives and eliminating them. The question of what makes an alternative relevant is also vital for Dretske's proposal, since the answer determines what's required for knowledge. Dretske's initial answer is also a superficial one: a relevant alternative is an alternative "that a person must be in a[n] evidential position to exclude (when he knows that *P*)" (2000, p. 57). Just as before, it merely postpones the problem by inviting a new question: what makes an alternative to be one that a person must be in a position to exclude in order to have knowledge?

Unfortunately, the literature on Dretske's theory of relevant alternatives has no consensual answer for us to import to the discussion of Stanford's problem (Black, 2020, sec. 3.a). The literature on the matter is split between two alternatives, presented in Cohen (Cohen, 1988, p. 101–103). One option, which is favored by Dretske, is to claim that an alternative *P* is relevant if and only if there's an objective possibility of *P* (that is, if, as a matter of fact, *P* is possible). The other option is to claim that an alternative *P* is relevant if and only if there's an epistemic possibility of *P* (that is, if we accept that *P* is possible).

The literature on Stanford's problem usually takes for granted the notion of relevance. By doing so, it also has no univocal notion of relevance, and the notion is implicitly understood in an inconstant way, alternating between interpretations in terms of objective and epistemic probability. I think that this conceptual obscurity in Stanford's argument hides some weakness in its defense. Thus, the matter shouldn't be treated unreflectively. What aggravates the problem is that no simple notion of relevance makes the New Induction as strong as it initially seems. In the remaining of this paper, I defend this claim by investigating how Stanford's argument interacts with specific definitions of relevance. I consider what I take to be the three straightforward interpretations of relevance derived from discussions of Dretske's theory: relevance as objective probability; relevance as epistemic probability assessed from current science; and relevance as epistemic probability assessed from past scientists.

1. Objective probability

First, consider objective probability. By that I mean the probability (as a matter of fact) for an event to occur in a given situation, regardless of how we assess the probability to be. Objective probabilities

are determined by the natural laws of the world, or (if we want to be nominalists about natural laws) simply by the way the world is (Hájék, 2019). For example, a particular radium atom will *probably* decay within 10,000 years, as expressed by the laws of physics. With this notion, we may define a relevant alternative:

Objective Relevance: *A theoretical alternative for domain D is relevant if and only if it expresses events with objective probability above a stipulated minimum.*

If we define relevance in terms of objective probability, then we get an account of eliminative inferences closely connected to contextualist epistemologies (Black, 2020). A rule of eliminative inference prescribes that we should track the relevant alternatives and eliminate all but one of them. But according to objective relevance, what makes alternatives relevant are matters of fact. Hence, what alternatives are required to be conceived will be also a contextual matter of fact. If we want to perform an explanatory eliminative inference about the cause of an event, we should be able to know what are the objectively possible causes for it (i.e. those with objective probability above the stipulated minimum), and then go on to assess and eliminate all but one of them. E.g., if we are in a region without any wolves, then I can reliably infer that a track in my garden has been caused by a dog, given its shape. But if wolves have invaded the city and I never dreamed about the possibility, then this becomes a relevant alternative and my inference is no longer reliable, even if the track was a dog's one after all.

However, assuming objective relevance undermines Stanford's argument. The problem is that, if we assume relevance to be determined by any factual or non-epistemic criteria, then Stanford's inductive sample is undermined. Stanford's list is composed mostly by historically abandoned theories, whose ontologies were excluded from current science. The facts and entities expressed by these theories are no longer considered to be real (e.g. miasmatic contagion, phlogiston, electromagnetic aether, and so on). So, they will not be judged as objectively probable by anyone who endorses current science, nor by anyone who recognizes the reasons why such past theories were abandoned. Phlogiston and miasma have no place in the natural laws of the world.

Notice that, for any domain mentioned in Stanford's list, the list mentions a successive line of theories developed and abandoned in that domain. When the abandoned theories are radically different from current scientific views, then such theories will be no longer relevant according to objective relevance (as far as we know), and only one example of Stanford's list will be actually relevant – the last theory to be accepted on that domain, which is current science. Thus, for any domain, the only example of unconceived alternative we'll have is that current science was once unconceived.

Alternatively, one may wonder if perhaps some domains on the list include *approximately* true theories, which were partially preserved in current science, but were not conceived by the previous generations of scientists. If that's the case, then according to current science these theories are *relevant* theories according to the criteria of objective relevance. However, in this case, these theories are no longer theoretical *alternatives* in a sense capable of maintaining (3) and threatening realism. In order to reliably perform an eliminative inference, we don't need to consider every approximately true theory (in fact, theories with mathematical formulas will have infinite approximated versions, and surely, we don't need to consider every one of them). If the unconceived alternatives are approximations from the theory we currently accept, then the only risk they raise is that we may have accepted an approximately true theory instead of a true one. But scientific realists accept that current science is only *approximately* true, and thus this conclusion is no threat for realism.

Thus, on the objective definition of relevance, past scientists failed to conceive only one theoretical alternative which was radically different from their own views and which we would consider relevant, and this alternative is precisely the current scientific view for that domain. If this is all the inductive evidence for the New Induction, then it's quite weak. In fact, it becomes quite hard to say that the New Induction is based upon evidence from the history of science, since one doesn't need much history to guess that current scientific views were once unconceived. At the same time, if past scientists didn't fail to conceive

radically different alternatives to their theories, we have no historical evidence to say that current scientists may be doing so as well. The most we can conclude is that, just as past scientists failed to conceive some details relevant to develop parts of their theories, so current science is probably failing to do it as well. But this is trivial and entirely compatible with realist's claims of approximate truth.

2. Epistemic probability

Assessing theoretical relevance in terms of objective probability did no good for the New Induction. Let us try again, then, with a notion of relevance defined in terms of epistemic probabilities. In this case, *relevant* theoretical alternatives are those judged as plausible ones:

Epistemic Relevance: *A theoretical alternative for domain D is relevant if and only if it has epistemic probability above a stipulated minimum.*

More simply, we may say that relevant theories are plausible ones. This notion of relevance is enough to distinguish unrealistic proposals from explanations that scientists would consider at least minimally plausible to be worth investigating. Performing eliminative inferences becomes a matter of finding all the hypothesis having initial plausibility, and investigating then throw until only one remains as the likeliest.

Now, if relevance is determined by judgements of plausibility, then a first crucial question is where these judgements of plausibility are grounded. Who makes these judgements? We from the present, together with current scientists, decide what's relevant for a domain of investigation in all its history? Or past scientists are the ones who determine what's relevant for their own context?

A first option is to ground the assessments of relevance in the present:

Diachronic Epistemic Relevance: *A theoretical alternative for domain D is relevant if and only if, according to current scientists, it has epistemic probability above a stipulated minimum.*

On this notion, we, scientists and philosophers from the present, are the ones who determine what are the relevant theoretical alternatives for a domain. When we assess past scientific activity, our judgements of relevance will become historically diachronic, since we are judging an epistemic activity with information unavailable in the past. Thus, when Stanford says that eighteenth century physicists failed to conceive a relevant alternative because they failed to conceive the quantum mechanical conception of light (2006, p. 21), he seems to be making a judgement about what *he* or contemporary physicists would judge as plausible and relevant. This option may seem attractive, since the problem of unconceived alternatives targets *current* science, so it has to imply that the unconceived theoretical alternatives would be relevant for us in the present, if the problem is intended to be relevant for us.

However, it should be clear that grounding the assessments of relevance on current science raises exactly the same problems faced by the approach with objective probability, since most theories from the past would be treated nowadays as improbable. The evidence that made us seem contemporary genetics as relevant also made us see the germ-plasm theory as false and proven wrong. Assuming diachronic epistemic relevance, only current science will be relevant.

The main alternative, then, is to employ epistemic probability as determining relevance in a *synchronic* approach:

Synchronic Epistemic Relevance: *A theoretical alternative for domain D is relevant if and only if, according to the scientists proponents of the theory, it has epistemic probability above a stipulated minimum.*

On this account, the assessments of past scientists determine what's relevant for an agent in their own context. This alternative fares better, since now abandoned theories don't become irrelevant just because of the reasons that made them abandoned. Thus, (1) seems to be true and Stanford's list of once endorsed theories fulfills the role of providing relevant theoretical alternatives for some domains. Furthermore, when we apply the synchronic relevance to assess current science, the result will

be that relevant theories will be the ones plausible according to us. Thus, (3) also seems to receive an acceptable interpretation: reliably performing an eliminative inference requires for us to conceive and eliminate the possibilities that we by ourselves consider as plausible ones. Thus, the problem of unconceived alternatives looks well built.

The notion of synchronic epistemic probability originates a new worry. If past scientists *S* didn't conceive a theory *T*, then they haven't formed a concrete judgement about *T*'s plausibility. Thus, the value expressing the synchronic epistemic probability of *T* for *S* will be an indeterminate parameter in the actual world. In order to determine it, we need to make a counterfactual judgement about how past scientists would have assessed *T* if they had conceived it. But are we sure that, if posterior theories were presented to past scientists, these theories would be treated as plausible ones?

Generally speaking, the problem is that for a theory to be scientifically treated as a competitor it must have some level of confirmation by the available evidence. But past scientists don't have access to the posterior evidence that make current theories acceptable. They also don't have the auxiliary hypothesis and the background knowledge required to understand and articulate a contemporary theory relating it to the facts they do know. So, saying that the quantum mechanical conception of light was a relevant alternative for eighteenth century physics is no trivial matter. To say the least, this claim requires a counterfactual judgement which is not well argued by Stanford's list nor by his treatment of the problem (not even in the examples of genetics, where he dives deeper).

The problem is aggravated if we rely on the Kuhnian image of science (Kuhn, 1962). Kuhn tell us that, once a domain of investigation overcomes its pre-scientific phase and a paradigm is accepted as a scientific consensus, theoretical alternatives which oppose the paradigm are excluded from the scientific arena, and become treated as unscientific or purely philosophical hypotheses. This is because the communitarian endorsement of the paradigm makes it to be much more empirically grounded and much more articulated than any new hypothesis could be in its initial state. The paradigmatic theory will be connected to much more evidence, making the new alternative sound like an untested and implausible hypothesis rejected by the well established paradigm. Thus, even if some lone wolf scientist had proposed the core of a contemporary theory to the previous paradigm, he would probably not have been taken seriously as offering a scientifically relevant hypothesis to the scientific community. New theories which challenge the paradigm are only taken seriously after the paradigm faces a crisis, where the emergence of anomalies begins to dismantle the epistemic confidence enjoyed by it. Interestingly, this conservative aspect of science would make the conclusion of Stanford's argument even more relevant, since it invites us to wonder whether this conservatism of paradigms is good for science, and whether we should not endorse a more tolerant and pluralist organization of inquiry (see Feyerabend, 1978). But if we assume the notion of synchronic epistemic relevance, then this conservatism undermines the claims that past scientists failed to conceive relevant alternatives simply because they failed to conceive posterior science. Synchronically, the core theories of posterior science would not have been relevant to past science before it faces a crisis. And what matters for the synchronic notion of relevance is what scientists would have judged at the time they realized an eliminative inference and accepted the paradigmatic theory, hence before the crisis.

With this problem, I think that assuming synchronic epistemic relevance puts the New Induction in a considerably worse shape than it initially seemed to have. At first, it seemed that Stanford's list alone provided inductive evidence to sustain the New Induction over a bunch of scientific domains. Now, for any one of these domains, the New Induction requires a sophisticated historical analysis showing counterfactually that past scientists would have judged posterior theories as plausible ones even when they were immersed in a rival paradigm. Notice that this problem affects not only the claim that a current scientific theory was a relevant unconceived alternative for past scientists, it also affects the claim that any posterior theory was a relevant unconceived theory. Thus, in the sequence of replaced theories of a domain, this problem undermines not only the last layer of the list, the layer of current theories, but it

also undermines all the layers except the first one. We don't know whether any of these theories would be relevant for the previous generation of scientists in their domain.

It's worth considering a possible response to this problem of counterfactual judgements. Let's say that the isolated core of a posterior theory wouldn't be accepted as plausible by their predecessors. Does that mean that there's no ground to say that past scientists slipped a relevant alternative, given that they didn't conceive current theories? Perhaps no. We can redefine 'theory' and say that the unconceived alternative is not the core theory taken alone, but a broader theoretical system composed by (i) the core of the posterior theory, added by (ii) all auxiliary hypotheses required to empirically ground it and maintain its explanatory value. This theoretical system, considered in its whole, was unconceived by previous scientists and could be treated as relevant for them, since now the new theory is articulated enough to compete with previous paradigm as a scientifically relevant alternative.

Interestingly, this interpretation can reveal another aspect of the New Induction: it may even occur that the core theory was effectively conceived long ago by the predecessors but was neglected because the relevant auxiliaries were unconceived, and so the problem of unconceived alternatives takes the form of a problem of neglected alternatives originated from a problem of unconceived auxiliary hypotheses. In this case, the conclusion would not be that there may be some relevant theoretical alternative entirely unconceived to current scientists, the conclusion would be that some relevant alternative may be being neglected because we haven't thought of the relevant auxiliaries to make the alternative plausible enough for the scientific arena.

Does that fix the problem? In some cases, it can. If we project forwardly the past failures to conceive relevant theoretical systems, then we may have a version of the new induction sustaining that current scientists still are failing to conceive some relevant theoretical system incompatible with current views, and that's enough to sustain the spirit of Stanford's argument. But in another cases, this solution may not do. The danger is that the auxiliary hypotheses are too disconnected from the evidence available to past scientists, so that even if the broader theoretical system was somehow conceived in its complexity, it would still be disregarded as an ad hoc maneuvering to save the core theory, instead of a relevant competitor. Thus, in order for this solution to work, the auxiliary hypothesis required to save the core theory cannot be too departed from the data available to past scientists. Showing whether this does or does not happen for some theory still requires an empirical case analysis which greatly surpasses the mere mentions of Stanford's list.

It may be tempting to save further the New Induction with a third interpretation of "theories", where the unconceived alternatives are understood as even broader theoretical systems composed by (i) the core theory, (ii) the relevant auxiliary hypotheses, plus (iii) the posterior experimental evidence that justifies such hypotheses in contemporary science. With this move, we inject the evidence available to posterior scientists into the background of past scientists, which allows us to say that even if the theoretical system was not considered relevant to them, it should be considered, and it would be considered if they have access to the posterior evidence. However, this interpretation trivializes the New Induction, since almost any skeptical hypothesis would be plausible and should be treated as relevant if we hypothetically assume the evidence in its favor. Even cartesian Demons would be plausible if we assume a scenario where they are talking to us and furnishing sufficient empirical evidence.

Finally, it's worth mentioning that there are some proposals of extending the New Induction which seem to navigate through these various interpretations of Stanford's argument. K. Wray (2016) and D. Rowbottom (2019) both argue that the problem of unconceived alternatives can be applied not only to hypotheses and theories but also to other aspects of scientific practice:

1. Unconceived Observations: assuming that certain observations are theoretically contaminated, we may not have the theories required to make determined observations which are relevant to choose the right theories in some domain (Rowbottom, 2019).

2. Unconceived Models and Predictions: assuming that the derivation of certain predictions depends on the construction of a model that shows how to apply the theory in a given situation, the possibility of hav-

ing unconceived models entails the possibility of having unconceived predictions relevant to the choice of theories. For example, the adequacy of Newtonian mechanics to deal with real pendulums (i.e. pendulums endowed with mass, friction, and three-dimensional motion) was not initially recognized, and the prediction was only considered successful after the development of the appropriate models (Rowbottom, 2019, p. 10).

3. Unconceived Experiments and Instruments: similarly, the appreciation of the predictive power of a theory may depend on experiments or technological instruments initially unknown. Here examples abound: any experiment unknown to the development of the theory and confirming a rival alternative would count as a case, just as any technological innovation relevant to later confirmation of the theory, such as microscopes or GPS (Wray, 2016; Rowbottom, 2019).

4. Unconceived Methodologies: if theoretical virtues (such as simplicity, scope, precision, fertility, and consistency, to use Kuhn's list) are treated as relevant for the confirmation of theories, and if the acceptance and application of these virtues is modified throughout scientific development, then we may think that Unconceived Virtues (or new interpretations of already accepted virtues) could affect the choice of theories so as to make neglected alternatives become relevant hypotheses (Rowbottom, 2019). The same can be said regarding the development of new methodological techniques (Wray, 2016).

These points give us a fruitful material to develop more specific versions of the problem of unconceived theories. But overall, they distinguish from the problem of unconceived theories only if we understand "theories" in a narrow sense referring to the core of a scientific theory. If we adopt the second mentioned interpretation, were unconceived "theories" refer to wide theoretical systems, it can include unconceived models, predictions, methodological technics, among others. And if we adopt the third interpretation, then we can include even unconceived experiments and observations into the problem, but dealing with the risk of trivializing it. So, while separating these multiple dimensions of the New Induction is elucidative and potentially relevant to other purposes, these multiple dimensions can still be reunited within a general version of the problem of unconceived theories-broadly-understood. More importantly, separating these multiple dimensions still does not remove the general problem from the uncomfortable situation: if we understand relevance in terms of synchronic epistemic relevance, then we'll need a deeper historical analysis capable of defending counterfactual judgements about how past scientists would assess posterior theories, models, methodological virtues, or whatever else. This is a problem created by the notion of synchronic epistemic relevance, and apply to any of these levels.

Conclusion

The problem of unconceived alternatives is one of the main arguments surrounding the scientific realism debate and its discussion is very much alive. I raised a conceptual worry in order to bring attention to a neglected obscurity in Stanford's argument. The worry is the problem of *theoretical relevance*: what does it mean to say that scientists failed to conceive *relevant* theoretical alternatives? The new induction requires some notion of relevance to be sustainable. What aggravates the problem is that no simple notion of relevance makes the New Induction as strong as it initially seems. We may treat relevance as objective probability, but then Stanford's sample is undermined, since we would judge only current scientific theories as relevant, and Stanford's list of abandoned theories becomes useless. We may treat relevance as epistemic probability assessed diachronically by current scientists (or philosophers), but then the same problem seems to apply. And we may treat relevance as epistemic probability assessed synchronically by past scientists, but then the New Induction relies on counterfactual judgements about how past scientists would assess posterior theories. These judgements would have to show what auxiliary hypotheses are necessary to make the posterior theories plausible to past scientists, and also whether these auxiliaries would be plausible by evidence accessible to past scientists. This is a complex matter for case analyses, which makes Stanford sample way more controversial than it initially was.

One might get the impression that discussing the notion of relevance is a red herring; that it's a proposal to avoid dealing with a real epistemic worry by appealing to an artificial conceptual problem. Perhaps further discussion reveals that this is right. But the epistemic worry raised by the problem of unconceived alternatives has already been widely discussed, and the notion of relevance has been almost entirely neglected, even though its interpretation can deeply affect the argument and the extent of its inductive support. While investigating the notion of relevance may initially seem unnecessary, I've argued that any obvious sense of relevance that we may attribute to the notion gets the New Induction into some particular difficulties. Hence, the proponents of the New Induction should stop treating the matter unreflectively. Taking the notion for granted may lead us into distorting and overstating the argument. And if the problem is a serious epistemic worry, then it must be stated clearly. A more precise notion of relevance will allow a better understanding of what difficulties surround the New Induction as an epistemic worry. Until now, such notion is missing.

References

- BLACK, T. 2020. Contextualism in Epistemology. *Internet Encyclopedia of Philosophy*. Available at: <https://www.iep.utm.edu/contextu/>.
- CHAKRAVARTTY, A. 2008. What you don't know can't hurt you: realism and the unconceived. *Philosophical Studies*, **137**(1): 149–158. doi: 10.1007/s11098-007-9173-1.
- COHEN, S. 1988. How to be a Fallibilist. *Philosophical Perspectives*, **2**: 91–123.
- DEVITT, M. 2011. Are unconceived alternatives a problem for scientific realism? *Journal for General Philosophy of Science*, **42**: 285–293.
- DOUVEN, I. 2011. Abduction. *The Stanford Encyclopedia of Philosophy*. Spring 2011. Available at: <http://plato.stanford.edu/archives/spr2011/entries/abduction/>.
- DRETSKE, F. I. 2000. Epistemic Operator. In: F. DRETSKE. *Perception, Knowledge and Belief: Selected Essays*. Cambridge, Cambridge University Press, p. 30–47.
- EGG, M. 2016. Expanding Our Grasp: Causal Knowledge and the Problem of Unconceived Alternatives. *The British Journal for the Philosophy of Science*, **67**(1): 115–141. doi: 10.1093/bjps/axu025.
- ENFIELD, P. 2005. Review of P. Kyle Stanford's *Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives*. *Philosophy of Science*, **59**(4): 881–895.
- FEYERABEND, P. 1978. *Science in a Free Society*. London, NLB.
- FORBER, P. 2008. Forever beyond our grasp? *Biology and Philosophy*, **23**: 135–141.
- VAN FRAASSEN, B. C. 1980. *The Scientific Image*. Oxford, Clarendon Press.
- GODFREY-SMITH, P. 2008. Recurrent transient underdetermination and the glass half full. *Philosophical Studies*, **137**(1): 141–148.
- HAJÉK, A. 2019. Interpretations of Probability. *Stanford Encyclopedia of Philosophy*. Aug 2019. Available at: <https://plato.stanford.edu/entries/probability-interpret/>.
- KUHN, T. 1962. *The Structure of Scientific Revolutions*. Chicago, University of Chicago Press.
- LAUDAN, L. 1981. A Confutation of Convergent Realism. *Philosophy of Science*, **48**(1): 19–49. doi: 10.1086/288975.
- LIPTON, P. 2004. *Inference to the Best Explanation*. 2nd edn. London, Routledge.
- MAGNUS, P. D. 2006. What's new about the new induction? *Synthese*, **148**: 803–819.
- MAGNUS, P. D. 2010. Inductions, red herrings, and the best explanation for the mixed record of science. *British Journal for the Philosophy of Science*, **61**: 803–819.

- ROWBOTTOM, D. P. 2019. Extending the argument from unconceived alternatives: observations, models, predictions, explanations, methods, instruments, experiments, and values. *Synthese*, **196**(10): 3947–3959. doi: 10.1007/s11229-016-1132-y.
- STANFORD, K. 2006. *Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives*. Oxford, New York, Oxford University Press.
- STANFORD, P. K. 2015. Catastrophism, Uniformitarianism, and a Scientific Realism Debate That Makes a Difference. *Philosophy of Science*, **82**(5): 867–878.
- STANFORD, P. K. 2019. Unconceived alternatives and conservatism in science: the impact of professionalization, peer-review, and Big Science. *Synthese*, **196**(10): 3915–3932. doi: 10.1007/s11229-015-0856-4.
- WRAY, K. B. 2016. Method and Continuity in Science. *Journal for General Philosophy of Science*, **47**(2): 363–375.

Submitted on December 5, 2020.

Accepted on March 4, 2021.