

Radu Dudău, Assoc. Prof.  
Faculty of Philosophy, University of Bucharest  
Splaiul Independenței 204  
060024 Bucharest  
radu.dudau@gmail.com

## Algorithmic rivals: do they pose a threat to scientific realism?

### Abstract

*André Kukla (1994; 1996; 1998; 2000) argues that there is no reason why propositional structures mechanically generated from given theories should be denied the status of genuine scientific theories. He indicates that such constructs are empirically adequate and truth-evaluable and asks for the reasons why these are ignored in scientific practice. I undertake to offer some decisive reasons: first, I argue that, unlike genuine theories, algorithmic rivals contravene to a proper conception of physical causation; and secondly, that they are parasitic upon the theories from which they are generated. These arguments disqualify algorithmic rivals from the epistemic competition. Thus, one direct way toward establishing the underdetermination of theory choice by evidence is rebuked.*

### 1. Introduction

One of the most frequently raised arguments against scientific realism is the underdetermination argument, which consists of the following two theses:

- (1) The *empirical equivalence* thesis (EE): for any given theory, there is an indefinite number of empirically equivalent rivals.
- (2) The *underdetermination* thesis (UD): any theory  $T$  is radically underdetermined, because no possible evidence could justify it over its empirically equivalent rivals.

Obviously, the truth of UD depends on the entailment  $EE \rightarrow UD$  and on EE's truth.<sup>1</sup> It is the aim of the present paper to reject the following claim, which is an attempt to establish EE: given any theory  $T$ , there are algorithmic procedures that generate empirical equivalents to  $T$  under *all possible* evidence. I call these constructs *algorithmic rivals*, following the nomenclature of their promoter in recent literature, André Kukla. Here are Kukla's (1998; 2000) two examples of such algorithms:

- (a) Given  $T$ , construct theory  $T_1$  by taking the observational consequences of  $T$  to be true while

---

<sup>1</sup> I am not going to discuss the scruples of those realists who object that UD's reliance on EE makes the former unintelligible because, to be formulated, EE presupposes an untenable neat dichotomy between the observable and the theoretical. In line with Paul Horwich (1982), I take it that a notion of evidential underdetermination may be construed without any reference to such a distinction.

denying that  $T$ 's theoretical entities exist.

(b) Given  $T$ , construct theory  $T_2$  by taking  $T$  to be true whenever an observation takes place, while taking  $T$  to be false when no observation is going on.

The reason why  $T_1$  and  $T_2$  should be a threat to scientific realism is conspicuous:  $T$  is empirically indistinguishable from  $T_1$  and  $T_2$  under all possible evidence.<sup>2</sup> But in spite of their popularity in the philosophy of science, such algorithms are disregarded in scientific practice. Scientists do not even bother to take this sort of propositional structures into consideration. But why is it that theory-candidates that are empirically successful and truth-evaluable are constantly overlooked? Kukla (1994; 1996; 2000) repeatedly complained about the lack of a resolution to this question. He labels this dereliction the problem of *scientific disregard*:

...scientists routinely and uniformly ignore certain propositional structures that seem to have a good measure of the empirical virtues. Let us call this the phenomenon of scientific disregard. The problem of scientific disregard is the philosopher's problem of explaining how and why this phenomenon takes place. (Kukla 2000, p. 22)

I agree with Kukla that the problem of scientific disregard needs a clear answer. However, unlike him, I argue that the phenomenon of scientific disregard is in place for good reasons. But before proceeding, a few preliminaries are required.

First, a matter of terminology: the hypothesized property that algorithmic rivals are deficient about, so that they cannot qualify as genuine rivals, is called by Kukla (2002, 22) *theoreticity*. Thus, even if both  $T_1$  and  $T_2$  had truth values and were empirically successful, they would always lose against  $T$  because of their lack of theoreticity. Correspondingly, propositional structures lacking theoreticity are called *quasi-theories*. The philosopher's task is then to establish what (if anything) makes  $T_1$  and  $T_2$  quasi-theories.

I argue that scientific disregard is the right scientific attitude. I start by explaining why  $T_1$  and  $T_2$  are to be treated separately. Next,  $T_2$  is dismissed as a genuine rival on the grounds of its belonging to a family of skeptical constructs which are supposed to be left behind in the scientific realism debate. Further on, I argue against  $T_1$  that, on the one hand, it cannot offer a proper framework for causal explanation. I also show that we can often causally infer to the existence of

---

<sup>2</sup> This formulation of the UD thesis is logically stronger than the one urging empirical equivalence under the so-far-known evidence. The present discussion takes for granted that the latter, weaker version of underdetermination is not a threat to scientific realism.

entities and low-level phenomenological laws, independently of the truth of any overarching theory about them. On the other hand, even when it is explanatory and predictively successful,  $T_1$  is parasitic upon  $T$ 's conceptual resources. Each of these shortcomings stands for  $T_1$  and  $T_2$ 's lack of theoreticity, so that the latter prove indeed to be non-starters in the epistemic race.

In spite of Kukla's manifest opposition,<sup>3</sup>  $T_1$  and  $T_2$  should be treated separately.  $T_1$  belongs to a class of theory-candidates typically embraced by sceptics or agnostics about unobservable entities, who urge restriction to belief in observables. Belief in observables is taken for granted while epistemic ascent to unobservables is deemed unwarranted. Ontologically, a defender of  $T_1$  has no qualms about the assumption that there is a realm of entities existing objectively and independently of our minds, language, and theories. At stake in confronting  $T$  with  $T_1$  are preeminently epistemic issues: how – and to what extent – can we ascertain this objectively existing realm? By contrast, this does not seem to be the relevant question with respect to  $T_2$ . The latter belongs to a family of sceptical scenarios according to which the world is ontologically different from what we can normally believe on the basis of our evidence.  $T_2$  entertains assumptions different from those of  $T$  about the ontological constitution of the world, though it maintains that they make no observational difference from those of  $T$ . Thus,  $T_2$  urges a clarification with respect to the ontological set up of the world.

Accordingly,  $T_1$  and  $T_2$  ought not to be treated on a par. Their rebuttal requires a differentiated argumentation. Let us start with  $T_2$ , whose rejection turns out to be rather uncomplicated in the context of our discussion.

## 2. The dismissal of $T_2$

A salient feature of the sceptical scenarios of the  $T_2$ -family is that they are constructed by way of systematic appendages to the ontological presuppositions of  $T$ , so that our actual epistemic procedures are not affected. The supposition that the fundamental make up of the world is very different from what we can infer on the basis of empirical evidence has several famous

---

<sup>3</sup> Kukla has protested against Laudan and Leplin's separate treatment of  $T_1$  and  $T_2$ : "if every proposed algorithm for producing empirical equivalent rivals is going to require us to come up with a new *ad hoc* rule to insure its disqualification, the theoreticity argument against [the thesis of empirical equivalence] loses all credibility." (1996, p. 151). However, I argue that the distinction between  $T_1$  and  $T_2$  is principled.

representatives: the Cartesian demon story, which tells us that the world of familiar objects does not exist and that we are being deceived by a powerful demon; its modern counterpart, the Putnamian brain-in-a-vat story, which says that I am a nothing but a brain-in-a-vat artificially stimulated to have all the experiences that I would have if I had a body and interacted in the normal way with the familiar world; the claim that God created the universe as we think we know it just five minutes ago, etc.

The possibility that the world's constitution changes drastically when observation ceases, and that, after observation resumes, the familiar setup is reinstated qualifies as a scenario of this kind. As a matter of fact,  $T_2$  evokes the consciousness-based problem of measurement in quantum theory: in the absence of measurement it is meaningless to ascribe reality to the properties of a quantum system. These properties come into existence once the quantum system interacts with an observer's mind. The presupposition must be that human consciousness is a substance which, though able to interact with matter, behaves quite differently from it.

Now the only 'unproblematic' way to account for such a dramatic dependence of the world's make up on the human mind is to embrace a form of idealism, be it subjective (e.g. solipsism) or objective. Notoriously, there is no logical way of eliminating these possibilities without begging the question. But, fortunately, our approach is not dependent on such a resolution. The debate over scientific realism is about the existence of *unobservable* scientific entities, as well as about our ways of coming to know them. Realism about observables is taken for granted, along with the epistemic procedures which establish their existence. Hence, scepticism about the external world is left behind. The parties in the debate are supposed to agree about this much. Thus, we need not be bothered by seeking solutions to a problem which is not at issue.  $T_2$  can therefore be discarded as irrelevant. What is at stake in the scientific realism debate is whether the ampliative inferential procedures taken for granted at the observable level ought to establish facts about scientific unobservables. It is only  $T_1$  that may threaten scientific realism in this respect.

### **3. The dismissal of $T_1$**

I put forward two arguments against  $T_1$ . First, I show that  $T_1$  is unable to offer a proper framework for causal explanations (3.1). The argument establishes that  $T$  is always preferable to  $T_1$ : the former can perform anything that the latter can, while the latter cannot, in general, support causal explanations. Second, even if  $T_1$  explains (other than causally), this can be done only *parasitically* upon  $T$ 's explanatory resources (3.2).

### 3.1 $T_1$ is unable to give proper causal explanations

Brian Ellis (1985) defines a *causal* explanation as the one giving information about the causal processes which result in something. He emphasizes that accepting the description of a causal relation as correct involves supposing that causes are *existent* events (1985, p. 173). A similar view is defended by Nancy Cartwright (1983), according to whom causal explanations have an essential “existential component”: if we accept that an entity  $X$  is the cause of something, then we are committed to believing in the existence of  $X$ . Not so with other sorts of scientific explanation; in Ellis’s classification, a *model-theoretic* explanation will give “information about how (if at all) the actual behavior of some system differs from that which it should have ideally if it were not for some perturbing influences...”; a *functional* explanation will explain the functional roles in a system, while a *systemic explanation* will offer “information about how the fact to be explained is systematically related to other facts” (Ellis 1985, p. 173). The important point is that only in causal explanations is belief in the existence of the posited cause unavoidable.

Since the above remarks are contentious, several issues will have to be discussed: do causal explanations *necessarily* involve existential commitment? How to defend the special status of causal explanation in the face of those who reject causal talk altogether? Why be existentially committed and not, for example, follow the way of the agnostic empiricists, who use causal talk without appealing to the ontology of real causes? And even if we take causality at face value, should we not limit ourselves to relations between observable events? What is the warrant for bringing theoretical unobservables into our view of causation? Of course, the answers to these questions are not independent from each other.

Regarding the necessity of existential commitment in causal explanations, Steve Clarke (2001) aptly notices that the logic of causal explanation merely establishes facts about the language in which causality is expressed:

In itself, this shows only that ordinary language users are implicitly committed to the unreflective ‘common sense’ realism about causes. But ordinary language users may be in error in making this implicit commitment. What would be telling against empiricist antirealists would be a demonstration that it is not possible to provide an antirealist redescription of causal explanation that does not involve existential commitment. (Clarke 2001, pp. 713-714)

Indeed, antirealists about causation turn themselves occasionally to unobservables for explanatory purposes, but explicitly deny belief in such entities. For example, van Fraassen (1980) makes use of  $T$ ’s theoretical posits for purposes of pragmatic explanation. He does not accept  $T$  as

true, but only as empirically adequate, being agnostic about its unobservables. There is only one step, that separates him from  $T_1$ 's partisan: the latter explicitly rejects  $T$ 's unobservable content, while taking  $T$ 's observable consequences to be true.

But I want to show that, in causal arguments, the ontological commitment must extend beyond the logical constraints of the language one happens to use. My strategy consists of a combination of Salmon's (1984) common cause analysis, and Cartwright's (1983) and Hacking's (1983) "experimental argument." Let us first introduce some terminology and tackle the Humean question of whether causal relations mean anything above the constant conjunction of observable events. According to Salmon (1984), the constituents of the world's causal structure are: *causal interactions*, by which "modifications in structure and order are produced; *causal processes*, by which "structure and order are propagated from one spacetime region of the universe to other times and places" (p. 179); and *causal laws*, which "govern the causal processes and interactions, providing regularities that characterize the evolution of causal processes and the modifications that result from causal interactions." (p. 132). We typically observe statistical correlations between events. These correlations can indeed be sometimes accounted for exclusively in terms of observable events, as in one of Salmon's examples: Adams and Baker are students who submitted virtually identical term papers in a course. The teacher is very likely to consider it highly improbable that the papers came out that way by pure chance. Instead he countenances one of the following reasonable possibilities: "(1) Baker copied from Adams, (2) Adams copied from Baker, or (3) both copied from a common source." (1984, p. 207). In other words,

There is either (1) a causal process running from Adams's production of the paper to Baker's, (2) a causal process running from Baker's production of the paper to Adams's, or (3) a common cause – for example, a paper in a fraternity file to which both Adams and Baker had access. In the case of this third alternative, there are two distinct causal processes running from the paper in the file to each of the two papers submitted by Adams and Baker, respectively. (Salmon 1984, p. 207)

If (3) is in place, we say that there is an *indirect causal relevance* between the considered events. The common cause – the reproduction of the original paper – is connected through causal processes to each of the separate effects. Let  $T$  be the theory positing the relevant causal mechanisms.  $T$  thus causally explains the statistical correlations between the events A and B.

In the above example, A and B are the teacher's establishing that Adams and Baker have, respectively, submitted virtually identical papers. In this particular case  $T_1$  will do as well as  $T$ : Since both  $T$  and  $T_1$  account for the correlations between some observable events in terms of

observable interactions and observable causal processes, there is no reason not to take  $T_1$  (instead of  $T$ ) as the theory providing the right causal explanation. The same point applies when two events are *directly* causally relevant to each other, when A is connected to B by a causal process through which the causal influence is transmitted. This corresponds either to (1) or to (2) in the term-paper example.

However, there are many familiar circumstances under which  $T_1$ 's causal explanations clearly fail. Consider the following situation: "A stone was thrown and broke the window." As  $T_1$ 's supporter would have it, it is perfectly all right to take 'throwing the stone' as the cause, the observable motion of the stone through the space as the causal process transmitting the causal influence, and 'breaking the window' as the effect in a causal connection between observable events. In these terms,  $T$  seems to have no monopoly on causal explanations;  $T_1$  can also explain causally. Thus,  $T_1$  could do everything that  $T$  can, and so would be a priori preferable on the grounds of its ontological parsimony. But this construal misunderstands the idea of explaining causally. In the window example, one may legitimately ask, *why* a normal windowpane actually breaks when hit by a stone.  $T$ 's advocate may locate an event C on the continuous spatio-temporal process going from A to B, C consisting of the absorption of the stone's kinetic energy into the pane's molecular structure. C *screens off* A from B, meaning that knowledge of C renders A and B statistically independent. Obviously, C's description must include terminology referring to microphysical entities. Thus, knowledge of the hidden causally efficacious entities and of their causal mechanisms is inherent to the scientific practice of asking 'why'-questions.

Now scientists typically form beliefs about such entities by interacting with them under carefully designed experimental conditions, that is to say, by controlled intervention into the causal processes of which they are part. Then, based on the experimental results, they decide which manipulated entities could possibly cause the process that they have intervened in, eliminating as many candidates as the evidence allows. There are two important sets of requirements that, fortunately, can frequently be satisfied jointly: on the one hand, by suitable experimental techniques, it is possible to single out *one* of a set of plausible candidates to the status of cause. On the other hand, different experimental methods may lead us to identify the same type of entity as the most probable cause of apparently different classes of effects.<sup>4</sup> Whenever these two aspects can be conjoined, the highly probable existence of a certain type of causally efficacious entity can be safely inferred.

---

<sup>4</sup> The latter aspect was most famously documented in Perrin's 1913 book, *Les Atomes*, in which he listed thirteen different methods of determining Avogadro's number. These included measurements of alpha decay, of X-ray diffraction, of black-body radiation, and of electrochemistry.

We see that the ‘inference to the most probable cause’ (IPC) is a special sort of inference to the best explanation (IBE). IBE is the typical vehicle of scientific realists, who argue both for the probable existence of the theoretical entities posited by the best established theories, and for the approximate truth of the theories themselves. From a scientific realist perspective, IPC and IBE are both to be performed by means of the best established theory. By inferring to the existence of some causally efficacious entity, the scientific realist also infers to the truth of a broad class of theoretical sentences about the properties of that entity and its relations to other entities. By contrast, the specific move of Cartwright and Hacking is to take causal explanation as being free of any overarching theoretical obligations. They take inspiration from the experimental practice, where scientists conduct experiments without turning to a unitary theoretical framework regarding the object of their experimentation. Scientists typically resort to established phenomenological regularities embedded into fairly narrow theoretical models, some of which are incompatible. As Hacking puts it,

Even people in a team, who work in different parts of the same large experiment, may hold different and mutually incompatible accounts of electrons. That is because different parts of the experiment will take different uses of electrons. Models good for calculations on one aspect of electrons will be poor for others. ... There are a lot of theories, models, approximations, pictures, formalisms, methods and so forth involving electrons, but there is no reason to suppose that the intersection of these is a theory at all (Hacking 1983, pp. 264-265)

The point is also emphasized, among others, by Peter Galison (1987), who brings forth sociological evidence that contemporary experimental practice displays an appreciable degree of independence from theory. Experimentalists rely on independent stratagems of judgment that do not depend upon the details of any particular overarching theory.

By performing IPC outside the broader frame of a theoretical explanation, one is an *entity realist* without also being a scientific realist. To accept IPC as a legitimate form of inference is to be committed to the existence of some type of causally efficacious entities, without thereby being committed to the veracity of any particular fundamental theory about that entity. Now, to be sure, it is debatable whether any argument for entity realism can be legitimized without thereby offering implicit support to scientific realism. I actually have doubts that this is the case. For one thing, I do not share Cartwright’s motivation regarding the distinction between high-level theoretical laws and low-level phenomenological laws with respect to their truth. I take it that fundamental theoretical laws are approximately true and that they are confirmed to be so just by relying on the truth of lower-level phenomenological laws. I follow Boyd (1984) in arguing that the methodology by which empirically successful theories are produced starts from the truth of those lower-level laws;

thus, successful theories are generated which, in their turn, will be incorporated into a further improved methodology. In this dialectic, it becomes obvious that causal arguments are foundational for scientific realism.

But the demarcation between entity realism and scientific realism is not important here, as long as at least one of them has been established. It is sufficient to put entity realism on a firm basis in order to show that  $T_1$  is out of the epistemic race. Recall that  $T_1$  resorts only to  $T$ 's observables. Thus,  $T_1$  is unable to provide a framework for causal explanations. Nonetheless, there is still an important potential objection that ought to be removed before being entitled to draw the conclusion of  $T_1$ 's lack of credit. The objection is the following: not only is the choice among competing empirically equivalent theories underdetermined by the empirical evidence, but also the inference to the existence of a particular cause is so underdetermined. To the inference that an entity is causally responsible for a certain set of observed effects, the sceptic will object that competing causal explanations are equally able to explain these effects. Clarke offers a good analysis of that possibility:

It seems logically possible, for example, that the observed and experienced effects that we explain by appealing to the existence of a single entity, the electron, might actually be caused by a *cluster of microscopic particles* that are invariably joined together. Alternatively, they might be caused by *an altogether different particle* from the electron, call it the 'flectron'. A third possibility is that they might be caused by an *evil demon* who wished to deceive us into believing in electrons, and who is able to supernaturally intervene in our activities. And it seems that we have no non-arbitrary way of choosing between beliefs in the electron, as against any of these possibilities, without appealing to pragmatic considerations. (Clarke 2001, pp. 718–719, my italics)

In line with him, let us take these possibilities in turn. In case there was an alternative explanation competing with the electron one, positing a cluster of particles as responsible for the observational effects we ascribe to electrons, then we would, *ceteris paribus*, identify this cluster with 'the electron', thus starting to develop a theory about the electron's composition. If, on the other hand, an empirically equivalent theory is at issue, which claims that the relevant effects are caused by an entity with properties different from those of the electron, then some experimental error from the part of the electron's advocates should be pointed out. Otherwise, all that can be said is that the posited entity is nothing but a different label for 'electron', which is not what we expect from a genuine empirically equivalent rival. Finally, with respect to the evil demon story, we rely on the previous section's considerations.

We can now summarize the steps of the argument:

- (1) Well-established theories committed to unobservables generally allow the formulation of frameworks for causal, as well as for other forms of explanation (abstract model, functional, and systematic explanations).
- (2)  $T_1$  precludes the search for causal explanations appealing to unobservables.
- (3) Causal explanations in terms of unobservables are essential to scientific investigation.
- (4) Therefore, we should always prefer  $T$  to  $T_1$ .  $T$  provides all of the sorts of explanatory frameworks enumerated above, while, by definition,  $T_1$  bars at least the possibility of explaining causally in terms of unobservables. This shows  $T_1$ 's lack of theoreticity.

### ***3.2 The parasitism objection against $T_1$***

There is another line of criticism of algorithmic rivals that Kukla discusses, which, since the argument over  $T_2$  is closed for our purposes, I shall investigate only with respect to  $T_1$ : the so-called parasitism objection, formulated by Laudan and Leplin (1991):

[ $T_1$ ] is totally parasitic on the explanatory and predictive mechanisms of  $T$  ... a [genuine] theory posits a physical structure in terms of which an independently circumscribed range of phenomena is explainable and predictable. (Laudan and Leplin 1991, pp. 9-10)

Recall that  $T$  is assumed to be empirically successful. Then, as the objection goes, if  $T_1$  *must* refer to  $T$  in order to explain and predict,  $T_1$  is nothing but a parasitic propositional structure. But undoubtedly, science does well without such parasites. Faced with Laudan and Leplin's parasitism objection, Kukla replies that even if the initial construction of  $T_1$  makes reference to  $T$ , this does not imply that  $T_1$  cannot be characterized so as to circumvent reference to  $T$  (Kukla 1996, p. 149). He points out that simple theoretical structures can be devised, so that their observational consequences can be described independently: if  $T$  is simple enough, its observational consequences can be described without reference to  $T$ . That implies that  $T_1$  can sometimes be described without reference to  $T$ .

Concerned about the possibility that this maneuver could be generalized, Leplin rephrases the parasitism criterion. In order for  $T_1$  to be a genuine theory, he urges that  $T_1$ 's class of

consequences be specified independently of  $T$ :

Kukla wants to eliminate reference to  $T$  by specifying directly what the empirical consequences are to be. But the determination of what to specify can only be made by reference to  $T$ . That is the point of charge of parasitism. Whether or not reference is made to  $T$  in identifying its purported rival is not the proper test of parasitism. (Leplin 1997, p. 160)

This is a reply that Kukla deems too strong to be fulfilled even by accepted scientific practices. He analogizes with the search for a theory  $U$  that unifies two disparate theories  $X$  and  $Y$ . “In this case, too, ‘the determination of what to specify can be made only by reference to’ the conjunction  $X \& Y$ .” (Kukla 2000, p. 28). In other words, since  $U$  explains the same range of facts as  $X$  and  $Y$ ,  $U$ ’s ‘specification’ can be made only by reference to  $X$  and  $Y$ . It seems thus that  $U$  is parasitic upon  $X \& Y$ , which renders Laudan and Leplin’s criterion unacceptable.

Nonetheless, I see some good reasons to reject this analogy. To urge the least, a unified theory must involve the presence of a level of concepts to which  $X$  and  $Y$ ’s concepts are logically reducible. Besides, the class of  $U$ ’s observational consequences, though largely overlapping with the one of  $X \& Y$ , is expected – as Kukla himself notices – to diverge from the latter. Thus,  $U$  is an independent theory; after the theory construction process has drawn to a close,  $U$ ’s formulation does not need to involve any reference to  $X$  and  $Y$ , let alone  $X \& Y$ . On the other hand, the relation between  $T_1$  and  $T$  is very different from the one between  $U$  and  $X \& Y$ .  $T_1$ ’s reference to  $T$  is parasitic because it is *ineliminable*. If, as Kukla argues, there are cases in which  $T_1$  can be reformulated so as to circumvent reference to  $T$ , these come down to a direct search for independent empirical equivalents to  $T$ . Even if this could be done in few particular cases – namely when  $T_1$  is a ‘simple propositional structure’ – there is no warrant that this maneuver can be generalized for *every*  $T_1$ . As a matter of fact, given the complexity of modern theories, the possibility of independent scientific descriptions of their observational consequences becomes overwhelmingly implausible. Consequently, Kukla’s stricture upon the above parasitism criterion misses its mark.

More effectively, Kukla maintains that there are cases in science where  $T_1$  plays a role:

Let us grant that the parasitism criterion succeeds in eliminating [ $T_1$ ] from the ranks of genuine theoretical rival. In that case, I would argue that it is far too strong a test for theoreticity, for there are circumstances where structures like  $T_1$  have an important role to play in the game of science. Consider the following scenario. (1) Theory  $T$  has been well-confirmed, so that its empirical adequacy is widely believed; (2) it is discovered that one of its theoretical principles is inconsistent with an even more

firmly believed theory; and (3) no one can think of any way to describe the empirical consequences of  $T$  except as the empirical consequences of  $T$ . In that case, we might very well come to believe a proposition which has precisely the structure of  $T_1$ : that the empirical consequences of  $T$  are true, but that  $T$  itself is false. (Kukla 1996, p. 149)

So we have a case where  $T$  is empirically successful, but due to the inconsistency of one of its theoretical principles with a more fundamental theory, we may only retain  $T$ 's observational consequences. Although we are assured of the fact that  $T$  must be rejected, what we do not know is how much of it could be retained. Yet, it is a sure thing that  $T$ 's observational consequences must not be rejected, so Kukla's scenario is logically sound.

My reaction, though not logically decisive, is to point to a double implausibility of this scenario. First, that the divide between the rejected and the retained parts of the theory should correspond exactly to the observable/unobservable divide strikes me as highly unlikely. If there are strong reasons to believe that  $T$  is wrong, we expect this to show up at the observable level, as well. Second, assuming that  $T$  is a well-established theory, given the coherence constraints upon the process of theory-construction, it is improbable that one of  $T$ 's theoretical principles will be inconsistent with an even more fundamental principle of that particular field. The consequence of these plausibility considerations is that, in the overwhelming majority of cases, Kukla's latter scenario cannot occur. Hence,  $T_1$  cannot play a systematic role in science.

Admittedly,  $T_1$  can occur by accident; yet this is by no means a sufficient ground for the algorithmic generation of  $T_1$  to become a methodological *rule* of theory construction. Still, again, no matter how high, methodological unreasonableness does not entail logical impossibility. What is eventually left for the realist in order to account for the very low probability of  $T_1$ 's accidental occurrence is simply to bite the bullet and admit that he can live with (few) situations in which  $T$  happens to be false. For after all his doctrine judiciously requires that *most*, not *every single one* of the well-established scientific theories be truthlike. Now, of course, by allowing that  $T$  might be false, scientific realism has to account for the threat of the so-called *pessimistic meta-induction*: if many theories of the past turned out to be false, why should we believe that our current ones fare better? Shouldn't we be wise and learn from history that our current theories will probably turn out false as well? Within the confines of this paper, I can merely sketch an answer to this objection. I basically follow Devitt (1991, pp. 162-165) in contesting the inference from the falsity of past theories to the probable falsity of current theories. Why should this be so? After all, scientific methodology has constantly produced better and better theories committed to unobservables. These,

in their turn, have led to an improved scientific methodology.<sup>5</sup> Thus, on the one hand, the truthlikeness standards became increasingly high. This is a relevant fact, given that the very evaluation of past theories is made from the perspective of current science. Hence, we can establish the falsity of the past theories only insofar as we accept the accuracy of modern theories' descriptions.<sup>6</sup>

Besides, as Kitcher (1993, pp. 140-149) shows, we can identify and eliminate many falsities of past theories, as well as establish correspondences between their successful problem-solving schemata and the mechanisms posited by contemporary theories.

#### **4. The insufficiency of Kukla's solution to the problem of scientific disregard**

Kukla explicitly rejects any solution to the problem of scientific disregard in terms of theoreticity. He argues that a Bayesian account offers a straightforward response. He explicates the fact that algorithmic rivals are disregarded as having zero prior probabilities:

According to Bayesianism, prior probabilities are free, subject only to the constraint of probabilistic coherence.... Further constraints apply only to how our opinions must change with the receipt of new information. The requirement of coherence, however, already has the consequence that infinitely many theories must be ascribed priors of exactly zero (or priors whose infinite sum converges on a number less than or equal to 1). Given that we have to assign zero probabilities, it's not at all surprising that there are theories which are utterly disregarded. It's only rational not to waste any time on impossibilities. (Kukla 2000, 31)

Kukla maintains that Bayesianism has a built-in explanation for scientific disregard. Nonetheless, the question to be raised immediately is, why should  $T_1$ , instead of  $T$ , be assigned zero prior probability? Kukla believes that the answer to this question is not important. He permits the utterly arbitrary choice of those theories which are regarded as impossible: "Given a different roll of the dice, we might have disregarded  $T$  and pursued  $T_1$ " (Kukla 2000, 31). All that matters, in his opinion, is that the disregarded hypothesis is deemed to be impossible. Yet, the subjectivity involved in the ascription of prior probabilities cannot be understood as completely free of methodological constraints. I have provided independent reasons why it is always  $T_1$  which is

---

<sup>5</sup> The dialectic between scientific methodology and truthlike theories was aptly emphasized by Boyd (1984).

<sup>6</sup> James McAllister (1993, p. 212) argues that "the theories deemed successful in the history of science were deemed to be so on the basis only of a set of criteria constructed in the light of imperfect knowledge about the properties of the properties of theories."

disregarded, and not  $T$ . Thus, I have shown that we have sufficient grounds to assign  $T_1$  zero prior probability. Contrary to Kukla, I gather that the theoreticity constraints are a useful complement to his Bayesian account. He claims that the advocates of the theoreticity approach look at the propositional structures they want to disregard and, to this purpose, come up with *ad hoc* requirements. Yet it is a matter of fact that algorithmic rivals cannot explain causally and that they are parasitic.

Kukla wants to derive support for the underdetermination thesis from his Bayesian account of scientific disregard. However, given my arguments to the contrary, he cannot derive more underdetermination than he has already put in by raising all methodological constraints (except those of probabilistic coherence) upon the ascription of prior probabilities.

### **Acknowledgments**

I am very grateful to Erik Olsson, Ludwig Fahrbach, and Daniel Cohnitz for helpful comments on earlier drafts of this paper.

### **References**

- BOYD, Richard (1984), „On the current status of scientific realism”, in: Jared LEPLIN (Ed.), *Scientific Realism*, Berkeley: University of California Press.
- CLARKE, Steve (2001), „Defensible territory for entity realism”, in *British Journal for the Philosophy of Science* 52, pp. 701-722.
- DEVITT, Michael (1991), *Realism and Truth*, Oxford: Basil Blackwell.
- ELLIS, Brian (1996), „What science aims to do,” in: David PAPINEAU (Ed.) *The Philosophy of Science*, Oxford: Oxford University Press.
- FRIEDMAN, Michael (1983), Book review: „Bas van Fraassen’s ‘The Scientific Image’,” *Journal of Philosophy* 79, pp. 274-283.
- GALISON, Peter (1987), *How Experiments End*, Chicago: University of Chicago Press.
- HACKING, Ian (1983), *Representing and Intervening. Introductory Topics in the Philosophy of Natural Science*, Cambridge: Cambridge University Press.
- HORWICH, Paul (1982), „How to distinguish between empirically indistinguishable theories,” in *The Journal of Philosophy*, vol. LXXIX, 2, pp. 61-77.
- KITCHER, Philip (1993), *The Advancement of Science: Science Without Legends, Objectivity Without Illusions*, New York: Oxford University Press.
- KUKLA, André(1994), „Non-empirical theoretical virtues and the argument from underdetermination,” *Erkenntnis* 41, pp. 157-170.
- KUK, André(1996), „Does every theory have empirically equivalent rivals?,” in *Erkenntnis*44, pp. 137-166.

KUKLA, André (1998), *Studies in Scientific Realism*, Oxford: Oxford University Press.

KUKLA, André(2000), „Theoreticity, underdetermination, and the disregard for bizarre scientific hypotheses,” in *Philosophy of Science* 68, pp. 21-35.

LAUDAN, Larry. & LEPLIN, Jared (1991), „Empirical equivalence and underdetermination”, *Journal of Philosophy* 88, pp. 448-472.

MCALLISTER, James (1993), „Scientific realism and the criteria for theory-choice”, *Erkenntnis* 38, pp. 203-222.

SALMON, Wesley C. (1984), *Scientific Explanation and the Causal Structure of the World*, Princeton: Princeton University Press.

VAN FRAASSEN, Baas C. (1980), *The Scientific Image*, Oxford: Clarendon Press.

#### **Note on the contributor**

Radu Dudău received Dr. Phil. from the University of Konstanz. He is currently associate professor at the Faculty of Philosophy, University of Bucharest. His current research interests include philosophy of science, epistemology, political philosophy.