Interview with Richard Dawid

Luca Moretti: Dear Richard, I'm delighted that you accepted to be interviewed by *The Reasoner*. Of course, I will ask you questions about your recent book, *String Theory and the Scientific Method* (CUP 2014). Before that, I would like to know about your intellectual story. I recall that you started your career as a physicist but converted to philosophy after a while. Could you tell us something about these events?

Richard Dawid: My pleasure, I'm looking forward to the interview. After my PhD, I spent two years as a high-energy physicist at Berkeley. That period, the late 1990s, was a particularly fascinating time for String Theory (ST). Some new conceptual ideas developed in those years substantially changed the understanding of ST and paved the way for its further development until today. Watching those developments, I felt that they raised novel and interesting philosophical questions at various levels. Thinking about them eventually made me switch from physics to philosophy.

LM: Did you have any background in philosophy?

RD: I was always interested in philosophy. I had read some philosophy, had joined a philosophy discussion group during my PhD in Vienna and had attended a few philosophical university seminars. But at the time I decided to enter philosophy, my knowledge was quite haphazard. Thinking back today, I'm a little stunned on what meager basis I made that decision.

LM: I remember you told me that you emailed eminent philosophers for advice. What did you ask them? Did you manage to meet any of them?

RD: Right, once I had developed some first philosophical ideas about what I intended to do, I wanted to clarify two things before seriously moving into philosophy. First, I wanted to know whether my ideas made a little sense to genuine philosophers. Second, I wanted to know whether it was fun discussing with genuine philosophers. Since the only philosophers I knew at the time were really famous ones, whose books I had read, I just emailed three of them: Hilary Putnam, Bas van Fraassen and Hartry Field. I asked them whether they were willing to talk to me about my ideas. All three were extraordinarily kind and agreed to meet. Unfortunately, Putnam had to cancel the day before we met for urgent personal reasons, but I met van Fraassen and Field and presented a sketch of my ideas to them. Van Fraassen was very supportive and gracious and seemed genuinely interested, a real pleasure to talk to. Field told me right from the start that he wasn't interested in the subject but was ready to comment on the general soundness of my reasoning, which he did with impressive acuteness. Both meetings substantially strengthened my conviction that it made sense for me to turn towards philosophy.

LM: Were your first philosophical ideas already about ST and the no-alternative argument?

RD: I was mainly interested in two issues that were both related to ST.

LM: Perhaps, before continuing, it would be helpful if you could shortly explain what ST is.

RD: ST aims at providing a unified theory of all physical interactions. The nuclear interactions, which are crucial for understanding microphysics, are today described by gauge field theory, which is based on the principles of quantum mechanics. Gravitation is described by general relativity. A coherent overall theory that covers both regimes faces deep conceptual problems. There are reasons to believe that ST can solve those problems. ST starts from the basic idea that elementary objects are not point-like objects, as gauge field theory assumes, but one-dimensional strings. Those strings are taken to be so small that their extendedness cannot be measured by present day experiments. But if ST is right, the movements and topological characteristics of strings can explain all observable properties of elementary particles.

LM: Thanks. Please now let's go back to my original question.

RD: Yes. First, I was interested in the phenomenon of string *dualities*. In ST, one encounters the phenomenon that seemingly very different realizations of the theory after close inspection turn out to be dual to each other. If two theories or models are dual to each other, they are related in a specific way that implies that they are empirically equivalent. Dual theories or models can be different in all respects normally taken to specify the ontology of a physical theory. They can imply different symmetry structures, different spacetime structure, different dimensionality of elementary objects, different kinds of interaction and so on. Duality relations even reach out beyond the limits of ST proper: it turns out that in specific contexts a string theoretical description is dual to a purely field theoretical one that doesn't contain any strings. Duality relations are abundant in string physics and constitute one of its core characteristics. At a philosophical level, duailities are fascinating for example because they seem to offer a straightforward argument *against* scientific realism: if I can move from a description that posits a certain set of fundamental objects to another one empirically equivalent that posits an entirely different set of fundamental objects, and if my theory suggests that such correlations are one of its core characteristics, a realist interpretation of any set of fundamental objects seems at variance with spirit and content of the theory.

LM: I can see it. I wonder why antirealists have never mentioned this intriguing argument.

RD: That's a good question. I first made this argument in a paper in 2003. The same point was emphasized later by Dean Rickles and Keizo Matsubara in their work on the philosophy of ST. But it was never picked up in the general realism debate. I think one reason is that philosophers of science mostly take ST as an unconfirmed speculation that, as such, can have no serious implication in philosophy of science. Which brings me directly to the second important philosophical issue related to ST I was and am still interested in. Despite the fact that ST hasn't found empirical confirmation, string theorists have a conspicuously high degree of trust in it. Clearly they *don't* understand ST as a mere speculation. Thus with ST, an empirically unconfirmed theory has acquired the position of a conceptually dominating force in fundamental physics. I think that this requires a substantially altered philosophical concept of scientific theory assessment and confirmation to account for this novel situation.

LM: So we have arrived at the topic of your recent book. I remember you told me that the original title was 'Delimiting the Unconceived'. Why did you choose just this title?

RD: Yes, my original title idea was 'Delimiting the Unconceived'. Nick Gibbons, the CUP editor, thought that for those who haven't already read the entire book that title was overly enigmatic. Today I think he was absolutely right. Still, the phrase 'delimiting the unconceived' catches quite well the basic idea of the book. All of us, scientists as well as everyone else, deal with the world based on theories we have developed about it. We know, however, that there are many other possible and potentially important theories we haven't thought of yet.

LM: This has been forcefully argued for by Kyle Stanford (2006). His point is that we can inductively infer from examining past science that our new theories are probably underdetermined by empirical data even if we are actually unable to think of the alternative theories that engender the underdetermination.

RD: Yes. The 'canonical' understanding would be that we know nothing about this realm of 'the unconceived'. My book argues that this is *not* true. We do know something about the unconceived. We don't know what it contains, obviously, but we can understand something about its limits. From our observations about the world we can learn something about the size of the spectrum of possible scientific theories that we have not yet developed. At its core, the book is an investigation into how this can work.

LM: So your book aims to answer Stanford's new underdetermination argument from unconceived alternatives. Your point is–it seems to me–that at least in the case of ST there is probably no alternative–not even an *unconceived* alternative.

RD: The book argues that there can be a scientifically viable line of reasoning that leads to that conclusion even in the absence of empirical confirmation. To evaluate the strength of such reasoning in a specific case is up to the involved scientists. Eventually, that is the conclusion, yes. But in order to develop the philosophical point, get there, one has to take a number of intermediate steps with respect to understanding various facets of underdetermination.

LM: I guess many of the difficulties to get to that conclusion hinge on the notion of an *alternative* theory. Cannot one argue that there are always *sceptical* alternatives–for instance the brain-in-a-vat hypothesis–or that we can produce alternatives by *conventionalist* manoeuvres, say, by changing the value of physical constants?

RD: You're right, before assessing the number of alternative theories, it is necessary to specify what counts as an alternative. That specification crucially relies on what we want achieve by counting alternatives. Let us go back to the initial observation that string theorists trust their theory in the absence of empirical confirmation. Why do they do that? The answer in my recent book is that they do so based on their assessment of underdetermination: they believe that the chances for a viable alternative to ST are small, from which they conclude that, assuming there is a viable scientific theory of all interactions at all, ST (or whatever ST ends up being when fully developed) is likely to be that theory. Note that physicists are not interested in the realism question here. They are interested in the more modest question whether ST is consistent with the empirical data at the theory's characteristic scale. This means that, when counting possible alternatives we should only count alternatives that can be empirically distinguished from ST at its characteristic scale. So the theories that are empirically equivalent to each other should be counted just as the same theory. Furthermore, we should only be interested in theories that pass for *scientific* in the eyes of physicists. If we have reasons to expect the number of alternatives of that kind to be very small, we have reasons to have trust in ST even in the absence of empirical confirmation. Based on similar reasoning, it doesn't make any sense to count theories with different parameter values as different theories. When a physicist assesses the viability of a theory with a free parameter whose value has not yet been fixed by empirical data, she does not insist on a specific parameter value. Therefore, her assessment of underdetermination will be based on a theory individuation that subsumes all parameter values under the same theory.

LM: I see where you're going. However, one might still doubt that if we have reasons to expect that the number of proper alternatives to ST is very small, we have evidence for ST in the absence of empirical confirmation.

RD: Well, a precondition for making this epistemic connection is to have trust in the success of the scientific method in the given context. Based on our observation that physicists have so often found viable scientific theories within the scientific contexts they were investigating, we can assume that there is some viable scientific theory for the contexts proper to ST as well. On that basis, we can say: if there are no scientific alternatives to ST, this theory must be viable. If there are very few alternatives, there should be a decent chance that, when developing ST, physicists have picked the viable theory. If there were a wide range of alternatives, however, knowing this fact wouldn't instill significant trust in ST.

LM: This is my last question. You have a paper forthcoming in the BJPS coauthored by Stephan Hartmann and Jan Sprenger titled 'The No Alternatives Argument'. Could you tell us what it is about?

RD: The topic of the paper emerges from the context we were discussing. If scientists assess the number of possible alternatives to their theory, how do they do it? In my book I identify three main argumentative strategies to that end. The most direct strategy is based on an inference from the observation that scientists haven't found any viable alternatives to the theory in question to the statement that there probably are no or few alternatives. I call this inference the 'no alternatives argument'. A second argument is based on the observation that the theory under scrutiny provides explanations of phenomena or conceptual characteristics of predecessor theories it was not developed to explain. And a third argument is based on the observation of a tendency of predictive success in the research field. Now an interesting question arises: what status can we attribute to such reasoning? Can we understand it as a form of theory confirmation? In the paper with Stephan and Jan we analyze this question for the case of the no alternatives argument in a Bayesian framework. We find that under very mild and plausible assumptions the no alternatives argument does amount to theory confirmation. This is interesting because Bayesian confirmation is normally taken to rely on empirical data predicted by the confirmed theory. But the observation that scientists haven't found alternatives to, say, ST cannot be predicted by ST itself. Still, it turns out that it confirms the theory in question. The paper also shows that the no alternatives argument on its own, though formally leading to confirmation, is ineffective because it does not allow assessing the significance of that confirmation. Thus, in order to have relevant and substantial confirmation, at least one of the other two argumentative strategies must be deployed in conjunction with the no alternatives argument.

LM: I would like to thank Richard for this interesting chat.

RD: Thanks, Luca, it was a lot of fun.