Making Sense of Non-Factual Disagreement in Science

Naftali Weinberger[1] and Seamus Bradley[2]

[1]Munich Center for Mathematical Philosophy, LMU Munich;
Contact: Ludwigstr. 31, 80539 München, Germany;
naftali.weinberger@gmail.com
[2]University of Leeds

February 4, 2020

Philosophical discussions of disagreement typically focus on cases in which the disagreeing agents are aware that they are disagreeing and can pinpoint the proposition that they are disagreeing about. Scientific disagreements are not, in general, like this. Here we consider several case studies of disagreements that do not concern first-order factual claims about the scientific domain in question, but rather boil down to disputes regarding methodology. In such cases, it is often difficult to identify the point of contention in the dispute. Philosophers of science have a useful role to play in pinpointing the source of such disagreements, but must resist the temptation to trace scientific debates to disputes over higher-level philosophical accounts.

1 Disagreement

Here's a standard set up for philosophical discussions of peer disagreement: "Suppose that you and I have been exposed to the same evidence and arguments that bear on some proposition... Suppose further that neither of us has any particular advantage over the other when it comes to assessing considerations of the relevant kind, or that he or she is more or less reliable about the relevant domain... Nevertheless, despite being peers in these respects, you and I arrive at different views about the question on the basis of our common evidence." (Kelly, 2005). Given this setup, philosophers then debate how one peer should rationally respond to discovering this disagreement. For example, should she be just as confident in her position as she was prior to discovering the disagreement?

As the quotation makes clear, the philosophical literature on disagreement typically focuses on cases where, when agents disagree, they are aware that they disagree and know what they are disagreeing about. We claim that disagreement in the sciences – particularly in interdisciplinary contexts – often has neither of these features. That is, scientists can in fact disagree without really being aware of the disagreement, and even when it becomes apparent that there is a disagreement, it isn't always clear (to those disagreeing) what is actually being disagreed about. Such cases raise novel problems that have not been considered in the current literature on disagreement. First, we need a way to account for the lack of transparency regarding the source of the disagreement. Here we identify one important source of this opacity: disagreements may result from differences in the background methodological stances of those disagreeing, rather than being about some fact within the domain being investigated. Such disagreements arise especially in interactions between scientists from different disciplinary backgrounds, where it is sometimes difficult to disentangle factual disagreements about the domain being studied from disagreements about methodology.[1] Second, in cases where those disagreeing are seemingly unable to articulate the source of the dispute, the question arises as to who would be justified in doing so and as to what methodology would be appropriate. We argue that philosophers can play such a role, but only if they are careful to resist the distinctly philosophical temptation to trace scientific debates to disputes over higher-level explanatory or methodological disagreements, without also considering more pragmatic factors that can lead individuals with different disciplinary backgrounds to disagree, miscommunicate, and/or talk past one another.[2] Understanding these factors requires careful attention to case studies, of which we will consider several in detail. Finally, there is the question of how such disputes are to be resolved.

In what follows, we distinguish between disagreements regarding the facts that fall within a science's domain of study and disputes about methodology. But we do not want to presuppose that methodological disputes cannot also be (or result from) factual ones. To have a neutral way of talking, we will refer to the facts studied within a domain – as opposed facts about a field itself or its methods – as "first-order facts". Philosophers typically represent first-order facts using propositions, and while such disagreements raise a host of scientific and epistemological questions, there is little uncertainty regarding the space of possible resolutions. If two agents disagree about (or assign different credences) to some proposition P, then the space of possible resolutions is constituted by the set of ways that the agents might change (or refuse to change) their belief regarding the truth or probability of P. In contrast, when a scientific dispute does not concern first-order facts – or where it is at least uncertain which facts are in dispute – it is far from obvious what a resolution might look like. If, for instance, a dispute concerns which of two frameworks is appropriate for modeling a phenomenon, there may exist no broader neutral framework

[1] The cases we consider also diverge from Kelly's in that it is not clear whether the different parties can be said to have shared evidence, or the same capacities to reason about the topic or to even be disagreeing about particular propositions (since the debates are often over methodologies). We bracket these differences in what follows, since our aim here is not to contribute to the existing debate over peer disagreement, but to note ways in which scientific disagreement differs from the standard characterization of disagreement in the literature.

[2] For further discussion of the important roles philosophers can play in the sciences, see Laplane et al. (2019); Andersen et al. (2019).

within which to compare the positions, and there may be no principled way to "split the difference". While we here offer no general account of when and how disputes that do not concern first-order facts can be resolved, the case studies we consider provide compelling evidence that such disputes are widespread, and make it clear why it can be hard to see what a resolution might look like.

In claiming that the resolution of scientific disputes is not simply a matter of the disputants weighing their evidence for their positions, we join a philosophical tradition going back at least to Kuhn (1962). Kuhn's emphasis on the role of non-epistemic values in scientific disputes has recently been more thoroughly explored in the literature on values in science (see, e.g., Douglas (2009)). While this literature has correctly emphasized that scientific disputes cannot be reduced to isolated disagreements about first-order facts, it nevertheless often presupposes transparency regarding the point of dispute. That is, it presupposes that it is clear what is being disputed, but unclear what types of evidence are relevant for resolving the dispute. In the cases we consider, in contrast, there is ambiguity regarding what is being disputed. Additionally, while disputes over non-epistemic values are one type of dispute that cannot be traced to disputes about first-order facts, they need not be the only type. While methodological and disciplinary disputes can in some cases derive from differences in non-epistemic values, we see little evidence that such differences play an important role in the cases we consider.

In the epistemological literature on disagreement, Rowbottom (2016) has recently raised the question of what it means for two agents to disagree. Like Rowbottom, we are similarly raising a question concerning the range of cases in which agents might be said to agree or disagree. But his discussion – like the disagreement literature more generally – is held against a backdrop in which we know the propositions under dispute and can ask about the agents' credences towards said propositions. We see our discussion as providing an independent basis for uncertainty regarding whether certain apparent disputes count as genuine disagreements.

To summarize: non-factual disagreement is a widespread phenomenon in the sciences, and one that deserves further attention. In this paper we will show that this phenomenon is widespread, and we will provide some suggestions as to how philosophers might facilitate better communication among groups of scientists.

The paper is organized as follows. Section 2 considers the consistency debate in epidemiology, which we argue results from a difference between the way that theoreticians and experimentalist think about models. This section serves as both a first-order contribution to the debate about the consistency assumption, and as our first case study of non-factual disagreement. Section 3 uses the ethnographic research of MacLeod and Nersessian to highlight similar communication problems among experimentalists and mathematical modelers in systems biology. Section 4 uses the case study of econophysics to highlight a distinct but related source of non-factual scientific disputes. Specifically, different groups of modelers use different kinds of idealizing assumptions. This case study further expands the range of disciplines where non-factual disagreement is an issue, and also demonstrates that the phenomenon is not restricted to communication failures between theorists and experimentalists: different groups of theorists can also get caught out by this phenomenon. Section 5 provides a final case study based on Fagan's work on different approaches to stem cell biology. Fagan's work provides an example both of a failure of scientists to communicate, as well as a philosophical attempt to bridge this gap. Section 6 takes a step back and considers how philosophers of science can play a positive role in helping scientists understand and resolve non-factual disagreements. Section 7 concludes.

2 The consistency debate in epidemiology

We begin with a debate in the epidemiology literature surrounding a modeling assumption called "consistency". Before delving into this literature, which will be largely unfamiliar to philosophers, we will first briefly consider a similar issue that has been raised by philosophers of science. It is increasingly common to explicate causal claims in terms of the possibility of changing an effect E via an ideal intervention on its cause C. One condition of an ideal intervention on C with respect to E is that any influence of the intervention on E must be *via* C (Pearl (2009); Woodward (2003)). That is, for an intervention on C to be ideal with respect to E, it cannot influence E by causal pathways not going through C. As Spirtes and Scheines (2004) emphasize, this condition rules out "ambiguous manipulations", for reasons we will presently explain. As we will see, the problems raised by such manipulations overlap significantly with those arising in the debate over consistency.

Spirtes and Scheines' example of an ambiguous manipulation is an intervention on the variable cholesterol. Will lowering cholesterol reduce the risk of heart disease? While this seems to be a well-posed question, talk of interventions on cholesterol raises potential difficulties. There are two types of cholesterol: LDL, which raises the risk of heart disease, and HDL, which lowers it. If one were to ask about the effects of an intervention on cholesterol without further specifying whether the intervention changes cholesterol by changing LDL or HDL (or some determinate mixture thereof), such an intervention would not be ideal with respect to heart disease. This is because whether the intervention raises or lowers one's risk of heart disease depends not merely on the amount by which it raises one's cholesterol, but further on what percentage of the increase was an increase in LDL as opposed to HDL. One can get around this issue by either replacing the variable cholesterol with the variables LDL or HDL, or by precisely specifying the intervention so that it is clear how it changes LDL and HDL. But the key point is that without such moves, talk of interventions on cholesterol – and thus, on these accounts, of the causal effect of cholesterol – is ambiguous.

In the consistency debate, there is a similar concern about whether variables that do not allow for unambiguous manipulations can be causally modeled. Yet the modeling framework in the background is different. The epidemiologists in the debate rely on the "potential outcomes" framework (Rubin, 1974). In this framework, causes are referred to as treatments, effects as outcomes, and $Y^i_x = y$ indicates that if individual i in a study were to receive treatment $X = x$, she would have the outcome $Y = y$. The potential outcome framework is inter-translatable with graphical frameworks more familiar to philosophers (Pearl, 2009; Spirtes et al., 2000). A key difference between the approaches is that while graphical ones determine whether the effect of X on Y is observationally or experimentally identifiable based on features of the graphical causal representation, the potential outcomes framework treats potential outcomes such as Y_x^i as undefined primitives, and provides assumptions under which these quantities can be measured. *Consistency* is one of these assumptions, and will be the focus of the present discussion. A separate (and more widely appreciated) assumption is *exchangeability*, which when satisfied, ensures that the relationship between X and Y is unconfounded.[3] Here we will not review exchangeability, but simply flag it to indicate that the issues regarding consistency have nothing to do with possible confounding.

The consistency assumption originates in Gibbard and Harper's discussion of counterfactuals (1978). It says that if it is the case that some event C in fact obtains, and it is also the case that were C to obtain, then E would as well, then E in fact obtains. In the potential outcomes framework, this is formulated as saying that if individual i *in fact* receives treatment $X = x$, then i's outcome is Y_x^i – i.e. the outcome i would have were X to equal x. The consistency assumption links counterfactual claims to what is in fact observed: $Y_x^i = y$ is understood as a counterfactual claim saying "if i were to receive treatment $X = x$ then the outcome would be $Y = y$", and consistency then says that if the antecedent is true in the actual world, then the consequent must be true as well.

While the consistency assumption may seem tautological, it has been a source of controversy in the epidemiology literature. Cole and Frangakis (2009) argue that consistency is in fact an empirical assumption, since saying that were individual i to receive treatment x she would have outcome Y requires that she would have outcome Y given any version of the treatment. The issue here is the same as in the discussion of ambiguous manipulations – if there are different versions of the treatment that have different effects, then the fact that a person was assigned a treatment does not fix her response to it, even if all other factors are held fixed and the world is deterministic. Since many treatments will potentially be ambiguous, whether consistency in fact obtains should be treated as an assumption rather than as a definition or axiom.

In a comment on Cole and Frangakis (2009), VanderWeele (2009) further claims that consistency can be decomposed into two assumptions. The first is that the counterfactual relationship Y_x^i is invariant across a range of versions of the treatment. The second is that the observed value of Y given that i receives one of these versions of the treatment is in fact Y_x^i. The reason Vanderweele takes the latter to be an assumption is that he sees the potential outcome Y_x^i as indicating the outcome for when one is assigned treatment x, and this outcome might not be the same as the one resulting from one voluntarily choosing to take the treatment. For example, the effect of a job-training course on employment may differ depending on whether one is forced to take the course or attends voluntarily. Note that the issue here is not confounding. That is, it is not that the group of individuals who would volunteer for the course would be causally different from those

[3] See (Pearl, 2009, section 6.5.3) for the relationship between the notion of exchangeability in the potential outcomes framework, and de Finetti's distinct statistical notion.

who would go only if compelled. Instead, the concern is that even for a single individual, being forced to take the course might lead to resentment that would make the course itself less effective (cf. Cartwright (2012), section 8).

In contrast to the epidemiologists just considered, Pearl (2010) argues that we should treat consistency as a theorem rather than an assumption. Instead of introducing potential outcomes as undefined primitives whose contents are to be clarified by the surrounding assumptions, Pearl at the outset defines expressions such as Y_x^i as representing counterfactuals about what Y would be were X to be x, and gives rules for evaluating these counterfactuals using causal graphs. Given these rules, consistency follows as a theorem. So what about alternative versions of the treatment? Pearl's answer is that if an action has effects other than via the treatment, these should be specified as part of the counterfactual. If assigning people to take a course will influence future employment not merely via course attendance, but also via resentment, then one must take into account the joint effects of attendance and resentment in evaluating the counterfactual. If, for example, the resentment resulting from the assignment would nullify any benefit from the course and result in the individual being unemployed, then it would simply be false to ignore the resentment and to assert that were the individual to take the course, she would be employed.

Here we've reached a level of abstractness at which it becomes difficult to determine what, if anything, the parties are disagreeing about. To see what is at stake, it helps to note why, in the absence of further clarification, each side in the discussion might initially be inclined to think that the other is saying something false. If one interprets counterfactuals the way that Pearl does, then the claim that there cannot be causal counterfactuals about treatments that influence the outcome via multiple avenues sounds false. It is true that if one accepts the particular counterfactual $Y_x^i = y$, one cannot also accept that there is some way that X influences Y such that it might be the case that $X = x$, but $Y \neq y$. But, for Pearl, this is just to say that one should not assert counterfactuals that one knows not to be true. Once one does accept that $Y_x^i = y$, there is no additional assumption needed to establish that there are no problematic avenues. On the other hand, potential outcomes enthusiasts do not build a logic of counterfactuals into the notation, so such additional assumptions are required. For them, when Pearl claims that counterfactual relationships obey consistency, he appears to be helping himself to knowledge that goes beyond what we would typically be able to infer from a given experiment.

Of course, Pearl is aware that if one treats potential outcomes as primitives, one needs to add consistency as an assumption. His position is that it is preferable to not treat potential outcomes as primitives, but rather to model them using counterfactuals at the outset, which gets one consistency for free. So what is being disputed here? Upon inspection, the dispute results from differing perspectives on the relationship between models and experiments. The tradition of treating potential outcomes as primitives is motivated by the idea that we can treat the relationship between a treatment and outcome in a particular experiment as an observed random variable. As such, Y_x^i is linked to a particular experiment rather than to an invariant relationship in the world that licenses counterfactuals, and any generalizability beyond that experiment is only licensed by additional assumptions. In fact, even in a particular experiment, there can be ambiguity in how to characterize the treatment.

In contrast, Pearl's models are not intended to capture particular experiments, but to represent counterfactual causal relationships among variables. He does not deny that discovering such relationships can be very difficult. But he sees it as crucial to first have models and semantics for the relationships that one is trying to discover, and then to model these difficulties as deviations from the ideal case. In advocating the treatment of consistency as a theorem, he has in mind a comparison to theorems in geometry such as the Pythagorean Theorem. Such theorems concern precisely defined ideal objects that do not exist in the real world. But the use of such objects enables one to explore abstract geometrical properties "while delegating the task of assessing the practical applicability of such properties to those who are more intimately familiar with the details of each specific application" (874).

We suspect that practitioners will not be content to think about their work as merely that of assessing the applicability of beautiful abstract models. Whether or not one finds Pearl's analogy apt, it should be noted that any non-operationalist theory of measurement will posit some conceptual distinction between the method of measurement and the quantity being measured. So, from an external philosophical perspective, one can see the consistency debate as one about how far our theoretical posits can and should diverge from our measurements. The position of the consistency-is-an-assumption advocates, if taken to the extreme of saying that we should not represent anything going beyond the content of our experiments, begins to look like an untenable form of operationalism. Pearl's position, when presented in a way that treats thorny and difficult empirical questions as secondary issues, begins to look like a realism achieved without the honest toil required to ground it empirically.

Of course, the reason why scientists debate consistency is not because they have a stake in whether operationalism is correct. In our view, the debate is best seen as a dispute regarding which of two research strategies is the more fruitful one to pursue. For Pearl, those who treat consistency as an assumption are ceding their ability to, in general, look at a causal model and read off its testable implications – for them whether a model obeys consistency is not built into the counterfactuals themselves, but must be established on a case-by-case basis. The advantage of Pearl's approach is that it gives causal counterfactuals a univocal interpretation across all models, and it neatly distinguishes the specification of a counterfactual quantity from issues of how it is to be established. In contrast, for those who advocate treating consistency as an assumption, Pearl's proposal that we should treat consistency as a theorem and only represent counterfactual relationships that obey consistency significantly downplays the difficulties involved in establishing such counterfactuals based on experiments. Particular experiments will at best establish a relationship between an outcome and a particular version of the treatment and thus caution must be exercised in making claims about the outcome given alternative versions of the treatment.

The sense in which the consistency debate is in fact a debate is that each side has a difficult time in seeing how the other side's strategy could work. Both sides see the other as making the wrong tradeoff between having robust theoretical concepts and having concepts that more closely reflect the quantities measured. Here we will not weigh in on what the appropriate tradeoff should be, or even on whether different tradeoffs might be more appropriate given different aims. We use this debate to illustrate how there can be disagreements in science that do not concern first-order scientific facts, but rather reflect a dispute over strategies. Such disputes are not easily resolvable, and often reflect sociological divides between disciplines as much as theoretical commitments. Yet when methodological disputes are not acknowledged, this can lead to scientists within or across disciplines talking past one another, and consequently taking themselves to disagree on the first-order facts. Here philosophers of science, who are trained to think about the logical structure of theories and models, have a potentially useful role to play in helping scientists avoid such miscommunications.

The debate just considered may seem so esoteric as to be of little practical consequence. Yet the issues that arise dispute between Pearl and the potential outcomes theorists keep arising as communities preferring each framework approach new problems. For example, Hernan and VanderWeele (2011) have used the consistency debate as a launching pad to consider issues of causal generalizability, while Petersen (2011) has offered a separate approach building on the work of Pearl and Bareinboim (2014). Moreover, the debate over whether the quantities in causal models should be identified with particular experiments has implications for whether epidemiologists can model the effects of a variable such as obesity without specifying a particular way of intervening on it. What may not be evident from the discussion so far is that there are (at least) two distinct communities with very closely related frameworks for doing causal inference in epidemiology, where the differences between the groups are exceedingly subtle, but still sufficient for preventing them from adopting a joint approach. Of course, simply diagnosing these debates as tracing back to different views of the relationship between models and experiments is not sufficient for resolving them. But it does help to clarify the source of the debate, and is suggestive of what would be required for each side to address the concerns of the other. Specifically, each side needs to be convinced that the alternative approach can help them address the theoretical or epistemological problems they are trying to address.[4]

The debate just considered involves a fairly high-level discussion between methodologists interested in spelling out the roles played by certain theoretical principles. If even methodologists can have disputes in which it is difficult to diagnose the source of the disagreement, it is unsurprising that this will occur among less methodologically-oriented scientists. Moreover, the example shows that even having the parties to a debate state their assumptions explicitly is not sufficient for pinpointing the source of the dispute. Here we've suggested that an important but unacknowledged source of the disagreement is the different perspectives that theoreticians and experimenters take towards formal models.

[4]Of course, one possible outcome of such discussions is that the frameworks of each party cannot address the concerns of the other party. This would not lead to a resolution of the dispute as much as a dissolution. Each side would use the tools that they need for their purposes, but since they differ in their purposes there would be no genuine disagreement. Given the similarity of the different frameworks in the consistency debate, this would be an unfortunate outcome.

One risk of having two communities that interpret the same models in different ways is that they may sometimes be falsely led to believe that they are agreeing. That is, if both communities accept the same models, but read the models as having different implications, then their apparent agreement regarding the model will mask underlying disagreement about what can be inferred from it. A more general point is that when considering communities with different but overlapping methodologies, determining what the parties *agree* about can be as difficult as determining what they disagree about.

3 Modelers and experimentalists

The consistency debate exemplifies a more general form of dispute that arises between theoretical modelers and experimentalists across many areas of science. Theoretical modelers value models that are precise regarding both the content of the model and the conditions across which the model can be generalized. In contrast, experimentalists are constrained by the types of quantities that they can feasibly measure, and are more hesitant to represent quantities that are difficult to measure. The experimentalists accuse the formal modelers of positing relationships that will only obtain under assumptions that are difficult to establish. The theoreticians deny that they are making any assumptions at all, but are merely being precise about the implications of the given models.[5] If one is not justified in accepting the posited causal relationships, they say, one should not blame the modeling framework, but should instead refuse to accept the posited models.

We suspect that this debate between theoreticians and experimentalists plays out across a wide range of sciences. This debate has been meticulously documented in MacLeod and Nersessian's work on interdisciplinary collaboration. These philosophers spent time in a systems biology lab in which mathematical modelers attempted to collaborate with experimenters, often leading to the frustration of both groups. Reading their quotations from the interviewees, one can almost hear the irritability in the voices of the scientists, who cannot understand why the other side won't just cooperate. MacLeod (2018) diagnoses the situation as follows:

> One problem is that mathematical model-building uses techniques that abstract and simplify biological network information, in order to generate computationally tractable parameter fitting problems. Justifications for these are based on mathematical and statistical principles. Molecular biologists have generally no expertise at all in assessing whether these really can produce adequate representations and are worth investing their time in. (707)
>
> At the same time modelers in the labs we studied underestimated at the start of their collaborations the amount of tacit, technical and biological knowledge that goes into experimentation, and the extent to which such knowledge goes beyond just "recipe following" in experimental procedures. Particularly they lack understanding of the dependence of those procedures on sound skills and control, their limitations, as well as technological and

[5]We also suggest an alternative reading of the theoreticians' position below.

> material constraints of the biological substances that are used... As a result modelers' practices are rarely well coordinated with constraints on experimentation. Modelers may build their models in apprehension of experimental results that are impossible to obtain for technical reasons they do not understand and cannot anticipate. (708)

Here it is salient how the attitudes of the scientists towards the models differ depending on whether they are constrained primarily by mathematical tractability or by experimental feasibility. The issue is not precisely analogous to that in the consistency debate, since Pearl was not concerned with tractability. But it similarly shows how differences can arise between communities that treat modeling as a first step in problem solving and those that see modeling as much more constrained by experimental practice.

We encourage interested readers to look further at Macleod's work, which contains a wealth of examples of the difficulties scientists face in attempting interdisciplinary collaboration. One difficulty in finding the source of scientific disagreements is that the written record of disputes across articles may omit sources of disagreement that only emerge in the face-to-face interactions among the sciences, so ethnographic work is invaluable. A further virtue of MacLeod's research is that he does not talk as if simply highlighting these differences is sufficient for resolving the problems of communication between different disciplinary communities. The problems he considers can only be resolved by the cooperation of different communities, and it is a serious worry that the modeling tools and practices of the communities cannot be straightforwardly combined. It is far from obvious how to address this worry, or even who would be capable of developing a path forward, since none of the disciplines considered contains all the tools necessary for tackling the relevant problems – this is why interdisciplinary collaboration was called for in the first place. In our discussion, we will also highlight the importance of not forgetting that scientists in particular disciplinary communities have been trained to use particular sets of methods to solve particular types of problems, and as a result will be suspicious of tools from different communities that do not help them solve problems in their home domain. This point may initially seem too mundane to be worth stating explicitly. But such pragmatic points are easily forgotten by philosophers as soon as they are invited to theorize about the conceptual disputes "underlying" a debate.

MacLeod's discussion also helps us add some nuance to the debate between modelers and experimentalists sketched at the beginning of the section. Echoing Pearl, we attributed to the theoreticians the position that they were not making assumptions at all in drawing certain implications from their models, but simply demanding that modelers be precise. But in the cases that MacLeod discusses, it does appear that the theoreticians are making assumptions, albeit not assumptions about first-order scientific facts in the domain. Specifically, they are assuming that their mathematical methods produce reliable results despite making certain idealizations, and, perhaps, that the strength of the mathematical methods can compensate for limitations in the data.[6] Naturally, experimentalists unfamiliar with these methods will be less confident in the them and

[6] We are indebted to an anonymous reviewer for highlighting the relevance of this element of MacLeod's discussion.

may be baffled as to how they could work despite such misrepresentations.[7] The use of idealizations can thus serve as a further source of non-factual disagreement. Our next case study makes this point clear.

4 Econophysics and Idealization

We have seen that the use of idealizations can lead to non-factual disagreements when different communities diverge in their comfort with idealizing assumptions. A further context in which idealizations cause difficulties is when different communities disagree about which idealizing assumptions are appropriate to a particular domain. This is especially likely to occur in interdisciplinary projects, as we will illustrate with the example of econophysics.[8] This case study offers an interesting contrast to the above, since it highlights that disputes arise not just between modelers and experimentalists, but also between groups of theorists with different backgrounds.

Econophysics is the term for a variety of interdisciplinary projects that take models from physics (typically statistical physics) and apply them to study economic systems. Thébault et al. (2018) discuss how there has been some debate between econophysicists and mainstream economists about the validity of these sorts of approaches. The economists claim that econophysics models are bad models of economic phenomena, and the econophysicists disagree. But what precisely are they disagreeing about? As Thébault et al argue, the economists are making an assumption that the standard modeling methodology in economics is the only way to model economic systems. The econophysicists – who typically have backgrounds in physics rather than economics – are assuming that a very different methodology is in play. The disagreement between the economists and the econophysicists should not be seen as a disagreement about the validity of the particular model, but as founded in a deep disagreement about which modeling practices are legitimate. It is not a factual proposition that is at stake, but a disagreement about whether it is a legitimate kind of explanation of real world phenomena that can be gleaned from the econophysics models that mainstream economists see as hopelessly unrealistic and flawed.

For example, some econophysics models treat money and exchanges of money in analogy to energy and the kinetic exchange of energy in gases. They then show that such systems yield distributional features we see in real world distributions of wealth. Economists criticize this kind of modeling as being a non-starter, since the exchanges of money are conservative, meaning that the overall amount of money remains constant. This model, according to the economists, is just inconsistent with some fundamental

[7] Feldman (2011) has argued that even in cases where the parties to a dispute do not initially have shared evidence "evidence of evidence is evidence". That is, even if agent A lacks some evidence E that agent B has, A has evidence (from B) that E exists, and should weigh it accordingly. In cases where two disciplinary communities rely on methods that render it mysterious to them how the methods of the other could work, and neither community can provide independent support for their methods, it seems implausible to suggest that members within each community should treat the testimony of the other as evidence for some independent evidence in favor of their position.

[8] Lisciandra and Herfeld (2019) discuss many further examples of interdisciplinary science.

features of economics that change the overall amount of wealth: growth and credit. This disagreement is not about some empirical fact, nor even about some economic stylized fact: it is about what the fundamental governing principles of economic modeling are. Economics and econophysics have different views on what kind of idealizations are legitimate; this seems to stem from the methodological backgrounds of the groups. For an economist, a model of the economy with no growth and no credit is fundamentally flawed, whereas for an econophysicist, treating money as a conserved quantity is a legitimate idealization because of the mathematical tools it allows us to bring to bear on the issue. For an economist, a model of economic agents should treat agents as rational decision makers whose choices are made consciously and on the basis of the best evidence available, whereas for an econophysicist, if the focus is on the aggregate behavior then we can reasonably ignore as much of the fine structure of the individuals as possible. The econophysics model looks implausible to an economist, since the "agents" essentially bounce around and exchange money at random. This does not look very much like an economic process at all. The scale, or the resolution of the model seems to be wrong.

The issue might actually be even deeper than that. Thébault et al (section 5) present various philosophical understandings of models and modeling and argue that some fit better than others with the kind of practice that the econophysicists engage in. Arguably a different philosophical outlook is natural for mainstream economic modeling. If that is right, then this disagreement really only bottoms out at the level of one's overall philosophical view of modeling in the sciences. This is not typically something scientists spend a great deal of time explicitly thinking about, but rather something they absorb through their socialization within a certain community of practitioners. As Thébault et al point out, this is fertile ground for philosophers of science to have an impact within science: we are trained to analyze methodological foundations, and we can help scientists who disagree to make clear the nature of their disagreement. However, as we emphasize in section 6, we should be skeptical of any claim that scientific debates can be traced back to debates between proponents of different philosophical theories.

The question of when idealizations are justified is a subtle one, and we will not attempt to resolve it here. Idealization appears to be an essential part of scientific practice, and the justification of an idealized model need not appeal to the possibility of replacing it with a de-idealized true model. Rather, it is sometimes possible to prove that the idealized model does as well as the (unknown) true model for certain predictive purposes (Fillion and Moir, 2018). But even if some idealized models can ultimately be given a rigorous justification, at the outset such models are adopted for more pragmatic aims such as making a problem more tractable. As a result, in cases where two communities adopting different idealizations interact to address a problem in a new domain, no one may be able to provide a rigorous proof for whether their idealizations are justified. This is not to say that the scientists lack any justification for their practices – one would hope that experienced and methodologically sophisticated modelers will develop a feel for when certain modeling techniques are promising. But such know-how cannot be easily shared across scientists trained in different communities. For this reason, disputes among modelers with different sets of idealizations can be especially difficult to resolve.

So we've seen that modelers and experimentalists can disagree on methodology, and we've seen that modelers can disagree with different modelers on methodology. But it's also true that experimentalists of different kinds can disagree on method as well. See, for example, Kvakkestad et al. (2007); Rocca and Andersen (2017) for a discussion of different experimental disciplines apparently disagreeing about the safety of genetically modified crops, but really disagreeing about details of method.[9] This is also an interesting case because the groups involved – molecular biologists and conventional plant breeders – are from much "closer" disciplines than the other cases we discuss. So these sorts of ultimately non-factual disagreement can be quite fine-grained: they can occur even among nearby disciplines.

5 Dynamical systems theory models in stem cell biology

We turn now to a case study of dynamical systems theory and stem cell biology first discussed by Melinda Bonnie Fagan (Fagan (2016)). In addition to providing a final example of non-factual disagreement, Fagan's discussion also serves as an example of a philosophical intervention in such a disagreement, and thus will set up our later discussion of how philosophers can play a fruitful role in such debates. This case study also emphasizes that such non-factual disagreement occurs across the sciences.

Some kinds of cells – known as "stem cells" – have the ability to turn into several different kinds of other cells. This feature – pluripotency – is of great interest to biologists since the therapeutic potential of such "programmable" cells is huge. Standard stem cell biology research typically involves careful experimental study of stem cells. Among the things biologists are attempting to understand is what genes are involved in causing a cell to develop into a stem cell, and in the possibility of inducing "stemness" through manipulation of a cell's genes. Mapping the "gene regulatory networks" controlling stem cells is a big part of what stem cell biologists do.

Dynamical systems theory is the mathematical analysis of systems and their behavior over time. DST modelers describe their target system using systems of differential equations which describe how the variables of interest change in response to changes in the other variables. Recently, some groups of DST modelers have turned their attention to modeling stem cells. They have, essentially, found a kind of DST model that exhibits the classic characteristics of "stemness"; the systems sit in some kind of equilibrium, but can be made to go towards one of several distinct steady states through some kind of manipulation. The intuition is that the first equilibrium state is the stem cell state, and the other steady states are the kinds of cell that the stem cell can be induced to become.

The DST groups have presented these results as providing an "explanation" of stem cell behavior. Stem cell biologists have, on the whole, ignored their work. Fagan considers why the experimentalists have ignored the work of the DST modelers. It will help to distinguish between two parts of her discussion. First, she imagines some criticisms that the experimentalists might make against the modelers. Plausibly, she suggests that the

[9] The case has some downsides compared to the cases discussed above, since there are also institutional and value-based reasons for disagreement in this case too. But method still plays a role.

experimentalists might doubt whether the highly idealized DST models in fact apply to the systems they are studying: they might doubt that the dynamical models developed by the DST researchers tracks the low-level details of the gene regulatory network, and they might also question the assumption that the simplistic connection between gene expression and phenotype presumed in the DST models. Additionally, DST models can accommodate the known features of stem cells, but they don't produce new information: they haven't taught us anything surprising about stem cells. Finally, the DST models focus exclusively on the properties of gene regulatory networks, while ignoring the cell-level molecular properties that are of interest to experimental biologists.

Second, after presenting these criticisms Fagan claims that the differences between the experimentalists and the DST models result from the former adopting a mechanistic account of explanation, and the latter adopting a "covering law" account. She concludes that in order for the DST theorists to convince the experimenters that their work is valuable, they need to acknowledge the legitimacy of mechanistic explanation and, presumably, incorporate mechanistic details into their accounts.

While we find it plausible that certain aspects of the debate over stem cell methodology can be traced to different views of explanation, Fagan's emphasis on explanation potentially obscures more mundane sources of the disagreement. At least some of the concerns that she imagines the experimenters posing to the theorists reduce to the following: we (the experimenters) do not know why we should think your idealized models even approximately represent the systems we study, and do not know how adopting them will help us address the research questions we want to solve. The inability of one research community to find another research community's work useful would seem to be sufficient for them to ignore it. Unless the DST modelers can explain to the experimentalists how *they* would benefit from the models, they will continue to be neglected, even if the explanatory paradigms of the two disciplines were similar.

In fact, Fagan does suggest a promising way in which mechanistic explanations might be fruitfully supplemented with dynamical ones (901). She presents Bechtel's notion of a "dynamic mechanistic explanation" (Bechtel, 2011), in which one begins with a mechanistic model of a system and then appeals to dynamical models in order capture more complex features of the system that elude the mechanistic model. This proposal is promising precisely because it focuses on a concrete example that the experimentalists would accept and makes precise the role that the dynamic models are supposed to play. But we see Fagan as overly emphasizing the role of philosophical theories of explanation in both understanding and attempting to resolve the debate (or lack thereof). Her explanation of why mainstream stem cell biologists have been ignoring the DST theorists strikes us as eminently plausible. To the extent that we have any disagreement with Fagan, it is only regarding her claim that we can treat the communities' different views of explanation as a *source* of the debate. This does not strike us as plausible, and in any event, the project of convincing the experimentalists that the DST models will matter for their work seems largely independent of determining whether the DST theorists rely on a legitimate alternative account of explanation.

One takeaway from this case study is that "disagreement" doesn't necessarily mean an actual dispute among researchers. In this case, the disagreement is really a case of one group apparently just ignoring the other. Why don't the two sides in the consistency debate or the econophysics debate just ignore one another? Why did those two debates generate a fractious and disputatious literature, while the DST/stem cell case didn't? In the econophysics case, it seems clear that the reason for the vocal dispute is because each side is trying to lay claim to the same subject matter: modeling economic systems. And this dispute over territory is obviously going to cause tensions.[10] In the DST case, it seems like the stem cell biologists simply don't see the DST modelers as doing something that "steps on their toes" in any way, and so have little reason to resent or object to their work.

6 Philosophers of Science and Scientific Disagreement

In the previous sections, we have considered a range of cases in which scientists are involved in a dispute, but where the dispute does not concern the first-order facts from the relevant domain, and where discovering the source of disagreement takes some work. Broadly, the disputes considered result from differences in background methodological assumptions. Resolving these disputes is not merely a matter of one party convincing the other of the truth of a particular set of propositions, but rather of coming to a consensus about which modeling frameworks are best suited for making progress in a particular domain. Yet, it may not initially be clear that the disputes concern methodology rather than first-order facts. With respect to the consistency debate, one could easily miss the subtleties concerning the question of whether consistency counts as theorem, axiom, or assumption, and might instead more directly interpret it as a debate over whether consistency is *true*. In the econophysics case, skeptics of econophysics are more likely to frame their opposition in terms of its simply being *false* that money is a conserved quantity, rather than granting that it is an idealization and asking when such idealizations are justified. So the appearance that the participants of a debate are disagreeing about a particular proposition may be misleading, and this has implications both for what is involved in diagnosing the sources of the dispute and for suggesting possible resolutions.

In cases where debates non-transparently depend on factors other than the disputants' beliefs regarding first-order facts, case studies are indispensable. The participants' presentations of their debate may not be sufficient for diagnosing the sources of the dispute, which may only be evident when considering broader methodological and disciplinary differences. But this is not to say that making a correct diagnosis is merely a matter of doing more ethnographic work. In analyzing the cases above, we have seen that many of the disagreements in fact concern questions about the legitimacy of modeling idealizations and of the virtues of particular axiomatizations over others. These are precisely the types of issues that philosophers of science have been trained to think about, and where they can make a distinctive contribution in seeking to resolve or make headway in the relevant debates.

That said, in order for philosophers to play a fruitful role they need to have a clear picture of the genuine sources of the relevant debate, and here we see philosophers as

[10] See Mäki and Marchionni (2009) for another case of "scientific imperialism" and resentment.

having certain tendencies that, left unchecked, can render their advice counterproductive. Like all the other disciplines considered here, philosophers have a toolkit that they are primed to use whenever possible, even in situations where the tools are not the most suitable. In particular, philosophers get especially excited when they can represent a debate as resulting from a clash of theoretical frameworks, often at the expense of focusing on more pragmatic sources of the debate.[11]

In the examples we considered in this paper, we can distinguish between 1) The reason why the parties are fighting (or ignoring each other) in the first place, and 2) the methodological reasons that the parties give[12] for their disagreement, which often have a philosophical flavor to them. Examples of the latter are Pearl's claim that his approach better approximates the Pythagorean ideal, and DST theorists' claims that their opponents are ignoring them because they have too narrow a view of explanation. Regarding the reasons why the parties are actually fighting, it would be too hasty to attribute these to pragmatic issues. But it is safe to say that most scientific disputes do not arise due to the scientists' commitments to certain philosophical positions about explanation, abstraction, or metaphysics. Nevertheless, in disputes with their opponents, scientists will sometimes raise or presuppose methodological principles with philosophical import. Here is where philosophers have a seemingly helpful role to play in making headway on the debate, but also where some caution is advised.

Suppose you are a philosopher studying explanation and you are asked to comment on a debate in which both sides make contentious claims about explanation. You read through the relevant literature, and discover that the two sides can be divided up so that their commitments more or less line up with contrasting positions in the explanation literature. Of course, the scientists, not being philosophers, will be imprecise in how they frame their explanatory commitments and how they contrast their positions with one another. So it looks like a perfect opportunity to bring your expertise in order to help clarify the debate. Herein lies a temptation that should typically be resisted. The temptation is to think that by clarifying the explanatory claims, one can help make headway in the relevant debate. The reason that the temptation should be resisted is that to the extent that the explanatory claims are not causes of the dispute, but rather effects of it, helping the parties to reinforce their theories of explanation will amount to helping them shore up the philosophical fortresses from which they are defending their positions, rather than seeking ways to resolve the debate. If, as we have suggested, certain scientific debates can be traced to methodological differences regarding which approaches are successful, then the role of philosophers should be to identify these differences so that they might be better addressed.

We want to be careful not to oversell the point. Of course, there have been scientific debates in which explanatory considerations have been central, and even in cases where they have not been they still may have played a role. Our point is that in order to get

[11] This practice of tying one's discussion of a case study in science back to some piece of philosophical grand theory seems to be an almost necessary part of getting one's paper published in a philosophy journal.

[12] Or perhaps it would be better to say "the methodological reasons that the parties *would* give if they thought about it".

a clear picture of the debate, one needs to be cautious to avoid the pitfall of placing too much credence in scientists' descriptions of the philosophical ideas that are allegedly at issue. Naturally, this reflects a more general point about learning about scientific practice from observing it, rather than blindly trusting scientists' descriptions of what they do. But scientific disagreements introduce new issues. As we have argued, it is not always evident what, precisely, scientists are arguing about, and the fact that they sometimes treat the differences as philosophical provides an especially tempting invitation for philosophers to intervene. We need to be aware of this temptation, and to keep it in check, if we are to accurately diagnose the bases for non-factual disputes in science.

7 Conclusions

In each of the cases we have discussed, scientists from different communities are disagreeing with each other, and the disagreements do not, at bottom, concern first-order scientific facts. These disagreements often result from broader methodological or foundational differences between the communities. Because these methodological principles are often tacit, unspoken aspects of the scientists' training, it is difficult for the practitioners to recognize the source of their disagreement with their colleagues.

Philosophers of science have the training to recognize the foundation of these sorts of scientific disagreement and have a role to play in resolving them. But they must be careful that they act as mediators – helping to facilitate dialogue – rather than giving scientists tools to reinforce their methodological silos.

References

Andersen, F., Anjum, R. L., and Rocca, E. (2019). Philosophical bias is the one bias that science cannot avoid. *eLife*.

Bechtel, W. (2011). Mechanism and biological explanation. *Philosophy of science*, 78(4):533–557.

Cartwright, N. (2012). Presidential address: Will this policy work for you? predicting effectiveness better: How philosophy helps. *Philosophy of Science*, 79:973–989.

Cole, S. R. and Frangakis, C. E. (2009). The consistency statement in causal inference: a definition or an assumption? *Epidemiology*, 20(1):3–5.

Douglas, H. E. (2009). *Science, Policy, and the Value-Free Ideal.* University of Pittsburgh Press.

Fagan, M. B. (2016). Stem cells and systems models: clashing views of explanation. *Synthese*, 193(3):873–907.

Feldman, R. (2011). Reasonable religious disagreements. *Social Epistemology: Essential Readings*, 137.

Fillion, N. and Moir, R. H. (2018). Explanation and abstraction from a backward-error analytic perspective. *European Journal for Philosophy of Science*, 8(3):735–759.

Gibbard, A. and Harper, W. L. (1978). Counterfactuals and two kinds of expected utility. In *Ifs*, pages 153–190. Springer.

Hernan, M. A. and VanderWeele, T. J. (2011). Compound treatments and transportability of causal inference. *Epidemiology (Cambridge, Mass.)*, 22(3):368.

Kelly, T. (2005). The epistemic significance of disagreement. *Oxford Studies in Epistemology*, 1:167–196.

Kuhn, T. S. (1962). *The structure of scientific revolutions*. University of Chicago Press.

Kvakkestad, V., Gillund, F., Kjolberg, K., and Vatn, A. (2007). Scientists' perspectives on the deliberate release of GM crops. *Environmental Values*, 16:79–104.

Laplane, L., Manovani, P., Adolphs, R., Chang, H., Mantovani, A., McFall-Ngai, M., Rovelli, C., Sober, E., and Pradeau, T. (2019). Why science needs philosophy. *PNAS*, 116(10):3948–3952.

Lisciandra, C. and Herfeld, C. (2019). Knowledge transfer and its contexts. *Studies in History and Philosophy of Science Part A*.

MacLeod, M. (2018). What makes interdisciplinarity difficult? some consequences of domain specificity in interdisciplinary practice. *Synthese*, 195(2):697–720.

Mäki, U. and Marchionni, C. (2009). On the structure of explanatory unification: the case of geographical economics. *Studies in History and Philosophy of Science Part A*, 40(2):185–195.

Pearl, J. (2009). *Causality*. Cambridge university press.

Pearl, J. (2010). On the consistency rule in causal inference:" axiom, definition, assumption, or theorem?". *Epidemiology*, pages 872–875.

Pearl, J. and Bareinboim, E. (2014). External validity: From do-calculus to transportability across populations. *Statistical Science*, pages 579–595.

Petersen, M. L. (2011). Compound treatments, transportability, and the structural causal model: the power and simplicity of causal graphs. *Epidemiology*, 22(3):378–381.

Rocca, E. and Andersen, F. (2017). How biological background assumptions influence scientific risk evaluation of stacked genetically modified plants: an analysis of research hypotheses and argumentations. *Life Sciences, Society and Policy*, 13(11).

Rowbottom, D. P. (2016). What is (dis)agreement? *Philosophy and Phenomenological Research*, 97(1):223–236.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66(5):688.

Spirtes, P., Glymour, C. N., Scheines, R., Heckerman, D., Meek, C., Cooper, G., and Richardson, T. (2000). *Causation, prediction, and search*. MIT press.

Spirtes, P. and Scheines, R. (2004). Causal inference of ambiguous manipulations. *Philosophy of Science*, 71(5):833–845.

Thébault, K., Bradley, S., and Reutlinger, A. (2018). Modelling inequality. *British Journal for the Philosophy of Science*, 69(3):691–718.

VanderWeele, T. J. (2009). Concerning the consistency assumption in causal inference. *Epidemiology*, 20(6):880–883.

Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford university press.