

Mechanistic Explanation: Integrating the Ontic and Epistemic

Phyllis Illari

Received: 10 July 2013/Accepted: 10 July 2013/Published online: 24 July 2013
© Springer Science+Business Media Dordrecht 2013

Abstract Craver claims that mechanistic explanation is ontic, while Bechtel claims that it is epistemic. While this distinction between ontic and epistemic explanation originates with Salmon, the ideas have changed in the modern debate on mechanistic explanation, where the frame of the debate is changing. I will explore what Bechtel and Craver's claims mean, and argue that good mechanistic explanations must satisfy both ontic and epistemic normative constraints on what is a good explanation. I will argue for ontic constraints by drawing on Craver's work in Sect. 2.1, and argue for epistemic constraints by drawing on Bechtel's work in Sect. 2.2. Along the way, I will argue that Bechtel and Craver actually agree with this claim. I argue that we should not take either kind of constraints to be fundamental, in Sect. 3, and close in Sect. 4 by considering what remains at stake in making a distinction between ontic and epistemic constraints on mechanistic explanation. I suggest that we should not concentrate on either kind of constraint, to the neglect of the other, arguing for the importance of seeing the relationship as one of integration.

1 The Ontic–Epistemic Distinction for Mechanistic Explanation

The original distinction between ontic and epistemic explanation is due to Salmon and situated against the background of Hempel's classic account of explanation, and related views. So, for example, Salmon writes: 'In its classic form—the inferential version—the epistemic conception takes scientific explanations to be arguments.' He contrasts this with his own conception of explanation: 'The ontic conception sees explanations as exhibitions of the ways in which what is to be explained fits into natural patterns or regularities ... [and] usually takes the patterns and regularities to be causal' (Salmon 1984, p. 293). There are more versions of the

P. Illari (✉)
University College London, London, UK
e-mail: phyllis.illari@ucl.ac.uk

epistemic conception, and further nuances of Salmon's views worthy of study, but my aim here is to examine the debate currently under way in the mechanisms literature.

Carl Craver and William Bechtel have taken inspiration from Salmon's distinction, and adopted its language, to disagree about the nature of mechanistic explanation: Bechtel claims that mechanistic explanations are epistemic, while Craver claims they are ontic. And this dispute is important to the mechanisms literature widely, because most new mechanistas agree with Craver. As Wright points out: 'Perhaps because of their common interests in causality, most New Mechanists have hitched their wagon to Wesley Salmon's ontic conception of scientific explanation' (Wright 2012, p. 376). Certainly Machamer et al. (2000) gives an ontic conception, Machamer (2004) seems to maintain it, and Glennan (2002, 2005) explicitly agrees. So if any of Bechtel's criticisms are right, that is of wide importance to the mechanisms literature. Since only Bechtel and Craver of the major mechanistas defend their view in any extended way, I will focus on their work. Further, their work needs examination as what their specific claims are takes some work to understand, as they do not simply adopt a pre-existing clear distinction.

I begin with Craver. There are two different things that Craver classifies as ontic explanation, which may both derive from Salmon. First, Craver holds that mechanistic explanation involves fitting a phenomenon into the causal structure of the world: 'I argue that good explanations in neuroscience show how phenomena are situated within the causal structure of the world (Salmon 1984; Craver 2007, p. 21)'. Here, explanation involves *showing or exhibiting* something about the causal structure of the world. Elsewhere, though, Craver says: 'Other times, the term explanation refers to an objective portion of the causal structure of the world, to the set of factors that bring about or sustain a phenomenon (call them objective explanations). . . Objective explanations are not texts; they are full-bodied things. They are facts, not representations' (Craver 2007, p. 27). Here, any reference to exhibition or showing is dropped. In this case the *mechanism itself* explains.

Bechtel holds that explanation is deeply concerned with understanding, and is essentially a human activity: 'Explanation is fundamentally an epistemic activity performed by scientists' (Bechtel 2008, p. 18). Given Bechtel's concern with understanding and the cognitive abilities of human beings in his overarching project of understanding mechanistic explanation, it is easy to assimilate his view to a psychologism where an explanation is anything that generates an entirely subjective 'aha' feeling on the part of the receiver of the explanation. This is a view that may well be rejected on the grounds that science is not in the business of making people merely *feel* as though they have understood the world. However, as Waskan (2011) argues, there is also a 'success' interpretation of 'understanding', which requires success in understanding the real world. As Bechtel firmly classifies his own view as epistemic, it seems likely that this is the sense he intends, which is consistent with the idea that explanations generate knowledge. So Bechtel seems to agree with his close collaborator: 'It is surely right to say that mechanistic explanatory texts aim to increase knowledge about mechanisms. . . Obviously, knowledge of how things work is an epistemic matter if anything is, which is just to say that analysis of

mechanistic explanatory texts properly requires a broadly epistemic conception of mechanistic explanation' (Wright 2012, p. 382). So, for Bechtel, and Wright, mechanistic explanations are texts, or descriptions and so on, that aim to increase knowledge about mechanisms. For the epistemic conception, the *text or description* explains. There are further complexities of Bechtel's real concern for the needs of actual cognizers in mechanistic explanation, which I will return to later, applying Waskan's arguments in favour of some aspects of psychologism being of real use to epistemic explanation.

Craver and Bechtel agree about a great deal regarding mechanisms and mechanistic explanation, which is worth spelling out, to help pinpoint their disagreement over ontic versus epistemic explanation. They have different accounts of mechanisms, so in this paper I will follow the view I have argued for elsewhere: 'A mechanism for a phenomenon consists of entities and activities organized in such a way that they are responsible for the phenomenon' (Illari and Williamson 2012, p. 120). In a very general way, finding mechanistic explanations involves finding and describing the phenomenon, and finding and describing the entities and activities, and their organization, by which the phenomenon is produced. We argue that this account captures a core consensus on what a mechanism is that the major mechanistas broadly agree on, while also offering understanding of what does not count as a mechanism, so the debate between Bechtel and Craver should not be affected by adopting this background view.

With this background in mind, Craver's ontic conception can be further spelled out: the relevant domain of ontic explanation is the real world, which has genuinely causal structure, some at least of which are entities and activities organized to form mechanisms, and at least some of those mechanisms are mechanistic explanations—mechanisms explain the phenomena they are responsible for.¹ While Bechtel is cautious about making explicitly realist claims about the causal structure of the world, he does not actually disagree with any of this except the final claim. Neither Bechtel nor Craver is an instrumentalist about science, so their disagreement is not a variation on the realist-instrumentalist debate. Further, Bechtel and Craver, along with all the mechanistas, take the thing explained to be a phenomenon, and do not seek answers to why-questions. Bechtel does, however, deny that mechanisms themselves are ever mechanistic explanations. Craver also seems to classify any *exhibition* of causal mechanisms as an ontic explanation, while Bechtel might well classify such explanations as epistemic. In such a case, though, Bechtel would still focus attention on the description or exhibition itself in a way that Craver would not. In so far as Craver still prioritises the mechanism itself in these cases, while Bechtel prioritises the description or exhibition, they still disagree. So I will take exhibitions of mechanisms which prioritise the mechanism itself as ontic explanations, and take such exhibitions which prioritise the description or exhibition as epistemic explanations. I will return to this point later, particularly in Sect. 3.

To motivate his view, Craver points out that we think that there are phenomena that we know of, and we think they have an explanation, although we don't yet know it—indeed we may never know it. So undescribed mechanisms can count as

¹ I thank an anonymous referee for suggesting a useful clarification along these lines.

explainers (Craver 2012, p. 7). He objects to explanations that include non-ontic entities, such as diagrams and equations, presumably on the grounds that equations and diagrams, as abstracta, cannot produce any worldly phenomenon. This is a bone of contention with Bechtel, who increasingly thinks that equations and other aspects of modelling mechanisms are essential to many cases of what he now calls ‘dynamical mechanistic explanation’ (Bechtel 2008). In turn, one reason Bechtel offers in favour of his view is that a phenomenon might remain when the mechanism that produced it has gone. This is not uncommon. For example a protein is produced from DNA, via mRNA, but the mRNA is usually broken down immediately after use, while the protein remains. Bechtel complains that a now-absent cause, such as my example of the mRNA, cannot currently explain in the ontic sense (Bechtel 2006, p. 34). Further, Bechtel says, mechanisms themselves, such as the protein synthesis mechanisms, were around long before we had any scientific explanations.

Bechtel and Craver disagree, and their disagreement impacts on the whole mechanisms literature. However, it is difficult to evaluate their dispute properly, as it is difficult to identify their core motivations. Further, in many places the arguments offered that do express their core motivations are question-begging—at least in debate with each other. For example, Craver (2007) builds an account of mechanistic explanation on criteria for good explanation that almost all look causal. Bechtel repeats in many places that explanation is an essentially human activity. While these claims probably do reflect their core commitments, Craver cannot build his view on the assumption that at least some explanations are causes, which seems to be a claim that Bechtel denies; and Bechtel cannot build his view on the assumption that all explanation essentially involves human cognitive activity, which seems to be a claim that Craver denies—at least, they cannot depend on these assumptions when engaged in debate with each other. Neither approach will move the debate forward, forcing a search for an alternative approach.

Further complicating any attempt to move this debate forward is the fact that the ontic-epistemic debate regarding mechanisms has changed from the original ontic-epistemic debate in what is at stake, its motivation, and its method, and it has not yet settled into a new frame. What is at stake has narrowed, since there is so much agreement in the mechanisms literature over the nature of mechanisms, as presented briefly above. Its motivation is now solely to do with understanding mechanistic explanation, and lacks once-popular targets such as Hempel’s view of explanation. Finally, while some arguments are still offered in the style of the classic debate—linguistic analysis of explanation (Wright 2012), or using cases of paradigm explanations to assess ascription conditions for ‘explanation’ (Wright 2012; Craver 2012)—a new line of argument is coming to prominence.

It will be a major part of the work of this paper to understand the new frame, and urge that the debate regarding mechanisms should settle more firmly into that new frame, where consensus can be generated by pursuing the new line of argument. Craver is explicit that he is arguing for normative constraints on explanation: ‘The second goal is explanatory normativity. The theory should illuminate the criteria that distinguish good explanations from bad’ (Craver 2012, p. 1). So in this frame Craver is not arguing about what an explanation itself *is*, but arguing for the

importance of ontic constraints in recognising, finding, and possibly even using good explanations. This is a different project, with different things at stake. Argument over constraints concerns more the function of good explanations in our essential scientific practices, and in designing those practices to do better, rather than a backwards-looking argument concerning common usage of the word ‘explanation’. Bechtel, and Bechtel and Wright, are not so explicit, but they recognise and are concerned about such norms: ‘explaining refers to a ratiocinative practice governed by certain norms’ (Wright and Bechtel 2007, p. 51), although they also sometimes appear to be arguing about what mechanistic explanations themselves are.² It is true that neither side consistently sticks to arguing over constraints, sometimes changing tack to argue in a more traditional way over what explanations themselves are. However, they are moving towards a focus on constraints, away from the traditional ontic-epistemic debate.

I will frame the debate in these terms of normative constraints on explanation, so I will be concerned with standards that we hold scientific explanations to, to make them good explanations, rather than norms or any other criteria to make something an explanation at all. It might be thought that this is merely to change the subject. On the contrary, this seems to me a positive move, towards a more useful philosophical project, away from unhelpful polarising pressures. Further, Craver holds that the ontic-epistemic debate has always been a normative debate, about regulative ideals for explanation (personal communication). So I will not be primarily concerned with what explanations themselves are, nor about fighting over whether paradigm cases of explanation can be captured by each of the ontic or epistemic conceptions. I will comment further on this during the paper, and particularly at the end of Sect. 3.

Within this new frame, I will argue that good mechanistic explanations must satisfy both ontic and epistemic constraints. I will argue for ontic constraints by drawing on Craver’s work in Sect. 2.1. I will argue for epistemic constraints by drawing on Bechtel’s work in Sect. 2.2. Along the way, I will argue that Bechtel and Craver actually agree on this. In brief, this is because while Craver argues for ontic constraints, he does not offer arguments against epistemic constraints, and some of his work seems to commit him to accepting them alongside ontic constraints. This argument is likely to generalise to the work of other mechanistas such as Machamer and Glennan. In parallel, while Bechtel argues for epistemic constraints, he does not argue against ontic constraints, and seems similarly committed to accepting them alongside epistemic constraints. In Sect. 3, I examine an additional argument of Craver’s, considering whether, even if both kinds of constraints are admitted, an argument can be made that one or the other is prior. I argue that such an argument cannot succeed without making question-begging assumptions about the primary nature or purpose of mechanistic explanation. The sensible conclusion is still to accept both kinds of constraints. I close in Sect. 4 by considering what remains at

² In the pre-publication manuscript, they write: ‘Mechanistic explanation is an epistemic practice. There are norms governing such a practice—namely, that explaining a target phenomenon requires an understanding of the systemic activities that locally produce it, which in turn requires revealing the mechanism’s internal structure, function, and organization’ (Wright and Bechtel 2007, p. 18). But this does not appear in the shorter published version.

stake in making a distinction between ontic and epistemic constraints on mechanistic explanation. I offer reasons against collapsing the distinction, instead arguing for seeing the relationship as one of integration.

2 For Both Kinds of Constraints

In this section, I will argue for two claims. First, we should accept both ontic and epistemic constraints on mechanistic explanation. Second, Bechtel and Craver agree on this, as their work commits both of them to accepting both sets of normative constraints. I will marshall arguments given by Bechtel and Wright, and Craver, in support of my view. I use only some of their arguments here, addressing further arguments of both in Sect. 3.

2.1 For Ontic Constraints: Craver

I will begin with the first sense of ontic explanation, where the mechanism itself explains by being responsible—causally responsible—for the phenomenon. Craver contends that this is a sense of explanation, specifically the sense in which we say that some known phenomena have unknown explanations. Here, ‘explanation’ is being used very much synonymously with ‘cause’: known phenomena have unknown causes. It seems to me that defenders of the ontic conception of explanation are free to retreat here, plant their flags, and forever refuse to budge. There is no way for an opponent to gain leverage on the view, as if Craver—and Machamer, Glennan and others—wish to use ‘explanation’ in that sense, then however many times Bechtel, and Wright, deny that it is a real sense of explanation, there is nothing to compel agreement.

Wright (2012) makes the best attempt to date to deny that this is a genuine sense of explanation. He does not present it as such, but what he does is mount an extended attack on the claim that ‘explanation’ is ambiguous that he identifies as the origin of Salmon’s arguments. This is an interesting approach, but it does not succeed. First, Wright seems at various points to make question-begging claims: ‘Biochemical pathways, changes in Ca_2^+ concentrations, phasic DA1 bursting or allostatic mesocorticolimbic dysfunction, oxygen generators, etc. are simply inapposite candidates for doing any explaining’ (Wright 2012, p. 388), which might suggest that he finds no ambiguity by refusing to countenance a purely ontic interpretation. Wright raises this concern himself and offers a response: ‘This inference might be objected to on the grounds that . . . [the arguments only follow] if we decide in advance that explanation is not synonymous with cause or causal mechanism or the like; and yet, this decision is precisely what is up for grabs. In response, this objection is also baseless. Explanation or explain may be closely related in semantic space to a term like cause or causation or mechanism, but they are neither synonymous nor substitutable salva veritate’ (Wright 2012, p. 391). However, a vaguer, more ambiguous term doesn’t have to be synonymous or substitutable salva veritate with a more specific term to have that more specific term as *one* legitimate interpretation. Second, and more generally, unrelated to the

specifics of Wright's arguments, his approach is hostage to what might be called definitional or usage stubbornness. If some philosophers choose to use or define 'explanation' in such a way, particularly a prominent group rather than a single individual, it is difficult to see how they could be forced to desist!

This is one reason why the move away from arguing about what explanations themselves are to considering normative constraints on explanations is an interesting one, as it is a way to move the debate forward. In so far as Craver is interested in normative constraints on explanation, presumably he is not interested in constraints on mechanisms—the mechanisms themselves simply *are*—he is presumably interested in how ontic features—features of mechanisms—constrain explanatory *descriptions* of mechanisms. To offer reasons that may convince ontic mechanistas to move their flags, or just leave them safely planted and go and look elsewhere; there is already plenty of work on what causes are, and on what mechanisms are, and so it is interesting to seek a distinctive task of an account of mechanistic *explanation*.

So we move to Craver's second sense of ontic mechanistic explanation, that ontic explanations show or exhibit how the phenomenon fits in the causal structure of the world, by showing or exhibiting the mechanism responsible for the phenomenon. As I have said, Craver still disagrees with Bechtel here as he prioritises the ontic constraints over the exhibition. For Craver, explanation cannot be solely to do with human representational processes, but must be evaluated with respect to mechanisms. Craver writes: 'The point of this section is that ontic explanations, the causes and mechanisms in the world, make an essential contribution to the criteria for evaluating explanatory communications, texts (models), and representations. Good mechanistic explanatory models are good in part because they correctly represent objective explanations' (Craver 2012, p. 11). So in Craver's most recent work, his view is that ontic constraints are normatively vital for evaluating mechanistic explanatory texts—identifying good ones. That is what has come to be most important in his view. Note that methods long-used in the debate over ontic-epistemic explanation such as linguistic analysis or application criteria decided by looking at cases of paradigm explanations are no longer of much interest. What is of interest is argument based on the important function of ontic normative constraints in building explanations.

Craver is right that this is an essential part of distinctively mechanistic explanation, and he offers novel arguments for this view. He says that ontic constraints help us to make several important distinctions between different kinds of explanations: 'In Sect. 2, I illustrate how appeal to ontic explanations is essential for marking several crucial normative dimensions by which scientific explanations are and ought to be evaluated: the distinction between how-possibly and how-actually-enough explanations, the distinction between phenomenal descriptions and explanations, the difference between predictive and explanatory models, and the requirement that explanatory models should include all and only information that is explanatorily relevant to the phenomenon one seeks to explain' (Craver 2012, p. 11). It is certainly true that we distinguish how-possibly explanations, which merely tell you how a phenomenon could be produced, from how-actually explanations, which tell us how a phenomenon is produced, by investigating whether how-actually explanations manage to describe

the actual worldly processes. One of the things that sets mechanistic explanation apart from, for example, laws-based explanation, is the identification of parts of the mechanism by which the phenomenon is produced. And, so that this is not mere story-telling, rigorous empirical constraints on what we admit as entities and activities and their organization is required.

Bechtel may emphasise the *description* of the mechanism, but he does not deny these claims. In some ways, he is more extensively committed to them than Craver is. Wright and Bechtel (2007) write: ‘Localization refers to mapping the component operations onto component parts. Decomposition refers to taking apart or disintegrating the mechanism into either component parts (structural decomposition) or component operations (functional decomposition)’ (p. 62–63). Further, Bechtel and Richardson (2010) is an extended exploration and defence of decomposition and localization as heuristics of mechanism discovery, where it is clear that the point of integrating functional and structural decomposition is to allow rigorous empirical constraints on the parts of mechanisms. Bechtel is perfectly clear that a mere functional decomposition is not a mechanistic explanation. A functional decomposition merely identifies possible operations (Craver’s activities). To get a mechanistic explanation requires mapping these operations to working parts (Craver’s entities) to have a mechanistic explanation. The mapping yields evidence that the explanation is no longer a how-possibly story, that the correct entities and activities have been found: ‘One of the strengths of decomposition and localization as a scientific strategy is that it facilitates an increasingly realistic representation of the explanatory domain, even when the initial representation is seriously distorted: failures of localization can be as revealing as successes’ (Bechtel and Richardson 2010, p. 8). For Bechtel, of course, would say that we appeal to real-world mechanisms, not that we appeal to ontic *explanations*, but this seems now to be disagreement merely about a particular use of a word.

Perhaps due to this implicit commitment, Bechtel does not argue against ontic constraints being relevant alongside epistemic constraints. Wright and Bechtel (2007) write: ‘Do the component parts and their operations and organization figure in our understanding of how and why depolarization occurs? Well, yes, in a flat-footed sense: without any of these things to implicate, mechanistic explanations would be without content’ (p. 50). They continue later: ‘In the context of mechanistic explanation, the model’s elements correspond to the parts of a mechanism, and their structure conforms, not to a theory, but rather to the mechanism’s constituency and interactivity’ (p. 53). Bechtel is interested in mental models, but he is concerned with our knowledge of mechanisms, not with *mere* mental models.

Craver is explicit about ontic constraints being causal constraints, while of course Bechtel, and Wright and Bechtel, are not, writing only of the components and organization of mechanisms. However, such parts and their organization are clearly real existents, and so ontic, even if they do not count them as ontic explanations. Further, it is clear that Craver counts reference to the components and organization of mechanisms as fitting phenomena into the causal structure of the world. There is little if any substantive disagreement.

So Craver's positive argument for ontic constraints on mechanistic explanation is absolutely right. But Bechtel's extensive discussion of decomposition and localization shows that he also recognises ontic constraints. He just doesn't want to call those constraints ontic explanations. Further, Craver doesn't argue directly against epistemic constraints as well as in favour of ontic constraints, although I will return to two further arguments of his in Sect. 3.

2.2 For Epistemic Constraints: Bechtel

Bechtel's core idea is that mechanistic explanation essentially involves conveying understanding of how the entities, activities and their organization produce the phenomenon. Mechanistic explanation is a cooperative enterprise worked on by many scientists, and so mechanistic explanation cannot ignore human representational processes. A mechanism producing a phenomenon certainly causes that phenomenon independently of us, but it is of no use to us whatsoever until we grasp that mechanism, describe it, and understand it. As I have said, both Bechtel and Wright use 'understand' on a success interpretation, where understanding is not purely psychologistic but requires better knowledge of the world, so that this is an epistemic conception. However, Bechtel and Wright are also concerned with how actual cognizers manage to achieve such success: 'After all, explaining refers to a ratiocinative practice governed by certain norms that cognizers engage in to make the world more intelligible; the non-cognizant world does not itself so engage' (Wright and Bechtel 2007, p. 51). As a result, Bechtel, along with Wright, is interested in exploring a space of reasons vastly more complex than traditional approaches to characterizing explanation in terms of inference or rational expectability, and particularly wishes to face up to the increasing importance of complex mathematical modelling in mechanistic explanation.

Craver is particularly suspicious of accepting equations, and other more complex mathematical creations like models or simulations, as part of mechanistic explanations. This is presumably because Craver cannot see where the ontic counterpart of such creations are, that act as legitimate constraints on mechanistic explanations. However, Bechtel is always interested in integrating equations, models and simulations with an account of genuinely mechanistic explanation. So, for Bechtel, acknowledging the importance of mathematical entities in mechanistic explanation goes beyond traditional deductive-nomological explanation, towards dynamical mechanistic explanation: 'First, the equations are advanced not as general laws, but as descriptions of the operations of specific parts of a mechanism. Second, the purpose of a computational simulation (like mental simulation in the basic mechanistic account) is not to derive the phenomenon being explained but to determine whether the proposed mechanism would exhibit the phenomenon. Finally, an important part of evaluating the adequacy of a computational model is that the parts and operations it describes are those that can be discovered through traditional techniques for decomposing mechanisms' (Bechtel 2011, p. 553). Note that this last is clearly phrased as a normative constraint. For Bechtel, the purpose of equations is to describe a real mechanism, and so the mechanism is their ontic counterpart. Further, he is trying to explain that he is not attempting to replace

Hempel's predictability requirement by deriving the phenomenon using simulations. Instead, he is trying to capture the way that simulations are used to find out whether the ontic system could produce the phenomenon to be explained. This, then, is a way of testing whether you have described the ontic system accurately, and Bechtel will only accept parts and operations of a computational model as describing a mechanism if those parts can also be discovered through more traditional techniques. So Bechtel's recent interest in the use of mathematical creations is entirely consistent with his earlier focus on decomposition and localization. Mathematically complex strategies are just ways of describing—and even discovering—more complex entities, activities, and forms of organization of mechanisms than can be described or discovered without using maths. Bechtel says this himself: ‘The need for simulation with mathematical models shows how adequately recomposing the sorts of mechanisms commonly encountered in biology and understanding their dynamics require tools beyond the qualitative one contemplated in the basic mechanistic account. Accordingly, Bechtel and Abrahamsen (2010) characterize such explanations as *dynamic mechanistic explanations*’ (Bechtel 2011, p. 553–554).

It is complexity of organization of real-world systems that demands such mathematical techniques: ‘When the organization being investigated remains relatively simple, it is possible to construct relatively simple explanatory models through mental simulation of the activity in the mechanism step by step. As more organizational complexity and co-operations are discovered, however, this becomes more difficult. With complex feedback loops, the mechanism can begin to behave in unexpected ways. To understand such behavior, researchers often need to offload some of the cognitive labor involved in constructing explanatory models of parallel complexity onto their (research) environment—e.g., by supplementing their own ability to mentally trace activity in a system with computer-based simulations. By supplying models of the various components and the manner in which they interact with each other, researchers can discover the consequences of organization’ (Wright and Bechtel 2007, p. 64–65). In so far as such techniques are aimed at uncovering real parts and complex forms of organization of real mechanisms, in a form in which cognizers can use it to come to have knowledge of these mechanisms, these concerns still fall within the epistemic conception of explanation. It is true that Bechtel is concerned with the actual practices likely to generate successful understanding—successful mechanistic knowledge—some of which are hostage to features of human psychology. In this, he seems to agree with Waskan (2011), who argues that while pure subjective psychologism about explanation is untenable, some psychological features are of interest to our knowledge, in a way that goes beyond mere descriptive accuracy. So, he thinks it is important to mechanistic knowledge that we understand *how possibly* the phenomenon came about. This requires some kind of mental grasp: ‘there is more to it than intelligibility than the mere feeling that sometimes tracks it. While the feeling may be had without explanation and vice versa, it is exceedingly difficult to find, or even imagine, a case where one genuinely has an explanation for a happening and yet does not understand how or why it might have occurred, and it is equally difficult to find or imagine cases where someone genuinely understands how or why possibly and yet

lacks an explanation. ... The study of the psychological underpinnings for intelligibility has, I believe, potential to tell us much about what explanations are, ... and about the manner in which we evaluate them' (Waskan 2011, p. 12). If Waskan is right about this, then some psychological features are of interest to the epistemic conception of explanation.

Craver initially seems to be very dismissive of similar concerns of Bechtel's: 'Such topics are the proper province of psychologicistic theorizing about scientific explanation. But that is a separate topic from the central topic of explanation that has occupied philosophers for over 50 years' (Craver 2012, p. 22). This is not obviously true. If, as Craver holds, the debate has always been about normative constraints on explanation, then Bechtel's view is that epistemic constraints are important normative constraints on explanation, and if Waskan is right, then some psychological features are highly pertinent to such epistemic constraints. But note that Craver's worries about using equations in explanations is partly due to his desire that the normative constraints on explanation allow us to discriminate genuinely mechanistic explanation from other forms of explanation: 'The HH model might be included in the explanatory text, but the equation is neither a cause nor a constituent of action potentials. This mistake is committed by those who think of the HH model as a law that governs the action potential ... rather than as a mathematical generalization that describes how some of the components in the action potential mechanism behave ...' (Craver 2012, p. 5). Craver's objection is not to using maths per se, and I have already explained above that Bechtel is interested in equations and simulations that help you describe—and possibly even discover—working parts of real mechanisms.

Considered more closely, what Craver is really rejecting is an exclusive focus on psychologicistic theorizing, on those aspects of our representations of mechanisms that are subject to the needs of us as knowers, to the detriment of understanding that mechanistic explanations must also satisfy ontic constraints. For Craver, the activity of describing, simulating and otherwise representing mechanisms is important to science, and he shows this when he is actually working with mechanisms. He recognizes that epistemic aspects of mechanistic explanation—including diagrams—are vital to the success of scientific practice: 'Discovering a mechanism involves specifying and filling in the details of a schema, that is, instantiating it by moving to a lower degree of abstraction. As we will see, diagrams and equations are often employed to depict graphically the schematic organization of mechanisms' (Darden and Craver 2002, p. 4). Later, we get: 'Crick (1959) generated a set of alternative hypotheses to resolve this anomaly, localized in various stages of the mechanism schema' (Darden and Craver 2002, p. 14). The same idea of a large community of scientists communicating and acting on a common representation of a mechanism is also implicit throughough the argument for the mosaic unity of neuroscience in Craver (2007). While Craver emphasizes the ontic constraints, he employs epistemic constraints. He is not at all inclined to argue that we should never pay attention to epistemic and even to 'psychologicistic' aspects of explanatory practices. Craver (2012, p. 1) is quite explicit: 'I do not claim that one can satisfy all of the normative criteria on explanatory models, texts, or communicative acts by focusing on ontic explanations alone. Clearly there are questions about how one

ought to draw diagrams, organize lectures, and build elegant and useable models that cannot be answered by appeal to the ontic structures themselves.'

Craver has a final concern worth mentioning: 'There is one further reason to avoid equating scientific explanation with the abilities of individual cognitive agents. Some phenomena might be so complex that they overwhelm our limited cognitive systems' (Craver 2012, p. 16–17). This is right, but it is not an individual cognitive agent that Bechtel or Wright have in mind, but exactly the kind of collaborative scientific work that Craver recognizes.

So Bechtel's positive argument for epistemic constraints on mechanistic explanation is right. However, Craver's work shows that he also recognises such constraints. Further, Bechtel doesn't argue against ontic constraints being relevant to mechanistic explanation alongside epistemic ones. Indeed, I have already argued that Bechtel's work on decomposition and localization shows a deep appreciation of the ontic constraints that belong distinctively to mechanistic explanation. To summarise, I have now shown that we need to recognize both ontic and epistemic constraints on mechanistic explanation, and that the work of both Craver and Bechtel shows commitment to both kinds of constraints.

With this in view, we can return briefly to their concerns described in Sect. 1, and see that they vanish. Bechtel's concern was that now-absent mechanisms cannot explain in the ontic sense, and that mechanisms have been around for much longer than we have had mechanistic explanations. But with the shift to focus on constraints, and the admission of epistemic constraints, we have satisfied that. If we admit both kinds of constraints, then until we describe mechanisms, we may have causes, but we have no interesting explanations. And we can describe currently absent mechanisms. Particularly in the general case, we can describe how mRNA usually operates, whether or not it is currently so operating. Craver should also now have no problem admitting non-ontic entities like diagrams and equations, so long as they are not involved in mere phenomenal or how-possibly explanations, but are used to describe, and to allow us to manipulate ideas about, worldly entities, activities, and their organization.

So as Bechtel and Craver's concerns are expressed in their earlier work, those concerns have now been answered. However, a more sophisticated argument is explicit in Craver (2012), and a correlative argument may be detected as implicit in Bechtel and Wright's work. I will examine these now.

3 Are Either Ontic or Epistemic Constraints More Fundamental?

To recall briefly, the frame of the ontic-epistemic debate has changed in the mechanisms literature. There is a background of agreement about what mechanisms are, and the process of finding mechanisms, so that the area of dispute is narrow. I have also indicated the importance of the shift of focus to normative constraints: ways of evaluating good explanations. In this frame, I have argued that we should accept both ontic and epistemic constraints on mechanistic explanations.

The further argument of Craver's weaving through those I have already described is for the conclusion that ontic constraints are more fundamental than epistemic

constraints. I will examine his argument in this section, and argue that we can provide a parallel argument for the importance of epistemic constraints—an argument that may be implicit in Wright and Bechtel's paper. To summarise, this is because Craver's argument requires an assumption about the primary aim of explanation, while a parallel argument for epistemic constraints merely requires a different assumption about the primary aim of explanation. I will argue that we should accept both aims, as Bechtel and Craver both do. However, if we hold neither aim to be primary, we will accept both ontic and epistemic constraints as essential for successful mechanistic explanation, but hold neither to be fundamental.

I have argued extensively that Craver admits epistemic constraints on explanatory texts. But he also writes of an: ‘asymmetric direction of fit between the representation-involving ways of talking about explanation and the ontic perspective’ (Craver 2012, p. 7). Presumably he means to imply that there is no perspective from which representation-involving ways have some kind of priority. He writes: ‘The ontic structure of the world thus makes an ineliminable contribution to our thinking about the goodness and badness of explanations’ (Craver 2012, p. 11). Presumably, again, he means to claim that ontic constraints are ineliminable in a way in which epistemic constraints are not. So ontic constraints are asymmetrically prior or fundamental or ineliminable, and epistemic constraints aren't. This claim reintroduces genuine disagreement between Craver and Bechtel, as while I have argued that Bechtel should allow ontic constraints, he will certainly not allow them priority.

A similar idea seems to be present in Wright and Bechtel's paper. Consider again part of a quote given earlier: ‘what our understanding *literally* proceeds ‘through’ is a network of linguistically- or graphically-expressed operations on representations’ (Wright and Bechtel 2007, p. 50). In the phrase ‘literally proceeds through’, it is not hard to see the implicit idea that epistemic items are more important to explanation than ontic constraints.

The problem is, the only way to get priority for one or the other set of constraints seems to be to grant some kind of priority to the need that that set of constraints meets. We might grant priority to the need that our scientific explanations reflect the nature of reality. Certainly, if we grant priority to this aim, then the ontic constraints will inherit that priority over epistemic constraints. But clearly Bechtel will not admit priority to any such aim, so this assumption cannot be introduced into an argument with Bechtel without begging the question. In the second place, Bechtel could construct an entirely parallel argument in favour of the priority of epistemic constraints, so long as he is allowed to make his favourite assumption. Bechtel seems to believe that, while satisfying ontic constraints is important to mechanistic explanation, the most important aim of mechanistic explanation is the building of an understandable, mentally manipulable, communicable, cognitive model. If such an assumption is admitted, then epistemic constraints will be prior.

Two exactly parallel arguments can be run. I run them here side by side in order to show that neither can succeed without begging the question about the other. In favour of the priority of ontic constraints, Craver writes: ‘The point of these examples is that models may lead one to expect a phenomenon without thereby explaining the phenomenon. These judgments of scientific common sense seem to

turn on the hidden premise that explanations correctly identify features of the causal structures that produce, underlie, or otherwise responsible for the explanandum phenomenon (see Salmon 1984). Expectation alone does not suffice for explanation' (Craver 2012, p. 10). In parallel, Bechtel, or Wright, could say: 'one can identify features of the causal structures that produce, underlie, or are otherwise responsible for the explanandum phenomenon without thereby explaining the phenomenon. This judgement of scientific common sense seems to turn on the hidden premise that explanations convey some kind of understanding, so that they allow scientists to mentally model the phenomenon—perhaps to expect it. Identifying causal structure alone does not suffice for explanation.'

Again, Craver writes: 'The methods that scientists use to discover how the world works, the standards to which they hold such tests, are intimately connected with the goal of science to reveal the causal structures that explain why the phenomena of the world occur and why they occur as they do. One cannot carve off the practice of building explanations from these other endeavors' (Craver 2012, p. 11). But Bechtel, or Wright, could say: 'The methods that scientists use to model the world, to understand it mentally and communicate it, are intimately connected with the goal of science to achieve understanding, a communicable understanding, that is crucial to the forward-progress of science. One cannot carve off the practice of building explanations from this need for understanding.'

I think that Craver is exactly right in what he says. What is unnecessary is merely the extra claim that ontic constraints are fundamental. The parallel arguments that I have constructed are also right, but similarly not entitled to the conclusion that epistemic constraints have some kind of priority. It seems that the most sensible conclusion to draw is that neither aim of mechanistic explanation is prior to the other. Ontic and epistemic constraints are both ineliminable, as both aims must be met, to generate a successful mechanistic explanation:

- Describe the (causal) structure of the world: to be distinctively mechanistic, describe the entities and activities and the organization by which they produce the phenomenon or phenomena.
- Build a model of the activities, entities and their organization that scientists can understand, model, manipulate³ and communicate, so that it is suitable for the ongoing process of knowledge-gathering in the sciences.

If this is right, then both epistemic and ontic constraints are essential for mechanistic explanation, but neither is prior. Without the first constraint, we are not explaining the production of a phenomenon by a mechanism; without the second we do not achieve the understanding essential to explanation. Bechtel and Craver should both be prepared to accept this: Bechtel doesn't want to claim that what is really in the world doesn't matter; while Craver doesn't want to claim that our ability to understand explanations doesn't matter.

I will close this section by pausing to reconsider two different ways of construing the ontic-epistemic debate, from its origins, and addressing a final argument of Craver's. Craver and Bechtel agree on the vast majority of their claims about

³ This does not require manipulation in Woodward's sense.

mechanisms, and on the experimental work used to discover them, so they are much closer together than those in the original debate. We are also now in a position to appreciate fully how the shift to consideration of normative constraints on explanations both reflects the current context of agreement, and enables further consensus.

The original ontic-epistemic debate might be interpreted as more concerned with what explanations themselves are, as it were, metaphysically. This conception of the issue seems to have gone along with a tendency for making a binary choice: either explanations are portions of the mind-independent world, or they are some kind of representational stuff. Wright (2012) detects a similar polarisation in the original debate, but suggests that it derives from Salmon's ambiguity claim: 'The general idea is that we can speak of different senses of explanation, which in turn licenses a bifurcation of scientific explanations into two kinds: those that reside in the mind-independent world versus those that reside in the world-dependent mind' (Wright 2012, p. 381). Whatever the source of polarisation in the traditional debate, it is clear that normative constraints allow consensus, as admitting one set of constraints is no bar to accepting another set. Once the new frame of the debate is fully embraced, admitting both sets of constraints is the sensible option. This shift to arguing about normative constraints is a helpful step forward, as it encourages a consensus position that can satisfy significant thoughts polarising the original debate. One natural reason to claim that explanations are ontic things is the thought that explanations cannot ignore worldly things. One natural reason to claim that explanations are epistemic things is the thought that explanations must meet our epistemic needs. But both of these thoughts are met by admitting both sets of normative constraints, with further debate about the metaphysical nature of explanations rendered unnecessary. As Craver is now explicit in accepting the move to considering normative constraints on good explanation, I think he is bound to accept that some of the regulative ideals governing what makes a good mechanistic explanation include representational aspects.

It is in this context that I wish to reconsider a third argument of Craver's that I touched on in Sect. 1. In more detail, he says that there are 'sentences' that we need ontic explanation to make sense of:

- A. Our world contains undiscovered phenomena that have explanations.
- B. There are known phenomena that we cannot currently explain ... but that nonetheless have explanations.
- C. A goal of science is to discover the explanations for diverse phenomena.
- D. Some phenomena in our world are so complex that we will never understand them or model them, but they have explanations nonetheless (Craver 2012, p. 7).

Such sentences do give us some reason to think that there could be a purely ontic sense of explanation, where 'explanation' is interpreted as meaning little if anything more than 'cause'. For Wright (2012), this is a mistaken, loose use of 'explanation'. I argued in Sect. 2.1 that no-one can be rationally compelled to give up such an interpretation. However, I hope that the need to insist on that pure sense has now been severely undermined by the argument that the core concern of advocates of the

ontic conception such as Craver can be met by considering ontic constraints on exhibitions of worldly mechanisms. If, instead, fruitless disagreement over the use of words can be avoided, and consensus on substantive issues reached, this is a good reason for flexibility.

Finally, if one is still seeking a remaining difference between Bechtel and Craver, it is that Bechtel is cautious about making straightforwardly realist claims, generally seeking to avoid anything that verges on metaphysics. Bechtel is not anti-realist; he is just more cautious than Craver. However, this difference is not deep, given the extent of the agreement over the nature of mechanisms I presented in 1. Both know very well that in our epistemic practices we might not know whether we have before us a how-actually or a how-possibly model, and both agree about the kinds of experimental work done to try to decide between the two options. I think this closes the final path for regenerating disagreement, but the large body of consensus has survived.

4 For Integration

If I have been persuasive about a view that Craver and Bechtel can both accept, it is natural to wonder what remains at stake in the ontic-epistemic distinction. I have criticised Craver for over-concentration on ontic constraints, and Bechtel for over-concentration on epistemic constraints, but does this really matter? This is worth pause for thought, as it reveals why it is important to recognise both sets of constraints.

On the epistemic side, it is not difficult to make up stories about the world that nevertheless enhance understanding and even knowledge of the world. Consider Plato's image in the *Phaedrus* of the soul or psyche as a chariot with one black horse and one white, that conveys the idea of the psyche in conflict as the two horses try to pull in opposite directions. Plato uses this to illuminate what he takes to be a serious problem: explaining how one person can be in internal conflict. The explanation is strictly false, of course, but illuminating about a real phenomenon. For a more mundane example, consider explaining to a small child that mummy is getting fat because there is a baby in her tummy.⁴ Mummy is not getting fat, and the baby is not in her stomach. This explanation is strictly false, but still illuminating, and, as it is expressed in the only concepts the child understands, may be an ineliminable step on the road to full knowledge. On the ontic side, it is also not difficult to say something about the causal structure of the world, or even to say something about the right portion of the world for a particular phenomenon. When an explanation is sought, a scientist in a lab might point or wave at something and say we think it's that thing, there, that's doing it. Such pointing conveys something to the original enquirer, or indeed to the wider scientific field, but little of detailed usefulness. It is probably in this kind of way that most of the population knows that their hair and eye colour has something to do with their genes, in their cells.

⁴ Thanks to Julia Tanney for originally suggesting to me this endlessly useful example.

Each kind of constraint alone gives us some kind of useful set of norms for evaluating, and attempting to build, mechanistic explanations. However, it is important to recognise both sets of constraints precisely because that turns our attention very quickly to the relationship between ontic and epistemic constraints. The real achievement of mechanistic (and possibly other forms of) explanation is satisfying both ontic and epistemic constraints simultaneously, to get a story constrained by all the empirical contact with the world that ingenuity can design; a story that we can understand, manipulate and communicate, that we can use, and use collaboratively, to help us manipulate, control and predict the world—and lead science to better knowledge. This is the ongoing challenge of mechanism discovery. Missing this is what is damaging about Craver's over-concentration on ontic constraints, and Bechtel's over-concentration on epistemic constraints.

There is a further interesting question concerning the relationship between these two kinds of constraints. It arises when we ask what aspects of our mechanistic explanations we should take to tell us about what is in the world, and what aspects tell us more about human psychology and scientific convention than about the world. A natural first pass is the aspects satisfying ontic constraints tell us about the world, the aspects satisfying epistemic constraints only about our own knowledge-building practices.

But the story is more complex than that. There are different kinds of epistemic constraints.⁵ One set is entirely determined by the kind of successful communication necessary for the knowledge-building of the scientific community, and it is particularly unconnected from the mechanism described. For example, if it is easier for scientists (and students) to understand diagrams of protein synthesis if we use different colours for the background cell parts, and the different active parts, such as DNA, mRNA, and ribosomes, that is quite independent of the worldly parts described in such diagrams. As another example, explanations tailored to the needs of a particular audience, such as the small child above, will take on aspects that are not a part of our more general scientific explanations. Note that these kinds of epistemic aspects of these kinds of mechanistic explanations are not in serious danger of being mistakenly interpreted as reflecting the world. I have suggested that Waskan is right to think that these considerations are legitimately part of the concern of the epistemic conception of explanation.

The second set of epistemic constraints is more connected to the world. We abstract, selecting only some stuff in the world to describe, seeking to manipulate and communicate only information about some stuff and not about others. We seek explanations that are as general, unified, and as simple as possible. This is vital to the forward-looking aspect of scientific explanation, the forward-drive to better explanations. This fact does, quite rightly, advise caution, as we should expect some such features of our explanations to reflect our epistemic limitations and preferences, at times perhaps even more strongly than they reflect the world. This is just to recast the familiar point that unity, simplicity, and possibly many other things are vital constraints on scientific theorizing, but they do not necessarily

⁵ I thank Michael Strevens for pressing me on this point.

indicate unity and simplicity in the world. In the language of this paper, they are epistemic constraints, and so subject to reflecting accidents of human psychology.

What is crucial is that we will often not be in a position to know whether a particular aspect of a particular mechanistic explanation is there to satisfy epistemic constraints alone, or ontic, or both. And this is another reason why it is important to recognise both sets of constraints. It would be simpler if we could entirely separate ontic and epistemic constraints, in our practices, or in our theories, so that we can know cleanly what is in the world, and what is an aspect of our theorising. This is not as damaging as might be thought. The human brain, which generates the psychological needs that epistemic constraints must meet, is itself a worldly entity, existing in the world through time, affected by its engagement with the world—particularly with the empirical engagement with the world that is the process of building mechanistic explanations.

This means that what looks to us intelligible, simple and unified is not a static feature of human psychology, but is affected by our empirical engagement. Newtonian action at a distance used to seem quite impossible to us; so did quantum mechanical indeterminism, and non-locality. Physicists sincerely describe quantum mechanical equations as elegant, a claim that can generate hilarity in those unused to working with such theories. If empirical engagement continually forms what we find intelligible, simple and unified, then epistemic constraints on explanation, even the more ‘psychologistic’ constraints, are *deeply* entangled with ontic ones. There may always be some aspects of our theories for which there is no fact of the matter, at a particular moment in time, whether they reflect ontic or epistemic constraints. And any successful disentangling will need to look at what is happening over time, rather than at a single time.

I will finish as I began, with Salmon, as in some modes this seems to me an insight of Salmon’s. He writes: ‘What constitutes an adequate explanation depends crucially, I think, on the kind of world in which we live; moreover, what constitutes an adequate explanation may differ from one domain to another in the actual world’ (Salmon 1984, p. 299). Salmon does not make change over time explicit, but I have argued that this claim amounts to more than the claim that explanations must meet ontic constraints as well as epistemic. The point is that the very process of science affects what epistemic constraints we need to satisfy. It affects them so deeply that it is not always clear whether some constraints are epistemic or ontic. In the end, this means it is best to understand that our good mechanistic explanations are always the result of a struggle to satisfy both ontic and epistemic constraints.

Acknowledgements I have to thank Bill Bechtel, Carl Craver, Toby Friend, Bert Leuridan, John Pemberton, Federica Russo, Christoph Schultz, Michael Strevens, Adam White, and other members of the audiences of Causality and Explanation in the Sciences, Ghent 2011; and the Causality in the Sciences Reading Group, London 2011 for comments and discussion leading to substantial improvement to the paper. Foundational work for this paper was laid down during my postdoctoral project ‘Mechanisms and Causality’ at the University of Kent, funded by the Leverhulme Trust, although it was completed far beyond the end of the project. I am still grateful for that funding, and for the opportunity to work with Jon Williamson. Finally, I owe great thanks to two reviewers for efforts above and beyond the call of duty to help me improve the paper, but sadly they are anonymous. Remaining errors are, of course, my own.

References

- Bechtel, W. (2006). *Discovering cell mechanisms: The creation of modern cell biology*. Cambridge: CUP.
- Bechtel, W. (2008). *Mental mechanisms: Philosophical perspectives on cognitive neuroscience*. Oxford: Routledge.
- Bechtel, W. (2011). Mechanism and biological explanation. *Philosophy of Science*, 78, 533–557.
- Bechtel, W., & Richardson, R. (2010). *Discovering complexity*. Cambridge, MA: MIT Press.
- Craver, C. (2007). *Explaining the brain*. Oxford: Clarendon Press.
- Craver, C. (2012). Scientific explanation: The ontic conception. In A. Hüttemann, M. Kaiser (eds), *Explanation in the biological and historical sciences*. Berlin: Springer.
- Darden, L., & Craver, C. (2002). Strategies in the interfield discovery of the mechanism of protein synthesis. *Studies in the History and Philosophy of the Biological and Biomedical Sciences*, 33, 1–28.
- Glenann, S. (2002). Rethinking mechanistic explanation. *Philosophy of Science*, 69, S342–S353.
- Glenann, S. (2005). Modeling mechanisms. *Studies in the History and Philosophy of Biology and Biomedical Sciences*, 36, 443–464.
- Illari, P. M., & Williamson, J. (2012). What is a mechanism? Thinking about mechanisms across the sciences. *European Journal of the Philosophy of Science*, 2, 119–135.
- Machamer, P. (2004). Activities and causation: The metaphysics and epistemology of mechanisms. *International Studies in the Philosophy of Science*, 18(1), 27–39.
- Machamer, P., Darden, L., & Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science*, 67, 1–25.
- Salmon, W. (1984). Scientific explanation: Three basic conceptions. *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 2, 293–305.
- Waskan, J. (2011). Intelligibility and the cape: Combatting anti-psychologism about explanation. In *Epistemology of modeling & simulation: Building research bridges between the philosophical and modeling communities*. Pittsburgh. <http://philsci-archive.pitt.edu/view/confandvol/2011eomsrbpbpmcph13april2011.html>. Accessed June 2013.
- Wright, C., & Bechtel, W. (2007). Mechanisms and psychological explanation. In P. Thagard (ed), *Philosophy of psychology and cognitive science*. Amsterdam: Elsevier.
- Wright, C. D. (2012). Mechanistic explanation without the ontic conception. *European Journal for Philosophy of Science*, 2, 375–394.