Organon F, Volume 26, 2019, Issue 1

Special Issue

Causality, Free Will, and Divine Action

Frederik Herzberg and Daniel von Wachter

Guest Editors

Contents

Organon F 26 (1) 2019: 2–4
https://doi.org/10.31577/orgf.2019.26101
ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

Preface

This issue of *Organon F* is dedicated to arguments from determinism or similar principles against libertarian free will and against the existence of divine interventions, such as miracles. Such arguments have been very influential especially since the 19th century, but they came up already in the 17th century.

In Thomas Hobbes we find the first clear statement of the argument from determinism against libertarian free will. In his book *Elements of Philosophy* of 1655 in the section *De corpore,* Hobbes reasons about what it is for an event to come to occur and concludes that every event must be necessitated by antecedent events, which is the doctrine that we call 'determinism' today. From this he derives that there cannot be libertarian free will: 'That ordinary definition of a free agent, namely, that a free agent is that, which, when all things are present which are needful to produce the effect, can nevertheless not produce it, implies a contradiction, and is nonsense.' (*Of Liberty and Necessity,* § 32)

Arguments of this kind, which assume that there is 'no room' for libertarian free will or that it is impossible, have been the main objection against libertarian free will especially since the 19th century, and they still are. In German-speaking philosophy, belief in determinism was promoted by Kant's principle of causality: 'every event is determined by a cause according to constant laws.' (*Prolegomena,* § 15) A different version of the argument, inspired by David Hume, gives up necessitation but refers only to laws that entail regularities of succession of the type 'All events of type x are followed by events of type y.' Many philosophers find the idea charming that the laws of nature and the description of the state of the universe at one time together entail descriptions of the state of the universe at all other times. Today the most widespread argument of this kind against libertarian free will refers to the principle of the causal closure of the physical, which is investigated by several contributions to this issue. Hobbes and Kant held that

© The Author. Journal compilation © The Editorial Board, *Organon F.*

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International Public License (CC BY-NC 4.0).

the truth of determinism is known a priori, others claim that it is known a posteriori, for example through conservation laws.

These arguments against libertarian free will are very similar to the arguments that German theologians put forward against the existence of divine interventions. The German theologian Ernst Troeltsch formulated this in 1898 in the following principle: 'No change can occur at one point without changes occurring before and after at other points, so that all events stand in a continuous, correlative interconnection and must necessarily constitute a single flow in which each and all hang together, and every event stands in relation to others.' The US American theologian Langdon Gilkey wrote in 1961: 'Contemporary theology does not expect, nor does it speak of, wondrous divine events on the surface of natural and historical life. The causal nexus in space and time which the Enlightenment science and philosophy introduced into the Western mind is also assumed by modern theologians and scholars.' By 'causal nexus' apparently he means determinism.

If such arguments against libertarian free will or against divine interventions are successful, then they are an effective and quick way of knowing something as important as whether we have libertarian free will and whether there are divine interventions. We can then know the truth about these issues without having to investigate specific evidence such as evidence for the resurrection of Jesus, evidence about brain events that have no event cause, or evidence from introspection. However, the question is whether determinism or some other principle from which we can derive that there is no libertarian free will is true and whether we can know it.

Robert Larmer investigates in his contribution 'methodological naturalism', i.e. the claim that scientists should always posit a natural cause for any event that takes place in the natural world. Richard Swinburne argues that the principle of causal closure is self-defeating: no one could ever be justified in believing it. Daniel von Wachter defends a 'principle of causal openness of the physical' that is contrary to the principle of causal closure. In a discussion note Michael Esfeld objects to Wachter that the principle of causal closure is best understood as a non-modal principle and that it can be known through the laws of nature. Wachter responds by distinguishing laws of nature from the differential equations that can be derived from laws. In his article Esfeld explains why physical theories with deterministic dynamical equations are to be preferred, but

then argues that the parameters in the equations do not refer to properties of the physical systems and that therefore there is no conflict between determinism in physics and free will. Ralf Bergmann investigates causal networks from the point of view of physics and argues that divine interventions would not violate the laws of nature. Uwe Meixner relates free will to quantum physics and develops the idea of a nonphysical organ of higher organisms. Ansgar Beckermann discusses arguments for dualism based on introspection and raises objections against agent causation. And Thomas Pink investigates whether freedom can exist as a form of power.

The contributors to this issue started their discussion on the topic at a conference that took place on 13–15 September 2017 in Vienna and that was a part of the project 'The Openness of the Universe for Free Will and Special Divine Action,' generously funded by the John Templeton Foundation.

Daniel von Wachter
International Academy of Philosophy
in the Principality of Liechtenstein

Organon F 26 (1) 2019: 5–24
https://doi.org/10.31577/orgf.2019.26102
ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

RESEARCH ARTICLE

The Many Inadequate Justifications of Methodological Naturalism

Robert A. Larmer*

Received: 24 May 2018 / Accepted: 4 September 2018

Abstract: Contrary to proponents' claims, methodological naturalism is not metaphysically neutral. Consequently, its acceptance as a practice requires justification. Unfortunately for its advocates, attempts to justify it are failures. It cannot be defended as a definition, or a self-imposed limitation, of science, nor, more modestly, as an inductively justified commitment to natural causes. As a practice, it functions not to further scientific investigation, but rather to impose an explanatory straitjacket.

Keywords: Ad hominem fallacy; nomological science; historical science; supernatural agency; inductive generalization; Robert Pennock.

1. Introduction

In 1983, Paul de Vries, a philosophy professor at Wheaton, a conservative Christian liberal arts college, advocated the practice of what he termed 'methodological naturalism' in relating scientific and religious

* University of New Brunswick
✎ Department of Philosophy, University of New Brunswick, 19 Macaulay Lane, Carleton Hall 211, Fredericton NB, E3B 5A3, Canada
✉ rlarmer@unb.ca

© The Author. Journal compilation © The Editorial Board, *Organon F*.

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International Public License (CC BY-NC 4.0).

beliefs.[1] His claim was that, as a matter of method, scientists, whatever their metaphysical beliefs, should always posit a natural cause for any event that takes place in the natural world. Since then, proponents, both secular and religious, insist that adopting such a method in no way commits one to any specific metaphysical position, since science is naturalistic only on the level of its methodology (MN), but is neutral with respect to metaphysics. Thus, given its presumed metaphysical neutrality, methodological naturalism provides a way in which science can be pursued by those with differing world-views.

The presumed insulation of methodological naturalism from any kind of metaphysical commitment is an illusion (Larmer 2003, 113–30). What one thinks to be the nature of reality cannot be neatly separated from the methods one uses to investigate it.[2] If, for example, one believes that there exist,

---

[1] Numbers writes that the term

> 'methodological naturalism' seems to have been coined by the philosopher Paul de Vries, then at Wheaton College, who introduced it at a conference in 1983 in a paper subsequently published as "Naturalism in the Natural Sciences," Christian Scholar's Review, 15 (1986), 388–396. De Vries distinguished between what he called "methodological naturalism," a disciplinary method that says nothing about God's existence, and "metaphysical naturalism," which "denies the existence of a transcendent God." (Numbers 2003, 320, Note 2)

Davis, however, notes the earlier use of the term, most notably by Edgar Brightman.

[2] Burtt, commenting on the presumption that methodology need have no links to metaphysics, notes that

> There is no escape from metaphysics, that is, from the final implications of any proposition or set of propositions. The only way to avoid becoming a metaphysician is to say nothing [...] If you cannot avoid metaphysics, what kind of metaphysics are you likely to cherish when you sturdily suppose yourself to be free from the abomination. Of course, it goes without saying that in this case your metaphysics will be held uncritically because it is unconscious; moreover, it will be passed on to others far more readily than your other notions, inasmuch as it will be propagated by insinuation rather than by direct argument [...] The history of mind reveals pretty clearly that the thinker who decries metaphysics will actually hold metaphysical

or may possibly exist, mental states which play a causal role in determining bodily behaviour, it makes no sense to adopt methodological behaviourism, since its adoption guarantees the development of psychological theories in which mental states either do not exist or play no causal role in such behaviour. Only if one has already established beyond reasonable doubt that mental states do not exist or, if they do exist, play no causal role does it make sense to insist on methodological behaviorism as a prerequisite of developing psychological theories. To insist on its employment in the absence of compelling reasons for disbelieving in the existence of mental states or their causal powers is to beg the question of whether its adoption is justified. Similarly, if one has not established beyond reasonable doubt that supernatural agents do not exist, or, if they do, they never intervene on natural processes, does it make sense to insist that explanations of physical events must restrict themselves to natural causes.

A further reason to question the easy acceptance of methodological naturalism is that there is no clear way to demarcate science from non-science.[3]

> notions of three main types. For one thing, he will share the ideas of his age on ultimate questions, so far as such ideas do not run counter to his interests or awaken his criticism. [...] In the second place, if he be a man engaged in any important inquiry, he must have a method, and *he will be under a strong and constant temptation to make a metaphysics out of his method, that is, to suppose the universe ultimately of such a sort that his method must be appropriate and successful* [...] Finally, since human nature demands metaphysics for its full intellectual satisfaction, no great mind can wholly avoid playing with ultimate questions [...] But, inasmuch, as the positivist mind has failed to school itself in careful metaphysical thinking, its ventures at such points will be apt to appear pitiful, inadequate, or even fantastic. (Burtt 1932, 224–26, emphasis added)

[3] The failure of the demarcation quest to provide a litmus test between 'science' and 'non-science' is generally acknowledged in the literature. This should come as no surprise, since, as John Earman notes, "it does not much matter what label one sticks on a particular assertion or an enterprise; the interesting questions are whether the assertion merits belief and whether the enterprise is conducive to producing well-founded belief" (Earman 2000, 3). A recent attempt to defend the possibility of such a litmus test is (Pigliucci and Boudry eds., 2013).

Not only do proposed demarcation criteria prove inadequate, they are typically employed polemically as discrediting devices, as "*machines de guerre*" (Laudan 1996),[4] whereby one can dismiss an opponent's position as 'unscientific' and thus unworthy of being taken seriously. Thus, for example, the arguments made by intelligent design theorists are routinely dismissed without serious examination, on the grounds they are deemed 'unscientific.'[5]

It is clear, therefore, that proponents of adopting methodological naturalism need to justify their insistence that scientific investigation must never take into consideration the possibility of a supernatural cause of observed phenomena. This, as we shall see, is no easy task.

2. Proposed justifications of methodological naturalism

(1) *The motivations of those questioning methodological naturalism are suspect.*

Those questioning methodological naturalism as a prerequisite of science are frequently dismissed on the basis that they are 'creationists.' Unfortunately, such dismissals equivocate on the term 'creationist' to the extent that it comes to mean anyone who questions the acceptance of methodological naturalism. If, however, one takes the term 'creationist' in its more usual

[4] Laudan goes on to comment that "many of those most closely associated with the demarcation issue have a hidden [...] (and sometimes not so hidden) agenda of various sorts" and that "if we [...] stand [...] on the side of reason, we ought to drop terms like 'pseudo-science' and 'unscientific' from our vocabulary; they are just hollow phrases which do only emotive work for us" (Laudan 1996, 344, 349).

[5] Such a strategy may be rhetorically effective, nevertheless it is logically fallacious. As Stephen Dilley remarks

> mere terminological labels do not change epistemic properties. Just as theists cannot lower the epistemic plausibility of [naturalist] hypotheses merely be deeming them 'arrogant bluster' so naturalists cannot lower the epistemic plausibility of God hypotheses by labeling them 'unscientific.' As an epistemic matter, each rival hypothesis must be evaluated on its evidential and conceptual merits. (Dilley 2010, 136).

use as referring to those who believe the earth is less than 20,000 years old, then it is clear there are many critics of methodological naturalism who are not creationists, and indeed not even theists.[6]

More fundamentally, such a justification fails, since it is clearly a case of the *ad hominem* fallacy. Whatever the motivations of those questioning methodological naturalism what should really be at issue is the arguments they present. If one wants to show that Richard Dawkins views on evolution are mistaken one must examine his arguments and not simply observe that he finds evolution an attractive theory on the basis that it provides a materialist origins story. Similarly, if one wants to dismiss critics of methodological naturalism one must do it by showing their arguments to be mistaken, rather than questioning their motives.

(2) *Science in principle excludes any recognition of the supernatural.*

Many proponents of methodological naturalism insist that, by definition, science cannot ever contemplate the existence of supernatural causes. Robert Pennock insists that methodological naturalism is "a scientific ground rule" (Pennock 2011, 184), that is to say, "as a point of method science does not countenance appeals to the supernatural" (Pennock 2011, 185).[7] On such a view, methodological naturalism constitutes a necessary, though not sufficient, condition of scientific investigation, and thus serves to at least partially demarcate science from other disciplines.

There are at least two reasons to reject such a justification of methodological naturalism. First, in the absence of argument as to why science must exclude recognition of the supernatural, it amounts to an arbitrary stipula-

[6] See, for example, (Monton 2013).

[7] A variation on this strategy is to insist that if the postulation of a 'supernatural' cause for a physical phenomenon became necessary, then that cause must be conceived as natural, since, by definition, the supernatural is "unknowable by means of scientific inquiry" (Forrest 2000, 14). She insists that "such confirmation would only demonstrate that this newly verified aspect of reality had all along never been supernatural at all" (Forrest 2000, 25). This, of course, empties the terms 'natural' and 'supernatural' of all content, since even God, understood as the ontologically distinct, and creator, *ex nihilo*, of all other entities, would in such circumstances have to be conceived as 'natural.'

tion. If one wants to claim that science prohibits ever recognizing a supernatural cause then one must provide reasons why this is the case, not simply define such an embargo into existence.

Second, demarcationist proposals are notorious for failing to provide necessary or sufficient conditions to distinguish science from other disciplines.[8] Making the important distinction between what may be termed 'nomological' science and 'historical' science further compounds the difficulty inherent in such an enterprise. Paul Draper notes that

> scientists engaged in nomological science formulate laws, models and other interesting if-then generalizations, often testing them by experiment and prediction, and making inductive generalizations based on observable data. In historical science, on the other hand, not all causal explanations fit the covering law model and many hypotheses about the past cannot be falsified and cannot be tested by prediction or experiment. Instead, they are judged on their simplicity, their fit with general background knowledge about the world, and their ability to explain specific known facts. What all this shows is that methodological naturalism cannot be adequately defended by describing something called *the* scientific method then arguing that it cannot be applied to the supernatural. For more likely than not, the method described will be characteristic of nomological science, while appeals to the supernatural would naturally be used to answer historical questions. (Draper 2005, 290)

(3) *Supernatural causation implies a chaotic universe.*

Not infrequently, advocates of methodological naturalism, attempt to justify its acceptance by claiming that taking seriously the possibility of supernatural agency implies a chaotic universe. They maintain that taking seriously such a possibility implies that "God may simply [...] zap anything into or out of existence [...] in any situation, any pattern (or lack of

[8] Del Ratzsch makes the point that "definitional attempts [to justify methodological naturalism] are *prima facie* problematic for the simple reason that no one actually has a completely workable definition of science (nor even necessary and sufficient conditions), and that proposed definitions have been historically unstable" (Ratzsch 2004, 441).

pattern) of data is compatible with the general hypothesis of a supernatural agent unconstrained by natural law" (Pennock 2001, 89). "We cannot live simultaneously in a world of natural causation and of miracles, for if one miracle can occur, there is no limit" (Lewontin 1983, xxvi) and "at every instant all physical regularities may be ruptured and a totally unforeseeable set of events may occur" (Lewontin 1983, xxvi).

Aside from the deep conceptual confusion involved in suggesting that choosing a certain methodology determines whether in fact the universe is chaotic, several criticisms are in order as concerns this attempted justification. First, it is generally recognized that science originated, developed, and took place in a Western European Christian intellectual environment. As Del Ratzsch notes,

> science works only in a very particular sort of reality and only with a very particular sort of conception of reality. The requisite picture—of a comprehendable, intelligible, uniform, predictable, even beautiful cosmos which can in principle make sense to finite minds like ours when observed via perceptual faculties like ours—is a picture of a cosmos structured like a *creation*. Although details are disputed, that Christian doctrines of creation and of divine voluntarism provided a hospitable matrix for science is not in dispute.

At the very least, it is historically clear that belief in the reality, of supernatural causation did not hinder the development of science.[9]

Second, it does not follow that admitting the reality of supernatural agency would imply that 'anything goes,' that on the hypothesis of theism God is liable at any moment to zap anything into or out of existence." What God freely wills will be accordance with his nature and not simply arbitrary or irrational. As Evan Fales comments, "it does not follow from the fact

[9] Ratzsch notes that,

> It is not necessarily irrelevant that it was not until nature was looked at as a product of design—i.e. as a creation—that science itself really got off the ground. Blanket stipulative prohibitions (definitional or otherwise) against *exactly* that initiating intuition would seem to demand extraordinary justification. (Ratzsch 2004, 443)

that God is a free agent that His purposes and behavior (including the occasional performance of a miracle) cannot be made intelligible or studied in systematic ways" (Fales 2010, 5).

Third, those maintaining the reality of supernatural causation offer criteria by which phenomena best understood as requiring a supernatural cause as opposed to a natural cause can be identified.[10] One may wish to dispute whether these criteria are effective, but the fact that they are proposed and that supernatural causation is advocated in the realm of 'historical' as opposed to 'nomological' science, makes clear that openness to the possibility of supernatural causation does not commit one to abandoning belief in an orderly universe, amenable to human investigation.

(4) *Allowing for the possibility of supernatural causation is a 'science-stopper.'*

Defenders of methodological naturalism often make the claim that taking seriously the possibility of supernatural causation is a 'science-stopper.' It is argued that at the psychological level scientists will become lazy and liable to abandon the search for natural causes for phenomena, and at the conceptual level that explanations in terms of supernatural causes are not falsifiable. Warnings abound that any openness to considering a supernatural cause will bring the scientific enterprise to a grinding halt.[11]

Once again, there are reasons to question such an assumption. We have already noted that belief in the reality of supernatural agency did not hinder scientific development.[12] We have also noted that the postulation of supernatural causation typically occurs regarding 'historical' rather than 'nomological' science.

---

[10] See, for example, (Larmer 2014, 79–87).

[11] Pennock claims that if "supernatural explanations are permitted [...] all empirical investigation beyond the purely descriptive could cease, for scientists would have a ready-made answer for everything (Pennock 2001, 90).

[12] Ratzsch notes that,

> neither science or scientists may be vulnerable to the temptations of intellectual sloth as presumed. Indeed, the history of science would suggest that the risks are not that great *on precisely this point*. His-

Neither is it the case that once proposed, or even accepted, that supernatural explanations cannot be challenged or discarded. Competing explanatory hypotheses are hardly unknown to science. To take a case in point, the fact that biological entities exhibit the informational patterns typically associated with intelligent agency, and that scientific investigation tends to emphasize the inadequacy of solely natural causes to account for the genesis of such structures, in no way implies that research cannot, or is not likely to, continue regarding the possibility of demonstrating that a plausible naturalistic account of such origins can be given.[13] If such a plausible account emerges then it will constitute reason to reject an explanation in terms of supernatural agency.[14]

> torically, no disaster such as that darkly suggested by Pennock occurred. In fact, if the history of science told by *critics* of teleology, creationism, intelligent design, and the like is accurate, during the 19th century previously entrenched supernatural design explanations *lost* the scientific battle to mere fledgling naturalistic explanations—hardly what one would expect if merely allowing currently *dis*enfranchised supernatural design explanations into the conversation were likely to destroy current mature and robust natural science. (Ratzsch 2004, 441)

[13] Ratzsch observes that,

> claims that design theories threaten the utter ruin of science [...] [are] less than wholly persuasive [...] Despite the popularity of such claims, I have not seen the slightest hint of even a presumptive example within the last several *centuries* where some design-friendly theory has challenged a 'proper' scientific theory and managed to displace, or even a case where some scientifically improper design theory which has ('unfortunately') already been in place within science has itself survived the challenges of legitimate science, thereby destroying legitimate scientific progress. (Ratzsch 2004, 138)

[14] Dembski comments that,

> If it could be shown that biological systems that are wonderfully complex, elegant and integrated [...] could have been formed by a gradual Darwinian process [...] then intelligent design would be refuted on the general grounds that one does not invoke intelligent causes when undirected natural causes will do. In that case Ockham's

Further, the insistence that it is never permissible to posit a supernatural cause for a physical event means that if such causation in fact takes place it can never be recognized. Unless one is prepared to defend metaphysical naturalism on an independent basis,[15] acceptance of methodological naturalism as essential to the pursuit of science requires that one understand science not as committed to pursuing the truth about reality, but rather as to pursuing the best natural explanation that can be formulated of a phenomenon.[16] Thus, acceptance of methodological naturalism requires that, no matter how implausible a naturalistic explanation for the origin of life might be it will be taken seriously so long as it is only slightly less implausible than competing naturalistic explanations. Even if life did originate through supernatural agency, and even if it bears the characteristics of things we know to be intelligently designed, acceptance of methodological naturalism prohibits ever contemplating such an explanation.[17] Similarly,

---

> razor would finish off intelligent design quite nicely. (Dembski 2004, 281)

An anonymous reviewer has criticized me on the basis that Dembski's theory of detecting design is 'wrong-headed.' I am puzzled by this criticism, since this, my only reference to Dembski, deals with a different point entirely.

[15] Emphasis must be placed on there being justification of metaphysical naturalism independent of any prior acceptance of methodological naturalism. On pain of begging the question, it will not do first to accept methodological naturalism, and then insist that, since there is no evidence of supernatural causes, that belief in metaphysical naturalism is justified on the ground of Occam's Razor. For a more complete explication of this point see (Dilley 2010) and (Larmer 2003).

[16] Proponents of methodological naturalism have the logical option of conceiving science antirealistically. Taking this option, however, removes any possibility of objecting to claims of supernatural intervention on the grounds that such claims are unscientific. I am not aware of any advocates of methodological naturalism who are antirealists concerning science.

[17] Ratzsch notes that,

> If (perhaps for overwhelmingly good reasons) science is restricted (even just methodologically) to 'natural' explanatory and theoretical resources, then if there is a supernatural realm which does impinge upon the structure and/or operation of the 'natural' realm, then the world-picture generated by even the best science will unavoidably be

methodological naturalism prohibits ever recognizing an event, no matter the context and circumstances in which it occurs, as a miracle.[18] Thus, if convinced that Jesus did in fact rise from the dead, a consistent methodological naturalist must presume the event does in fact have a natural explanation.

(5) *Theological considerations require there be no supernatural intervention in nature.*

Methodological naturalism is sometimes defended on explicitly theological grounds. These take the form of assertions that the perfection of God implies that His purposes in the world must be accomplished exclusively—perhaps in the case of the more conservatively minded an exception being made for 'salvation history'[19]—by means of secondary created causes. Nature, it is asserted, is fully-gifted and any supernatural intervention by God in creation would be coercive and inconsistent with God's perfect love.[20]

There are a least two reasons to see such theological justifications as carrying little weight. First, even the most cursory examination reveals that they are question-begging and employ rhetorically loaded language. We are informed that "the notion of God [as] [...] meddling with matter, [...] is offensive [...] it would be a very poor sort of god who created a universe

> either incomplete or else wrong on some points. Unless one assumes philosophical naturalism (that the natural constitutes the whole of reality) that will be the inescapable upshot of taking even mere methodological naturalism as an essential component of scientific procedure" (Ratzsch 2002, 4)

(I am grateful to the anonymous referee who brought this quotation to my attention.)

[18] For a defense of the traditional conception of miracle as a supernatural intervention in nature see (Larmer 2014, 7–46).

[19] Any such qualifications require some principled reason why methodological naturalism *must* be applied in one area of investigation, but not in another. The issue should be whether some events are best explained in terms of supernatural causation. If this is the issue, then it will not do to insist a priori that such explanations *must* be restricted to 'salvation history.'

[20] See, for example, (Van Till 2002, 114).

that wasn't right and then tinkered with it at later stages" [Davies 2012, quoted in (Ratzsch 2001, 198, Note 19), emphasis added], and that "a God who uses the openness of his created universe [...] to insert additional causal events from time to time into that universe to produce particular events or trends [...] would be a meddling demigod, a moral monster, and a contradiction of himself" (Jenkins 1987, 63).

Reasons, however, why God's perfection requires no supernatural intervention in the natural order are noticeably lacking, as are arguments why such interventions should be understood as meddling or tinkering.[21] One is expected to accept that belief in supernatural intervention in nature necessitates viewing God as either a moral monster or a bumbler, who, not getting things right the first time, must adjust an ill-thought out plan in an *ad hoc* manner. The possibility that God at times acts directly within creation, that inorganic chemicals were never intended to have the capacity to self-assemble into living entities, is ruled out a priori, with no need to consider actual empirical evidence.

Second, such justifications are inconsistent with their proponents' claim that methodological naturalism is metaphysically neutral. The methodology they insist upon is grounded in their metaphysical commitment to a non-interventionist God. Far from being metaphysically neutral, it is an outworking of their deistic or semi-deistic view of God's relation to creation, that, is to say, their insistence that God be viewed as acting exclusively, or almost exclusively, through the instrumentality of secondary material causes.

(6) *Methodological naturalism is inductively justified.*

Philosophically astute proponents of methodological naturalism, both theists and non-theists, have increasingly tended to defend it not as necessarily constitutive of doing science, but as based on a well-evidenced

[21] An anonymous reviewer made the point that the fact that such reasons are not given does not mean that they do not exist. Fair enough, but such reasons must be given if the objection is to be taken seriously. If, and when, they are, then the force of such reasons can be evaluated. Even in such a case, they would have to be weighed against the empirical evidence for divine intervention.

inductive generalization that proves fruitful.[22] Writing from a Christian perspective, Patrick McDonald and Nivaldo Tro, are happy to agree that methodological naturalism is not an essential or definitional aspect of science, that there "are no clear a priori reasons to exclude the supernatural from science; [rather, it is] an empirically validated methodology and as such should be honored unless and until a better framework comes to the fore" (McDonald and Tro 2009, 203). Similarly, metaphysical naturalists Maarten Boudry, Stefaan Blancke, and Johan Braeckman reject the conception of methodological naturalism as an intrinsic limitation of science, defending it "as a provisory and empirically grounded attitude of scientists, which is justified in virtue of the consistent success of naturalistic explanation and the lack of success of supernatural explanations in the history of science" (Boudry et al. 2010, 227).

This of all the proposed justifications of methodological naturalism is the most promising. Nevertheless, it is far from convincing. First, at the level of actual practice, methodological naturalism functions not so much as a provisional inductive generalization helpful in guiding scientific investigation, but rather as an absolute prohibition on ever taking seriously the possibility of supernatural causation.[23] Richard Lewontin's comment regarding intelligent design and his "willingness to accept claims which are against common sense" (Lewontin 1997) is revealing. He writes,

> it is not that the methods and institutions of science somehow compel us to accept a material explanation of the phenomenal world, but on the contrary, that we are forced by our a priori

[22] Elliot Sober in his (2011) comments that

> arguments against introducing the claim that God exists into scientific theories have often been *in-principle*; they attempt to show that this postulate *necessarily* prevents science from reaching one of its goals [...] The argument I would offer is more modest. Naturalistic science has been a success [...] The modest defense I would offer of methodological naturalism is simply this: *if it isn't broken, don't fix it.* (Sober 2011, 375)

[23] An anonymous reviewer sees this point as feeble. With respect, I disagree. Questions of how methodological naturalism is employed are relevant to its presumed justification.

> adherence of material causes to create an apparatus of investigation and a set of concepts that produce material explanations, no matter how counterintuitive, no matter how mystifying to the uninitiated. Moreover, that materialism is absolute, for we cannot allow a Divine Foot in the door. (Lewontin 1997)

Fundamental explanatory virtues such as simplicity, coherence, scope, etc. serve to adjudicate competing explanations, whether they be natural or supernatural, without any need to invoke methodological naturalism as adding some further necessary criterion.

Second, and more importantly, the inductive generalization which is presumed to justify methodological naturalism is far less secure than generally portrayed. This is so for several reasons. First, at a historical level, there are many factors which explain the move away from theories of supernatural agency, that, is to say design inferences. Many of these "are scientifically irrelevant, and to the extent that they drove the history, to that extent the history too will be irrelevant" (Ratzsch 2005, 128–36). Further, it seems fair to observe that the rejection of present day design inferences should be based on present day science. By way of analogy, Ratzsch notes that

> were one attempting to show that phlogiston theory really should not be resurrected within present science, one surely would not have to rely on reference to difficulties raised a couple of centuries back. If phlogiston theory is indeed dead, we'd better be able to develop a case out of present science for thinking so. If we can't then present science is in extraordinarily deep trouble. (Ratzsch 2005, 136)

Second, as has been noted, appeals to supernatural agency are not typically found regarding how things work, but rather how they come to exist in the first place. Questions of the origin of the universe and its apparent fine-tuning and questions of the origin and development of life are questions of 'historical' rather than 'nomological' science. It is in these areas of scientific investigation that one encounters appeals to supernatural agency, which is to say, design inferences.

Once this distinction is made, the success that 'nomological' science has had in finding natural causes of phenomena cannot automatically be

taken as providing a strong inductive argument for adopting methodological naturalism in historical sciences. The fact that turtles are easy to catch hardly provides warrant for thinking that cheetahs will be easy to catch, and the fact that natural explanations in nomological science have enjoyed great success, scarcely warrants the assumption that explanations in terms of natural causes in historical science will enjoy the same degree of success.

Indeed, any inductive argument regarding cosmological and biological origins seems to be on the side of design inferences. A little over a century ago it was possible to view the universe as eternal and the structure of living cells as relatively simple. This is no longer so; the more we know the harder it is to avoid positing design regarding origins. Our best cosmology points to our finely-tuned universe coming into existence without any natural cause.[24] Similarly, although structures that manifest redundant order, e.g. crystals, or simply complexity, e.g. mixtures of random polymers, are easily found in nature, none of them exhibit the specified complexity, that, is to say, information that exists in DNA, RNA, and proteins. Such structures recalcitrantly resist explanations in terms of any known natural causes, so much so that James Tours, one of the top-ranked chemists in the world, is prepared to say in his 2016 Pascal Lecture "The Origin of Life: An Inside Story" that there presently exists no naturalistic account of life's origin that is even faintly plausible (Tours 2016).

It is important in this regard to realize that the argument does not commit the fallacy of *ad ignorantium*. The inference to supernatural agency is based not simply on the repeated inability of scientists to produce plausible natural explanations in these areas, but on the fact that the phenomena being investigated display characteristics that, in our experience, are the product of intelligent agency. As Stephen Meyer notes,

> the inadequacy of proposed materialistic causes forms only part of the basis of the argument for intelligent design. We also *know* from broad and repeated experience that intelligent agents [...]

[24] See, for example, (Spitzer 2010, 13–74). Also, (Craig and Sinclair 2009). An anonymous referee asks whether these sources consider quantum gravitational theories. The answer is yes.

> produce information-rich systems [...] Experience teaches that whenever large amounts of specified complexity or information are present in an artifact or entity whose causal story is known, invariably creative intelligence—intelligent design—played a role in the origin of that entity. Thus, when we encounter such information in the large biological molecules needed for life, we may infer—based on our knowledge of established cause-and-effect relationships—that an intelligent cause operated [...] [the argument] asserts the superior explanatory power of a proposed cause based upon its proven—its *known*—causal adequacy *and* based upon a lack of demonstrated efficacy among the competing proposed causes. (Meyer 2009, 376–77)

Further, there are scientific disciplines where design inferences are routinely made. Archeologists, for example, are willing to identify primitive tools precisely on the basis that what they are looking at bears the type of structure we know to be produced by intelligence, and the SETI project is looking to see if there are signals from space that exhibit a pattern requiring intelligence to explain. These appeals to design appear to be accepted because they are not presumed to challenge a naturalistic account of how conscious intelligent agents originated.[25] It is only when the design would have to be attributed to a supernatural intelligent agent that the design is judged to be only apparent and not genuine.

This, however, is to ignore the fact that the question of whether an object is designed, that, is to say the product of intelligent agency, is distinct from the question of the designer. If one goes to a faraway planet and finds symbols on a cliff face that turn out upon inspection to demonstrate the impossibility of trisecting the angle and doubling the cube, one will have no hesitation in identifying that mathematical proof as the result of intelligence, even if one has no idea as to the identity of the agent responsible. Given that the recognition of intelligent design is logically prior to the question of the designer's identity, it will not do to reject what appear to be instances of design on the basis that, if recognized as

[25] Anyone familiar with contemporary philosophy of mind and what naturalist philosophers term the 'hard problem of consciousness' will realize it is very far from clear that any plausible naturalistic account of consciousness can be given.

genuine, they constitute evidence for a supernatural designing intelligence.[26]

3. Conclusion

I have argued that methodological naturalism, contrary its proponents' claim, is not metaphysically neutral. I have further argued that its acceptance as a practice requires justification. Unfortunately for its advocates, attempts to justify it are failures. It cannot be defended as a definition or a self-imposed limitation of science, nor, more modestly, as an inductively justified commitment to natural causes.

Its rejection, however, in no way prohibits scientists from searching for natural causes of physical phenomena. The issue is not whether it is legitimate to look for natural causes of phenomena, but rather whether science must or should in all circumstances confine itself to attempted explanations in terms of natural causes, no matter how inadequate such attempted explanations prove. Whether in science or other endeavors, one needs to be

[26] An anonymous reviewer has objected that "the idea that design can somehow be 'detected' [...] while completely abstracting from the identities and characteristics of particular designers, strikes me [...] as wrongheaded." Three points are in order. First, the reviewer provides no response to the example I give. Second, the recognition of a pattern that is analogous to those produced by intelligent agents such as ourselves—e.g. the recognition that biological entities contain numerous sophisticated machines—is evidence that the designer's nature is not completely foreign to our own. Third, insofar as recognition of biological design might plausibly be thought to provide an argument for God, the theistic hypothesis is not silent concerning God's character. As Draper, no friend to theism, comments,

> moral perfection is built into the theistic hypothesis. Because we are not entirely in the dark about the preferences of such a being (at least other considerations held equal) some facts about nature are more probable on theism than on metaphysical naturalism [...] Furthermore, building moral perfection into the theistic hypothesis does not make that hypothesis ad hoc if [...] God's moral perfection is made likely by other attributes that are plausible attributed to a personal ground of being. (Draper 2005, 295)

free to go where the evidence leads. If the best explanation for some events or structures is direct supernatural agency then science needs to be able to have access to it. Any methodology which precludes such access acts not to further scientific investigation, but rather to place it in an explanatory straitjacket.

References

Boudry, Marteen, Stefaan Blancke, and Johan Braeckman. 2010. "How Not to Attack Intelligent Design Creationism: Philosophical Misconceptions about Methodological Naturalism." *Foundations of Science* 15: 227–44. https://doi.org/10.1007/s10699-010-9178-7

Burtt, Edwin Arthur. 1932. *The Metaphysical Foundations of Modern Physical Science*. 2nd edition, revised. London: Routledge and Kegan Paul.

Craig, William Lane, and James D. Sinclair. 2009. "The Kalam Cosmological Argument." In *The Blackwell Companion to Natural Theology*, edited by William Lane Craig and James Porter Moreland, 101–201. London: Wiley-Blackwell. https://doi.org/10.1002/9781444308334.ch3

Davis, Ted. 2012. "Science and the Bible: Intelligent Design, Part 3." Available at https://biologos.org/blogs/ted-davis-reading-the-book-of-nature/science-and-the-bible-intelligent-design-part-3. [Accessed February 22, 2018.]

Dilley, Stephen. 2010. "Philosophical Naturalism and Methodological Naturalism." *Philosophi Christi* 12 (1): 118–41. https://doi.org/10.5840/pc20101218

Dembski, William. 2004. *The Design Revolution: Answering the Toughest Questions about Intelligent Design*. Downer's Grove, Illinois: InterVarsity Press.

Draper, Paul. 2005. "God, Science, and Naturalism." In *The Oxford Handbook of Philosophy of Religion*, edited by William Wainwright, 272–303. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195331356.003.0012

Earman, John. 2000. *Hume's Abject Failure*. Oxford: Oxford University Press. https://doi.org/10.1093/0195127382.001.0001

Fales, Evan. 2010. *Divine Intervention: Metaphysical and Epistemological Puzzles*. New York: Routledge and Kegan Paul.

Forrest, Barbara. 2000. "Methodological Naturalism and Philosophical Naturalism: Clarifying the Connection." *Philo* 3 (2): 7–29. https://doi.org/10.5840/philo20003213

Jenkins, David. 1987. *God, Miracle, and the Church of England*. London: SCM.

Larmer, Robert. 2003. "Is Methodological Naturalism Question-Begging?" *Philosophia Christi* 5 (1): 113–30. https://doi.org/10.5840/pc2003518

Larmer, Robert. 2014. *The Legitimacy of Miracle*. Lanham, Maryland: Lexington.

Laudan, Larry. 1996. "The Demise of the Demarcation Problem." In *The Demise of the Demarcation Problem*, edited by Michael Ruse, 337–50. Amherst, New York: Prometheus Press.

Lewontin, Richard. 1983. "Introduction." In *Scientists Confront Creationism*, edited by Laurie R. Godfrey, xxiii-xxvi. New York: Norton.

Lewontin, Richard. 1997. "Billions and Billions of Demons." *The New York Review of Books*, January 9, 1997. https://www.nybooks.com/articles/1997/01/09/billions-and-billions-of-demons/. [Accessed June 15, 2017.]

McDonald, Patrick, and Nivaldo Tro. 2009. "In Defense of Methodological Naturalism." *Christian Scholars' Review* 38 (2).

Meyer, Stephen. 2009. *Signature in the Cell*. New York: Harper Collins.

Monton, Bradley. 2013. "An Atheist Defence of Christian Science." *European Journal for Philosophy of Science* 5 (3): 43–54. https://doi.org/10.24204/ejpr.v5i3.218

Numbers, Ronald L. 2003. "Science without God: Natural Laws and Christian Beliefs." In *When Science and Christianity Meet*, edited by David C. Lindberg and Ronald L. Numbers, 265–86. Chicago: University of Chicago Press. https://doi.org/10.1093/acprof:oso/9780195320374.003.0003

Pennock, Robert T. 2001. "Naturalism, Evidence, and Creationism: The Case of Phillip Johnson." In *Intelligent Design, Creationism and Its Critics: Philosophical, Theological, and Scientific Perspectives*, edited by Robert T. Pennock, 77–98. Cambridge, MA: MIT Press.

Pennock, Robert T. 2011. "Can't Philosophers Tell the Difference between Science and Religion? Demarcation Revisited." *Synthese* 178 (2): 177–206. https://doi.org/10.1007/s11229-009-9547-3

Pigliucci, Massimo, and Maarten Boudry, eds. 2013. *Philosophy of Pseudoscience—Reconsidering the Demarcation Problem*. Chicago: University of Chicago Press.

Ratzsch, Del. 2001. *Nature, Design and Science*. Albany: SUNY Press.

Ratzsch, Del. 2002. "Design Theory and Its Critics: Monologues Passing in the Night." *Ars Disputandi* 2: 1–16. https://doi.org/10.1080/15665399.2002.10819755

Ratzsch, Del. 2004. "Natural Theology, Methodological Naturalism and 'Turtles All the Way Down'." *Faith and Philosophy* 21 (4): 436–55. https://doi.org/10.5840/faithphil200421448

Ratzsch, Del. 2005. "Intelligent Design: What Does the History of Science Really Tell Us." In *Scientific Explanation and Religious Belief*, edited by Michael G. Parker and Thomas, M. Schmidt, 126–49. Tübingen, Germany: Mohr Siebeck.

Sober, Elliot. 2011. "Why Methodological Naturalism." In *Biological Evolution: Facts and Theories. A Critical Appraisal 150 Years after The Origin of*

Species, edited by Gennaro Aulette, Marc LeClerc, and Rafael A. Martinez, 359–378. Rome: Gregorian Biblical Press.

Spitzer, Robert J. 2010. *New Proofs for the Existence of God*. Grand Rapids, Michigan: Eerdmans.

Tours, James. 2016. The Origin of Life: An Inside Story. Available at https://www.youtube.com/watch?v=_zQXgJ-dXM4. [Retrieved June 11, 2017.]

Van Till, Howard. 2002. "Is the Creation's Formation Economy Incomplete? A Response to Jay Wesley Richards." *Philosophia Christi* 4 (1): 113–18.

Organon F 26 (1) 2019: 25–39
https://doi.org/10.31577/orgf.2019.26103

ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

RESEARCH ARTICLE

The Implausibility of the Causal Closure of the Physical

Richard Swinburne*

Received: 26 October 2018 / Accepted: 7 November 2018

Abstract: Much recent neuroscientific work, and in particular the programme initiated by Benjamin Libet, seeks to show "the causal closure of the physical"—that mental events never cause physical events, and in particular that our intentions never cause brain events and thereby our intentional bodily actions. But no one is justified in believing any scientific theory unless they are justified in believing that it successfully predicts certain events. Someone is justified in believing that certain events predicted by some theory did occur, if they apparently remember having perceived these events or if some other scientist apparently testifies that they have perceived these events. But we believe our apparent memories of our past perceptions of events because we believe that perceiving those events has caused brain events which have caused our present apparent memories of them; and we believe the apparent testimony of others because we believe that their intentions to testify have caused brain events in them which in turn have caused the words of their testimony to come out of their mouths. So someone could only justifiably believe the theory that mental events never cause physical events if they believe that either their past perceptions or the intentions of other scientists to tell them what they perceived, both of which are kinds of mental events, have caused brain events, which are physical events. So that theory is self-defeating; no one could ever be justified in believing it, or more generally be justified in believing the theory of the causal closure of the physical.

* Oxford University
50, Butler Close, Oxford OX2 6JG, United Kingdom
richard.swinburne@oriel.ox.ac.uk

© The Author. Journal compilation © The Editorial Board, *Organon F*.

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International Public License (CC BY-NC 4.0).

Keywords: Causal closure of the physical; mind-brain causation; physical determinism; Libet.

I begin with some definitions.[1] Although "physical events" are often defined in terms of events involving properties designated by predicates belonging to a true and complete physics, or ones supervenient thereon[2], this is not a very helpful definition since we do know much about what a future true and complete physics would be like. But there is surely one property which all the events with which a future physics would be concerned, would have; that is the property of being publicly accessible events. So I shall define "a physical event" as "an event to which no one person has privileged access." If I can show that the causal closure of the physical is implausible on this definition of physical events, it will surely be implausible on any other currently available definition. I shall contrast a physical event with a "mental event", defined as one to which its subject (the person whose it is) necessarily has privileged access by experiencing it; and "a pure mental event" as a mental event which does not entail the occurrence of a physical event. By a person having privileged access to some event, I mean that necessarily he or she is in a better position than anyone else to know that it is occurring; whatever ways others have of finding out whether or not it is occurring, the subject can also use—but they have an additional way, by experiencing it. Among physical events are brain events; anyone who learns how to use the relevant apparatus can find out as much about my brain events as can I or anyone else. Perceptions such as me seeing a tree are mental events since whatever ways others can use to discover whether or not I am seeing a tree, such as observing my behaviour and studying my brain events, I could also use—I could watch a film of my behaviour and study the brain events just as well as anyone else—but I have an additional way of knowing about my perception, that is by actually experiencing seeing

[1] The ideas contained in this paper were first published in (Swinburne 2011). I am grateful to the editor of the *Journal of consciousness studies* for permission to reuse them. For a more developed presentation of these ideas see my book (Swinburne 2013), especially chapter 4.

[2] For definitions along these lines, see (Montero 2009).

the tree. Perceptions are however not pure mental events, since seeing a tree entails that there is a tree present—and that is a physical event. But me seeming-to-see-a-tree, that is me believing that I am seeing a tree does not entail the existence of the tree.

The class of pure mental events includes all conscious events, that is apparent experiences, events of kinds such that they occur only in so far as we are aware that they are occurring. Conscious events include sensations such as pains and feelings of nausea and patterns of colour in my visual field. They also include occurrent thoughts which cross my mind, such as "today is Thursday;" if I am not in any way aware that I am having a pain or the thought "today is Thursday," I am not in pain and the thought isn't crossing my mind. They also include apparent perceptions—such as my apparent perception of a tree, apparent awareness of one's beliefs and desires, and apparent awareness of truths of reason such as being conscious of "$23\times3 = 69$" as I think about it.

Also among conscious events are intentions. By my intention I mean what I am trying to do, my purpose which I am trying to fulfil. Intentions include short-term intentions, such as intentions to move an arm or utter a sentence; medium-term intentions such as an intention to walk to the railway station or to eat lunch, which are purposes which I am trying to fulfil during the whole of a longer period; and long-term intentions, such as to write a book or pursue some career, which I try to fulfil by actions at different periods over a long stretch of time. If my body performs some movement of a kind which I normally make intentionally, but which on this occasion was simply an unintended reflex, then there was no intention in what I was doing. Other people can reach well justified conclusions about whether I had some intention in making some movement and what that intention was, from a study of my public behaviour (including what I say) and perhaps one day from a study of my brain. Yet whatever ways others use to find out about my intentions, I could also use; I could watch a film of my behaviour and study my brain just as well as anyone else. But I have an additional way, not available to anyone else, of knowing whether I had some intention and what that intention was—by actually experiencing the intention; and necessarily no one else can experience my intentions—for the person who experiences an intention is the person who has the intention.

Pure mental events include, as well as conscious events, what I call continuing mental states, that is events—such as beliefs and desires—which may occur while the subject is not conscious of them but of which she can become conscious by asking herself what she believes or desires, or by being aware of them when she is planning to take them into account in her actions. They are pure mental events because, while the subject can use all the ways available to others for discovering what she believes or desires, she has a way of discovering these states which is not available to others. The suspect who tells the police that he did not commit the crime has a way of knowing whether he is telling them what he believes, which is not available to the police—he is consciously aware of whether his intention is to tell them what he believes.

Our ordinary experience of life seems to show us very strongly that our intentions often cause our bodily movements. Consider first an intention which may or may not succeed in producing the intended effect—for example, suppose that I am trying to lift a weight. I may not succeed in lifting the weight, but if I do succeed, it seems to me strongly that it was me trying which caused the weight to rise, and so that lifting the weight was an intentional action. It seems to me strongly that this is so, because to try to do something just is to do what one believes will make it more likely that that thing will occur. And clearly the weight will not rise unless I try to raise it, and the harder I try the higher it will rise. There is a continuum of actions between actions like this in which I have to try for a short time in order to produce the intended effect, and those bodily actions which are so easy for most of us to perform and take so short a time, that we are seldom aware of a time when we are trying to bring about the effect and have not yet succeeded in doing so. But clearly there is a difference between me intentionally raising my arm and my arm just rising in a way unintended by me. That difference seems to be that me intentionally moving my arm consists in me causing the bodily movement, that is me having the intention to cause the bodily movement causes the bodily movement.

It is a fundamental epistemic principle which I call the Principle of Credulity, that what we seem to (that is, apparently) experience is probably so, and hence we are justified in believing (that is, it is rational for us to believe) that it is probably so—barring counter-evidence. If it seems to me

that I am seeing a tree, then probably I am seeing a tree. If it seems to me that I am hearing you speak, then probably I am hearing you speak—all this in the absence of evidence that I am subject to some illusion. If we could not rely on our apparent experiences of the world in this way—even after we had checked it by looking again or listening for a longer period—we could have no knowledge about the world at all. So since our experience seems to tell us so strongly that our intentions cause our bodily movements, it is—in the absence of counter evidence—very probable that they do. Science has shown us that all our bodily movements are caused more directly by brain events initiating sequences of events in our nerves which cause the bodily movements. So it is very probable that that our intentions cause the bodily movements by causing the brain events which cause those bodily movements, and so that mental events often cause physical events.

The doctrine of 'the causal closure of the physical' (CCP) is the doctrine that physical events are caused only by physical events. It follows from CCP that such common-sense views as that I came to this building because I had the intention to do so, and so the intention caused my leg movements which brought me here are false. Now it is logically possible that such common sense beliefs are simply illusions. But given my previous argument, it would need substantial evidence (that is, counter-evidence to what strongly seems to be the case) to show that they are illusions, and so that CCP is probably true. CCP is an empirical doctrine about which kinds of event cause which other kinds of event, and so a justified belief in CCP requires a justified belief in some scientific theory which entails it. In this paper I argue the epistemological thesis that noone could ever be justified in believing any theory which entails CCP, and that claims that recent neuroscientific work provide that justification are not merely false, but couldn't possibly be true—because of what constitutes a justified belief in a scientific theory. For a justified belief in a scientific theory requires a justified belief that it makes successful predictions, and that means both a justified belief that it predicts certain events and a justified belief that those events occurred. In this paper I will be arguing that (at least one of) those justified beliefs couldn't be had if CCP were true. Hence CCP is in a crucial sense self-defeating; if it were true, we could not be justified in believing it.

So how can anyone have a justified belief that some scientific theory predicts certain events? Scientists in the relevant field will have calculated that it makes these predictions. And if a scientist can hold all the calculations in her mind at one time, it will be for her a deliverance of reason, evident a priori, that the theory does make these predictions. Alas, for any scientific theory of any complexity most experts in any field will be unable to hold in their minds at one time all the relevant calculations; even as the scientist reads through the text of her calculations, she depends on her memory towards the end of the calculations for her belief that the initial calculations were correct. Later in life all that she may remember is that it did seem to her earlier that the theory made those predictions. She may have a diary in which she recorded this, which will be—as it were—her testimony about this to herself and others. Non-scientists and scientists less central in the field will depend on the testimony of those whom they regard as experts, that they have made those calculations. So what makes someone's belief that the theory predicts certain events justified is (if it can be had) experience (of oneself currently 'seeing' that the calculations are correct), memory (of having made calculations in the past), or testimony (from oneself or others that they have made certain calculations); or rather, since all of these sources may be mislead, it is apparent experience, memory, and testimony which provide our justified beliefs that the theory makes true predictions—justified in the absence of counter-evidence, often called "defeaters."

And how can anyone have a justified belief that the events predicted in fact occurred? They will normally depend on the evidence of the same three sources. Certain observers will (apparently) in a wide sense experience these events—that is if they are physical events, they will perceive them, or if they are conscious events they will experience them (in a narrow sense). Later, the observers may (apparently) remember having experienced the events; and others will depend on the (apparent) testimony of observers about these (or the observers may depend on their own apparent written testimony.) Alternatively, a believer may have a justified belief that the events predicted occurred because it is a consequence (deductive or probabilistic) of some other justifiably believed theory that they did. But in that case a justified belief in that other theory would itself depend on the evidence of the same three sources.

When we believe that our apparent experiences are real experiences, as by the principle of credulity we are surely normally justified in doing—in the absence of counter evidence, we believe them because we believe that they are caused by the events apparently experienced. You believe that your apparent experience of observing some distant star through a telescope is a real experience and so that there is really a star there, because you believe that your apparent experience was caused by the star emitting light which impinged on your eyes, and caused the apparent experience of it. If you came to believe that your apparent experience was caused by a malfunctioning of the telescope, and had no other reason for believing that the star was there, you would cease to believe that the star was there. So if your apparent experience is a real experience, then a physical event is causing a mental event; and that of course is not ruled out by CCP.

Our apparent memories of our experiences are themselves apparent experiences of our past experiences, and so again by the principle of credulity we are justified in believing that the apparent memories are real memories in the absence of counter evidence. When we believe them, we believe them because we believe that they are caused by the past experiences. In the case of memories of experiences less than a few minutes earlier, it may be that the past experience causes the present memory of it directly (that is, without causing it by causing some other event which causes the memory). But we know well, as neuroscience has also shown,[3] that in the case of memories of events more than a few minutes ago the causation is indirect; the past experience causes a brain event (a "trace" in the brain) which later causes the memory of it. And so our reliance on apparent memory for our knowledge of past events depends on assuming that a mental event (a past experience) causes a physical event (a brain event).

It also follows from the principle of credulity, that when it seems that someone is telling us something (orally or in writing) we are justified in believing that they are telling us that thing, in the absence of counter evi-

[3] For example, a patient from whom portions of both temporal lobes including their hippocampus were removed, proved unable to recall anything which had happened to him more than a few minutes earlier. For a description of this patient's condition see, among many other places, (Thompson 2000, 392–93).

dence. We assume that their intention to tell us what they apparently remember has caused the words reporting the experience to come out of their mouth or to be typed into their computer. If we came to believe that the words came out of some person's mouth because they suffer from fluid aphasia (which produces streams of unintended words), or that the person's fingers pressing keys on his computer keyboard were caused to do so by a neuroscientist controlling the neural impulses to those fingers, we would no longer believe that testimony. The apparent testimony would not be real testimony. And when we also believe someone's apparent testimony to have experienced some event, we assume that their past experience has caused their apparent memory of it (again, a mental-to-physical causation). And so relying on apparent testimony to a past event involves relying on two separate processes of mental-to-physical causing ('causing' in the sense of the mental event being a necessary part of the total cause of the physical event.)

A scientist takes his (apparent) observations, experiences and calculations as probably correct, at least when he has looked carefully and checked—in the absence of contrary evidence. Almost all scientific knowledge relies on (apparent) memory (e.g. of the results of experiments or calculations only written up the following day). And for all science, we all rely most of the time on the (apparent) testimony (written and spoken) of observers to have made certain observations (or had certain experiences) and of theoreticians to have done certain calculations. And the wider public relies entirely on the (apparent) testimony of scientists with respect both to their calculations and to their observations. A solitary scientist would be justified in believing his apparent perceptions while he was apparently perceiving without assuming that there was any mental-to-physical causation. But no one would be justified in believing a large-scale scientific theory without having a lot of evidence about the occurrence of events at different times and places and having his calculations confirmed by other scientists. So it would seem that we cannot have any significant justification for believing any scientific theory without relying on apparent memory and apparent testimony, and so without relying on processes which consist in mental events causing physical events. Hence it would seem that no one will ever be justified in believing CCP, since the justification would consist in believing in the occurrence of processes which would be ruled out by CCP itself.

There are two different kinds of way in which scientists have attempted or could attempt to establish CCP. The first way is to attempt to show that when mental events occur, they never make any difference to the pattern of later physical events; and the obvious way to set about this would be to show that the most plausible candidate for mental events which make such a difference, that is our intentions, never make any difference to our brain events.

The results of a recent neuroscience research programme, initiated by the work of Benjamin Libet in the 1980's, have been interpreted by many scientists as showing that conscious events (and so 'mental events' generally in my sense) never cause brain events. In the original and most influential Libet experiments[4] subjects were instructed to move their hand at a moment of their choice within a short specified period (for example, a period of 20 seconds). The subjects watched a very fast clock, and reported subsequently the exact time at which they formed their intention to move their hand. They reported the intention to move their hand occurring (on average) 200 msecs (milliseconds) before the time at which experimenters recorded the onset of activity in the subjects' muscles initiating the hand movement. Experiments of other kinds, Libet claimed, showed that subjects report the time of sensations as occurring 50 msecs before the time of the brain events which caused them. That led Libet to hold that subjects misjudge the time of all conscious events by 50 msecs, and so he concluded that their intention occurred (on average) 150 msecs before the muscle activation. However electrodes placed on subjects' scalps recorded (on each occasion of hand moving) a build-up of electric potential, called 'the readiness potential' (RP), which was presumably caused by the occurrence of a particular kind of brain event which occurred (on average) 550 msecs before the muscle activitiy and so 400 msecs before the occurrence of the intention. The last thirty years have seen very considerable progress in understanding the neural basis of intentional actions, made possible by new techniques

[4] For Libet's description of his own work, see (Libet 2004). For accounts and interpretations of the development of this work on the neural basis of intentional actions over the last twenty years, using new methods of discovering what is happening in the brain at different times, see the surveys in (Hallett 2007), (Haggard 2008), and (Banks and Pockett 2007).

which allow neuroscientists to identify far more precisely than by measurements of electric potential on the skull, which areas of the brain are active some exact number of milliseconds before the time at which subjects claim that they form some intention. The results of this work seem initially to give very considerable support to the view that a prior brain event of the kind which gives rise to RP is a necessary condition (although not necessarily a sufficient condition) for the occurrence of a simple intentional bodily movement of the kind studied by Libet.

So, if the subjects' reports are at all accurate, there is this succession of events: a brain event (B_1), followed by a conscious event (the intention represented by 'M_1' in the diagram), and also followed by other brain events (represented by B_2 and B_3) and then (later than M_1) a brain event (represented by 'B_4') which directly causes the muscle activity and so the movement.

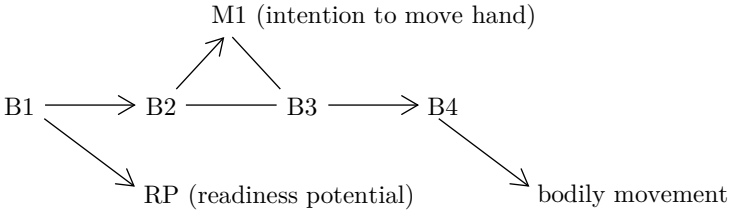

Diagram to illustrate the Libet experiment. 'M' represents a mental event. The 'B's represent brain events; events to the left of the diagram are earlier than events to the right. Arrows represent evident causal action. Lines without an arrow indicate possible causal action.

Many neuroscientists have argued from this kind of evidence to reach the extraordinary conclusion that the intention does not cause the movement; but rather a brain event B2 (caused by B1) causes both the intention (M1) and the subsequent brain event B3 which (via B4) causes the bodily movement, without the intention causing anything at all. Thus one group of neuroscientists concluded that Libet's data "contradict the naïve view of free will—that conscious intention causes action. Clearly conscious intention cannot cause an action if a neural event that precedes and correlates with

the action comes before the conscious intention."[5] But that is a totally unjustified conclusion. One event X may cause another event Z by causing an event Y which causes Z. So it is equally compatible with all the data showing the correlations between an earlier event (B_1), an intention (M_1) and the brain event (B_4) which causes the movement, and also the most natural explanation of those correlations, to suppose that B_1 causes (via some intermediate brain event B_2) the intention (M_1), and that the intention causes (via some intermediate brain event B_3) the brain event (B_4) which directly causes the movement. (By 'causes' I mean 'is a necessary part of the cause of.') To show that the intention was not also a necessary part of the cause, you would need to show that an earlier event B_1 causes the very same sequence of brain events with or without subjects having the requisite intention (to produce that bodily movement). And the Libet programme is nowhere near attempting to show that.

But the major problem with Libet-type experiments is that, in order to show that or anything else, the programme depends on the apparent testimony of subjects for information about the time of the occurrence of their mental events. And so all Libet-type experiments designed to show that mental events do not cause physical events require the experimenter to assume that sometimes mental events do cause physical events. Such experiments might serve to show that certain kinds of bodily movement (for example, those studied in Libet-type experiments, or—more obviously—some very quick reactions such as jumping out of the way of a passing car) which we might have regarded as formed by our intentions are in fact not formed by our intentions. But they could only show this if we assume that other bodily movements (such as the movements of our lips in telling others about our intentions) are formed by those intentions. Clearly the same general objection applies to any attempt to establish CCP, which relies on evidence about which mental events occur when. We could only get enough evidence to show the theory to be true by relying on the apparent testimony of other people, and that involves assuming that that theory is false.

But there is another way by which it might seem that scientists could establish CCP, and that is on the basis of evidence merely about which

[5] (Roediger, Goode, and Zaromb 2008). For similar quotations from other neuroscientists see (Mele 2009, 70–73).

physical events occur when, without relying on any evidence about which mental events occur when. Could not observation merely of physical events provide us with a justified belief in some very general deterministic physical theory from which it follows that every physical event has another physical event as an immediately prior necessary and sufficient cause? If so, it would follow that no brain event could have a mental event as a necessary part of its cause. For if a deterministic physical theory of this kind were true, the physical event would have occurred anyway; whatever mental events the subject had could have made no difference to the brain events. Suppose that scientists discover that for each physical event F studied in large samples of many different kinds of physical events (including apparently intentional bodily movements), there is some other immediately prior physical event E which a certain deterministic theory claims will have F as its immediate effect. That would seem to be powerful evidence in favour of that theory and so in favour of CCP.

The problem however remains that our evidence about which physical events occur when in these large samples of events must come from the reports of scientists about the results of their experiments, that is from the apparent testimonies and memories of scientists about what they claim to have perceived. But we have seen that we are only justified in relying on apparent memory and testimony if we are justified in assuming that past observations cause brain traces which cause present memories and that people's intentions cause words to come out of their mouths. So again relying on the evidence of the apparent memory and testimony of scientists about what they have observed already seems to presuppose the falsity of CCP.

It is however possible to avoid this objection, as an objection to this second way of trying to establish CCP, by modifying our understanding of 'memory' and 'testimony,' in such a way that we could be said to 'remember' some past physical event without assuming that the memory involves remembering a past experience of that event; and we could be said to give 'testimony' to the occurrence of a past event without assuming that that involves testifying to a past experience of that event. We could understand 'memory' simply as memory of the occurrence of events, and not only of events which are experiences of the occurrence of other events. A subject

could be said to 'remember' past physical events in virtue of those events causing traces in his brain, which at a later time cause the apparent memory of those events without any mental-to-physical causation being involved. The memory traces caused in a person's brain by the occurrence of some event sometimes cause that person to become aware later of details of that event of which they were not at the time aware; and it does not seem too unnatural a use of the word 'remember' to say that they 'remembered' those details. So the only assumption that someone would need to make order to rely on his 'apparent memory' of some past physical event, would be that the past event caused his present belief that it happened; and that does not involve any mental-to-physical causation. And we could come to understand 'testimony' to amount merely to the public utterance of sentences reporting that an event occurred caused by a chain of events in the utterer, itself caused by the event reported, a chain which need not include any conscious events. The 'testimony' would not be testimony that the testifier had observed the events, but merely testimony that the events had occurred. Thus someone's eyes could receive light rays from physical events, and—because those physical events caused brain events in that person—subsequently report them, without that causal chain proceeding through any conscious events. So the only assumption that someone A would need to make in order to rely on B's 'apparent testimony' that some past event E occurred, would be that a past event caused a brain event in B which caused B's mouth to produce words stating that E occurred ('caused' in the sense that it was a necessary part of a sufficient cause); and that does not involve any mental-to-physical causation. Counting as a witness's 'testimony' to an event any utterance of his claiming that the event occurred which was caused by the event, certainly seems to involve giving a stretched meaning to 'testimony.' We might reasonably want to claim that science ought to rely on the reports of observers about their own experiences, rather than merely on words coming out of their mouths (whose content is not further justifiable) that certain events affected their brains. Still, it is worthwhile to see if we can save CCP by stretching our understanding of 'memory' and 'testimony' in these ways. We would not then need to make any assumptions about conscious events causing physical events in order justifiably to believe what we learn from 'memory' and 'testimony' about the occurrence of physical events.

And in that case the evidence that events occurred which were in fact predicted by some physical theory entailing CCP would not be undermined by the use of that evidence presupposing the falsity of the physical theory.

There is however a further (and to my mind insuperable) difficulty in supposing that we could have a justified belief that some deterministic physical theory gave true predictions about relations between physical events. This is that we would also need, not merely a justified belief, that certain relations between physical events occurred, but also a justified belief that these relations were predicted by that deterministic theory. But anyone who had not calculated for himself what that theory predicted about the relations between physical events must depend on the evidence provided by the apparent testimony of scientists to have calculated this and 'to have seen' (that is, had a conscious belief) that that was what the theory predicted, that is evidence of the conscious events of scientists, not merely evidence of the occurrence of physical events. But if the deterministic physical theory were true, the scientists could not have been caused to give that testimony by their intention to tell what they believed about what the theory predicted. Hence no one (and so no other scientist) could justifiably believe what any scientist reported about his calculations, and so believe that the theory made the predictions which he claimed that it did (as well as believing that the predicted events occurred), since believing what the scientist reported would undermine the credibility of his apparent testimony to it. Neither—for the same reason—could any scientist rely on his own testimony to himself recorded in a diary that he had previously calculated the consequences of the deterministic theory. Nor could a scientist rely on his own apparent memory of having calculated these consequences, since this would involve his past experiences (of his calculation) causing the brain event which caused his memory. Only if a scientist could hold in his mind at one time all his calculations (or perhaps do all the calculations within less than a few minutes), from which it apparently followed that the deterministic theory predicted certain events, could he have a justified belief that that theory made successful predictions, and so a justified belief in CCP. For most scientific theories and most scientists, this is most unlikely. The crucial difference between having a justified belief that certain physical events occurred, and having a justified belief that a certain theory predicts certain

events, is that for the latter we need evidence of conscious events (we or others 'seeing' that certain calculations are correct), while for the former (given extended senses of 'memory' and 'testimony') we need only evidence of the occurrence of physical events.

I conclude that (with the above very small exception) no one could have a justified belief that any deterministic physical theory made certain predictions. So neither by the route of trying to show the inefficacy of mental events, nor by the route of trying to show the total efficacy of physical events, could anyone (with the stated very small exception) have a justified belief in CCP. Hence we should believe that things are they seem to be, that often our intentions do cause our bodily movements, which clearly they do by causing our brain events.

References

Banks, William P., and Susan Pockett. 2007. "Benjamin Libet's work on the Neuroscience of Free Will." In *The Blackwell Companion to Consciousness*, edited by Max Velmans and Susan Schneider, 657–70. Oxford: Blackwell Publishing. https://doi.org/10.1002/9780470751466.ch52

Haggard, Patrick. 2008. "Human Volition: Towards a Neuroscience of Will." *Nature Reviews*, *Neuroscience* 9: 934–46. https://doi.org/10.1038/nrn2497

Hallett, Mark. 2007. "Volitional Control of Movement: The Physiology of Free Will." *Clinical Neurophysiology* 118 (6): 1179–92. https://doi.org/10.1016/j.clinph.2007.03.019

Libet, Benjamin. 2004. *Mind Time*. Cambridge, MA: Harvard University Press.

Mele, Alfred R. 2009. *Effective Intentions*. Oxford: Oxford University Press.

Montero, Barbara. 2009. "What is the Physical?" In *The Oxford Handbook of the Philosophy of Mind*, edited by Brian McLaughlin, Ansgar Beckermann, and Sven Walter, 173–88. Oxford: Oxford University Press.

Roediger, Henry L., Michael K. Goode, and Franklin M. Zaromb. 2008. "Free Will and the Control of Action." In *Are We Free?*, edited by John Baer, James C. Kaufman, and Roy F. Baumeister, 205–25. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195189636.003.0010

Swinburne, Richard. 2011. "Could Anyone Justifiably Believe Epiphenomenalism?" *Journal of Consciousness Studies* 18 (3–4): 196–216.

Swinburne, Richard. 2013. *Mind. Brain, and Free Will*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199662562.001.0001

Thompson, Richard F. 2000. *The Brain*. 3rd edition Worth Publishers.

Organon F 26 (1) 2019: 40–61
https://doi.org/10.31577/orgf.2019.26104
ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

RESEARCH ARTICLE

The Principle of the Causal Openness of the Physical

Daniel von Wachter*

Received: 16 July 2018 / Accepted: 3 December 2018

Abstract: The argument from causal closure for physicalism requires the principle that a physical event can only occur through being necessitated by antecedent physical events. This article proposes a view of the causal structure of the world that claims not only that an event need not be necessitated by antecedent events, but that an event cannot be necessitated by antecedent events. All events are open to counteraction. In order to spell out various kinds of counteraction I introduce the idea of 'directedness.'

Keywords: Causal closure; determinism; free will; miracles.

1. Introduction: the principle of causal closure

(1.1) The Principle of Causal Closure (PCC) expresses a substantial conviction about causality and about everything that happens. Contrary to PCC, in this article I shall defend the 'Principle of the Causal Openness of the Physical' (PCO). I shall proceed by addressing the following points:

- Note how influential PCC has been.
- Which version of PCC is required for an argument against the existence of physical events that have a non-physical cause?

* International Academy of Philosophy in the Principality of Liechtenstein
✎ International Academy of Philosophy in the Principality of Liechtenstein, Fürst-Franz-Josef-Str. 19, FL-9493 Mauren, Liechtenstein
✉ epost@von-wachter.de Website: http://von-wachter.de

© The Author. Journal compilation © The Editorial Board, *Organon F*.

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International Public License (CC BY-NC 4.0).

- State the Principle of Causal Openness.
- Relate PCO to the phenomenon of superposition.
- Introduce the idea of directedness and relate it to PCO.

(1.2) Many believe that PCC constitutes a strong reason for believing that there are no souls interacting with the body and that there are no divine interventions. Thus David Papineau:

> Over the latter half of the last century English-speaking philosophy became increasingly committed to naturalistic doctrines. Much of this naturalistic turn can be attributed to the widespread acceptance of the thesis that the physical realm is causally closed. (Papineau 2009, 53)

> [I]t was the empirical evidence for causal closure that persuaded philosophers to be physicalists. Once mid-century physiological research had established that all physical effects had physical causes, even in bodies and brains, philosophers quickly figured out that general physicalism followed. (B. G. Montero and Papineau 2016, 188)

Jaegwon Kim formulated PCC in the following way, which some call 'Two-way Causal Closure' (TCC) (Montero 2003, 175):

> If you pick any physical event and trace out its causal ancestry or posterity, that will never take you outside the physical domain. That is, no causal chain will ever cross the boundary between the physical and the nonphysical. (Kim 1998, 40)

(1.3) Unless the words 'never' and 'will ever' are taken to include possible worlds or the word 'cannot' is added, this formulation says only:

> *No past and future physical events have nonphysical causes.*

Let us call this nonmodal version of PCC 'NMPCC.'

Can NMPCC be used for an argument for physicalism? No, it requires a modal version. Imagine that there are no nonphysical objects. In that case NMPCC is true. But how can one then know that NMPCC is true? By investigating the evidence for the existence of souls and of God and by finding that there is little evidence or by finding that there is evidence

against the existence of souls or God. By this investigation one can know the truth of NMPCC, but NMPCC provides no reason for believing that no physical events have nonphysical causes. There is no argument from NMPCC to physicalism here, because the belief in physicalism justifies the belief in NMPCC, but the belief in NMPCC cannot be justified without belief in physicalism.

(1.4) The physicalist who wants to use PCC in defence of his view needs a modal version of PCC. And indeed that is what Kim and the other proponents of PCC mean. In another book Kim uses the formulation quoted above but adds: 'If x is a physical event and y is a cause or effect of x, then y too must be a physical event' (Kim 1996, 147). In a more recent book he formulates PCC as 'If a physical event has a cause at t, then it has a physical cause at t'[1] and elucidates this by saying: 'There can be no causal influences injected into the physical domain from outside' (Kim 2005, 16).

(1.5) Papineau states PCC using the notion of a 'sufficient cause.'

> *Every physical effect has a sufficient physical cause.* (Papineau 2009, 53)

This is nearly identical to determinism, which is usually defined as the claim that 'Every event has a preceding sufficient cause.' 'Sufficient' is usually taken to be equivalent to 'necessitating,' so that Papinau's PCC amounts to Thomas Hobbes's formulation of determinism: 'All the effects [events] that have been, or shall be produced, have their necessity in things antecedent.'[2]

(1.6) What does the term 'sufficient cause' in Papineau's PCC mean? If it meant only that there is all that is required[3] to cause the event, then

---

[1] This version of PCC 'does not by itself exclude nonphysical causes [...] of physical events' but only if *the principle of causal exclusion* is added: 'If an event e has a sufficient cause c at t, no event at t distinct from c can be a cause of a (unless this is a genuine case of causal overdetermination)' (Kim 2005, 17).

[2] '[...] quaecunque producta vel erunt vel fuerunt, necessitatem suam in rebus antecedentibus habuisse' (*De corpore*, 9.5).

[3] John Bramhall (1655, 172) criticised Hobbes for confusing 'sufficient' in the sense of 'enough' with 'necessitating.' Christian August Crusius (1743) objected to Leib-

Papineau could simply say 'Every physical effect has only physical causes' or 'Every physical effect has only non-probabilistic physical causes' (although this is in conflict with quantum mechanics). This statement can be justified only by defeating all the argunents for the existence of nonphysical objects, that is, it cannot be justified without presupposing physicalism.

The defenders of PCC who use PCC for an argument against interactionist dualism must mean by a 'sufficient physical cause' 'a physical state of affairs that alone suffices to determine the effect.'[4] So they use 'sufficient cause' equivalently to 'necessitating' or 'determining.' Some philosophers, in particular Leibniz, believed that it is a metaphysical principle that everything that exists has a determining reason; one that entails or necessitates 'why it should be so and not otherwise.'[5] They believed that it is impossible that something exists that was not determined. PCC makes this more specific, claiming that an event has to have a determining preceding cause. Such a cause makes it impossible for an immaterial object to prevent the effect, and an event that had such a cause could not have been prevented by an immaterial object once the cause had occurred. Determining causes exclude interventions. That is the idea behind the argument from PCC. It is not just the claim that there are no nonphysical causes of physical events but the claim that the physical is causally 'closed.'[6]

(1.7) Belief in PCC has been influential for longer than, as Papineau suggests, since 1950. Consider Hobbes's line of reasoning (stated in my words):[7] If an event occurs, then its cause was complete ('entire') and sufficient, otherwise it would not have occurred. If a cause is complete, then the effect has to follow, otherwise the cause is not complete. A complete

niz's 'Principle of Sufficient Reason' that Leibniz should have called it 'the principle of the determining reason.'

[4] (B. Montero 2003, 174). Menzies (2015) argues that the argument from PCC is *invalid* if 'causal sufficency' is interpreted as 'nomological sufficiency' and 'causation' as 'counterfactual difference-making.'

[5] Leibniz, *Monadology*, § 32.

[6] Various versions of PCC are discussed by Gibb (2015, 628), Lowe (2008, chap. 2), and B. Montero (2003, 174).

[7] (Hobbes 1655, chap. 9), similarly (Hobbes 1654).

cause is a sufficient cause, and a sufficient cause cannot but produce the effect; the effect cannot but follow it. Hence, *every event is necessitated by preceding events.* Hobbes applied this to free will and concluded:

> [T]hat ordinary definition of a free agent, namely, that a free agent is that, which, when all things are present which are needful to produce the effect, can nevertheless not produce it, implies a contradiction, and is nonsense; being as much as to say, the cause may be sufficient, that is to say, necessary, and yet the effect shall not follow. (Hobbes 1654, § 32)

Determinism was extremely influential in Western Europe through Hobbes, Leibniz, and Christian von Wolff. Although Immanuel Kant did not affirm it for the 'things in themselves' because he considered free will to be undeniable, he spread determinism by claiming that the 'principle of causality' is *a priori*, it is a principle of thought that cannot be denied:

> Everything that happens is at all times determined before through a cause according to constant laws.[8]

Although there have always been critics of determinism, these are much less known today than the determinists. Especially in the 19th century in Germany, many considered determinism to be the hallmark of rationality and of science. Those who endorsed it were declared to be part of the 'Enlightenment.' From around 1780 onwards in Germany some popular writers declared Wolff and Kant to be representatives of 'the Enlightenment,' Hobbes and Leibniz were considered to be precursors of the Enlightenment. To those who accepted determinism, the human soul interacting with the body, free will, and divine interventions, which had previously been accepted by most European thinkers, were considered to be not rational and not up to date. Belief in determinism led many Christian theologians and churches to give up belief in miracles. Thus determinism and PCC is a philosophical doctrine that changed the world.

(1.8) Let us formulate PCC in a way that is clear and serves the purpose of the physicalist:

[8] 'Alles, was geschieht, [ist] jederzeit durch eine Ursache nach beständigen Gesetzen vorher bestimmt.' (*Prolegomena*, § 15)

> *There cannot be a physical event that is not necessitated by preceding physical events. (NPCC)*

In other words, a physical event can only occur by being necessitated by preceding physical events. That event S_2 at time t_2 was *necessitated* by event S_1 at time t_1 means that it was impossible that S_1 occurs (i.e. an event happens that is exactly like S_1) but S_2 does not. Unless indicated otherwise, I shall mean by 'PCC' always 'NPCC.' NPCC is what the defenders of PCC mean or should mean, because the nonmodal versions of PCC can only be justified by an investigation of the evidence for nonphysical objects.

(1.9) If NPCC is true, we might well be able to know it and use it as a powerful argument against interactionist dualism and against the existence of divine interventions. We know some modal truths through insight, through thinking about our experience, or through some special kind of experience. This way we can then know something about the individual cases about which the modal truth says something without investigating them individually. For example we know that it is impossible that x has a cause that is later than x. This way we know that there are no cases of backward causation without having to investigate any evidence for alleged cases of backward causation.

If we could know NPCC, that would be a great philosophical success. It would mean not only that we can know something as substantial as that through philosophical reasoning, but that we could know it without considering the evidence. It would mean that we can know that we have no soul (that interacts with the body) without considering the evidence and the various arguments for and against it, and that we can know that there are no divine interventions without considering the evidence for the various divine interventions in which people believe. It would be a shortcut argument, a knock-out argument. While Papineau thinks that PCC is supported through physics, Hobbes and Kant believed that determinism can be known a priori, just by thinking about how an event can occur and about causality.

(1.10) I shall now propose a principle, which I call the Principle of the Causal Openness of the physical (PCO), which contradicts PCC. Causality as we know it is not as PCC describes it. The physical is causally open to non-physical intervention.

2. Causal openness

(2.1) The principle which I want to defend in opposition to PCC and which I call the *Principle of the Causal Openness of the Physical* (PCO) is:

> *Physical causes are open to cooperation and to intervention.*

(2.2) Openness to cooperation: Each event is so that there can be an additional event in conjunction with which it can cause something different. More precisely:

> *If event S_1 (at time t_1) was the complete physical cause of event S_2 (at time t_2), then in addition to S_1 an event at t_1 could have existed in cooperation with which S_1 could have caused an event at t_2 that does not contain S_2.*

(2.3) Openness to intervention: Arguments from PCC against non-physical causes assume that a complete physical cause necessitates its effect and therefore excludes that something non-physical impedes the causing. Contrary to this, PCO states that all causes are liable to intervention:

> *Even if event S_1 (at time t_1) was the complete physical cause of event S_2 (at time t_2), then it was possible that S_1 occurs but something prevents S_2.*

Cooperation and intervention are two ways how an event can be prevented from causing what it would have caused on its own, had nothing else acted on it or on what follows. They are two ways of counteraction. Cooperation is done by events that are synchronical with the event in question. Intervention is done at some time later, and it can be done by either a physical event or by whatever other kinds of things exist.

(2.4) One important negative implication of PCO is:

> *No event can necessitate a later event.*[9]

[9] I have defended this claim also in (Wachter 2009, § 5.10) and (Wachter 2012). That causes do not necessitate their effects is also argued by Mumford and Anjum (2011, chap. 3). Gibb (2015, § 2) explains why according to the powers theory of causation not all causes are sufficient. This is a weaker claim than mine.

Thus, not even non-probabilistic event causes necessitate their effects. By 'later' I mean that the beginning of the second event is no earlier than the end of the first event. If we are speaking about events occurring at certain points of time, then I mean that the second event occurs at a later point of time.

(2.5) I am assuming in this article that the laws of nature do not or cannot change. So the reason wherefore I claim that no event necessitates a later event is not that the laws can change. Further, I am assuming that if there is a God, he continues to sustain things with their causal powers in being. So the reason wherefore I claim that no event necessitates a later event is not that God may stop sustaining things.

(2.6) Applied to processes, PCO implies:

Causal processes are stoppable.

(2.7) Let me elucidate the counterfactual and modal claims that I have used. By saying that 'event x necessitated event y' I mean: it could not have happened that although x occurs y does not. The occurrence of x is incompatible with the non-occurrence of y. Or: the scenario in which x (or an event exactly like x) occurs but not y is not possible. Speaking more extensionally: In every possible world in which x (or an event exactly like x) occurs y occurs too.

(2.8) By saying 'it was possible that S_1 occurs but something prevents S_2' I mean that the occurrence of S_1 is compatible with something preventing S_2 and thus S_2 not occurring. Extensionally speaking: In some possible worlds, S_1 (or an event exactly like S_1 in the actual world) occurs but S_2 does not because something prevents it.

(2.9) I am assuming that being possible is something different from being consistent (i.e. being not self-contradictory), and being necessary is something different from being analytic. When I say 'x necessitated y' I do not mean that the description of x is contradictory to the negation of the description of y. For every description of something there is the question whether the existence of something to which the description applies is possible. For self-contradictory descriptions, like 'a married bachelor,' this question does not arise. Modal questions arise with respect to consistent descriptions. That the predicates P and Q are contradictory means that in

the relevant language the meanings of 'P' and 'Q' are such that 'P' is used in order to say about something, among other things, that it is not Q (or vice versa). That the existence of something that is P and Q is impossible means that, while 'P' and 'Q' are not contradictory, the properties to which 'P' and 'Q' refer cannot be had by one thing at the same time; they are not combinable. We have the two different words 'contradictory' and 'impossible' for a good reason. They refer to very different phenomena.[10]

For our case here that means that on the one hand I am not saying that, for example, 'S_1 necessitates S_2' is contradictory, and on the other hand I am not just saying that 'S_2 was prevented although S_1 occurred' is consistent.

(2.10) As PCC is used against substance dualism, I want to specify what an interaction between the soul and the body and thus a violation of PCC would be like. In my view, in an action (in my view not just in a free action but in any action) the agent brings about an event directly in the sense that it is not the result of a causal process and thus has no preceding cause but occurs through an action. I call such an event a *choice event*. This bringing about directly is not a kind of process causation, it is a phenomenon *sui generis*. It is not obvious that only souls or God can bring about choice events. Someone who holds that there are no souls could hold that material persons can bring about choice events. But if there are souls or bodiless persons, then they can bring about choice events. I am inclined to think that this is the only way immaterial objects cause physical events. The soul can bring about a property change (perhaps the firing of a neuron) in the brain which has no preceding cause but is due to the soul. Similarly, God, while he is sustaining the universe, might bring about events in the universe directly so that they have no preceding complete cause. Perhaps only he can bring about physical events which have no material cause at all. He can bring about as choice events in the physical world not only property changes, but he can bring about entirely new things. The relevance of all this for this article is that PCO implies that the physical world is open to choice events.

[10] For a defence of this view of possibility see (Wachter 2009, chap. 3).

3. Superposition

(3.1) Before I present in defence of PCO a more detailed account of causal processes, I want to point out that PCO is suggested by the phenomenon of superposition. Newtonian forces can be superpositioned and can be calculated by vector addition. Every force can be combined with other forces. Every force can be counteracted by another force. If various forces are acting on a body, then they have the same effect as if only the force that is their sum acted on the body.

(3.2) If body B with mass m is exerting a gravitational force F_B on body A, which also has mass m, and nothing else is acting on the body, then A accelerates with $a = F_B/m$. But if there are more forces acting on A, then it moves in a different way. For example, there could be a body opposite of C with mass m so that A does not accelerate at all because the gravitational forces exerted by B and C cancel out each other. F_B contributes to how it moves, but it does not determine how A moves.

(3.3) If body A was accelerated by body B through the gravitational force F_B, then A's movement at time t_2 was caused partly by B at t_1. Let us call the effect, A's movement at time t_2, 'S_2' and its complete cause at t_1 'S_1.' S_1 led through a causal process to S_2. Had there been an additional force acting on A, A would have moved differently at t_2 and already right after t_1. Hence *S_1, the complete cause, caused but did not determine (necessitate) S_2*. That is, it could have happened that although S_1 occurs something prevents S_2. This is what PCO says about this case. Not only something after t_1 and before t_2 could prevent S_2 but also something at t_1 could prevent S_2, by forming a cause together with S_1 which then causes something other than S_2. States of affairs can be counteracted by cooperating states of affairs.

It is true that the following was impossible: S_1 occurs, S_2 does not occur, and nothing prevented S_2. It is also true that 'given that S_1 occurred and nothing prevented S_2, S_2 had to occur.' But the sentence 'S_1 determined S_2' cannot reasonably be understood as saying just this. Note that this true statement, 'Given that S_1 occurred and nothing prevented S_2, S_2 had to occur,' is of no use to an argument against interactionist dualism or against divine interventions, because it leaves open that a soul or God could prevent

S_2. The sentence 'S_1 determined S_2' in the sense which is required for an argument against interactionist dualism is false.

(3.4) Someone might want to reply: 'S_1 determined S_2 *given the circumstances*.' The 'circumstances' should refer to certain states of affairs that obtained at t_1 in addition to S_1. However much you include in these circumstances, it will always be true that the state of affairs that consists of S_1 plus the circumstances did not necessitate S_2, because the state of affairs that consists of S_1 and the circumstances is compatible with the existence of some further state of affairs which prevents S_2. If one replaces 'given the circumstances' by 'given that nothing else was acting on S_1' (I call that the 'no-further-causes clause'), it is still not true that S_1 determined S_2, in the sense of 'S_1 necessitated S_2,' because S_1 is compatible with the existence of something that prevents S_2. What is true is that the following is impossible: S_1 occurs, and S_2 does not occur, and nothing prevented S_2.

(3.5) Finally, someone might want to hold that 'It is impossible that $\vec{F} = \sum_{i=1}^{n} \vec{F}_i$ is the resulting force acting on A at t_1 but A does not accelerate (at and after t_1) with $a = F/m$.' If 'force' here were be understood as just any action on A, then the statement would be true. But if 'force' refers just to forces in the sense used in physics—that is, to forces which obey $a = Fm$, which are exerted by physical objects, and which are described by laws of nature—then it does not include the action by souls or by God. With that the statement would have the implications that the physicalist desires, but it cannot be justified and is false, because an agent can counteract the resulting force. As mentioned above, in an action the agent brings about an event directly in the sense that it has no preceding cause but is due to the agent. It is to be investigated whether such events exist, but I see no reason for claiming that they are impossible. In any case, even if such actions on physical objects were impossible, that would not be because somehow the physical is closed or because physical causes exclude the intervention by immaterial agents. So this would not support the argument from causal closure. Physical events are open to cooperation and to counteraction.

4. Directions

(4.1) Now I want to introduce some ideas and concepts with which we can understand and describe causal openness better. Imagine a rolling billiard ball A hitting billiard ball B at time t_2. A's hitting B causes B's movement at a certain time t_3 after t_2. That is a clear example of causation, more specifically of 'event causation' or, as I prefer to call it, of 'process causation.' It is also quite clear that it is true to say that A's movement at a certain time t_1 before t_2 caused B's movement at a certain time t_3 after t_2. A's rolling between t_1 and t_2 is a clear example of a *causal process*.

(4.2) Allow me to comment briefly on the method of our philosophical investigation. While some assume that in philosophy we investigate only concepts, I now invite you to think not about concepts, nor about possible counter-examples to certain definitions, nor about the meaning of 'cause' or 'causal process,' but about typical cases of event causation and causal processes and about what is going on in them. It seems to me that philosophers when they investigate causation are often misled because they think about the concept of causation instead of thinking about what a typical case of causation is like. The concept of a cause subsumes a variety of cases. The assumption that all concepts have a definition of the standard form, consisting of 'necessary and sufficient conditions,' is questionable. At any rate, the task of finding the correct definition of causation which captures all those and only those cases, discussing all those peculiar counter-examples that have been advanced in the literature against the various definitions that have been proposed, is rather different from the task of describing certain cases of causation. I ask the reader now to think not about concepts, not about possible counter-examples to certain definitions nor about what we mean by 'cause' or by 'causal process,' but about what typical cases of causation and of causal processes are like and what is going on there. Causal processes are a phenomenon that we find in the world. What they are we can know only through experience and through thinking about causal processes.

(4.3) Imagine that at time t_2 ball A hits ball B. Through the collision the process of the rolling of A from t_2 does not continue in the direction into which it was heading before t_2. In this sense it was stopped. We can

see all the time that causal processes are stopped. What follows is obvious but was not recognised by Thomas Hobbes (see above pp. 43–44): *At least some causal processes can be stopped.* They are stoppable.

(4.4) *All* physical causal processes are stoppable; non-stoppable processes, as posited by Hobbes, are impossible. We will consider in more detail what the stopping can consist in, but here is a first consideration in defence of this: Small and medium-sized processes can be stopped by processes of the same kind. A billiard ball can be stopped by a billiard ball, a tidal wave can be stopped by a tidal wave, a planet can be stopped by a body of a similar size. Is also the process of the whole universe stoppable? A bigger process requires something bigger, more powerful to stop it, but the size of a process does not influence its modal properties. If small physical processes, like a rolling billiard ball, are stoppable, then big ones of the same kind are also stoppable. Of course, if there is nothing besides the universe, then there is nothing that could stop the process that is constituted by the universe, but nevertheless it is stoppable. The statement 'There is nothing besides the universe' implies that there is nothing that could stop it, but the things and events that constitute the process of the universe are compatible with the existence of additional objects, either material ones or immaterial ones, which, if powerful enough, could stop the universe.

(4.5) A and B can also be considered together as constituting one process. The process that at t_1 includes A and B is not stopped at t_2, but it contains two processes that collide at t_2 and are thereby stopped. We can draw the boundaries of processes in various ways. What is not arbitrary though is what belongs to the process at t_1 which led to the rolling ball A at t_2. For a certain event S_2 at t_2 it is a fact of the matter what at t_1 belonged to the event S_1 which was the stage of the process that led to S_2 at t_2 and thus is at t_1 the complete cause of S_2.

(4.6) There are connections between the stages of a causal process. Its stages are not as unconnected as the Moon at t_1 is unconnected to the Pleiades at t_2. Rather, the process has a *direction*. It cannot stop without something stopping it. It cannot stop or change its direction by chance.

(4.7) The process has a direction at each of its stages. A stage of a process we can call an 'event' or a 'state of affairs.' I often use the term 'event' instead of 'state of affairs' just because it is shorter. A stage of a process

consists of certain properties being at certain things or places at a certain time. So we specify an event or state of affairs by saying which properties at which place or thing at which time we mean. So the event S_1, which is the stage of the process at t_1 (or at an interval ending at t_1), has a direction. It is directed towards S_2 at t_2. We can call the state of affairs in virtue of which the direction exist its 'basis.' I use the terms 'direction' and 'directedness' equivalently. Thus we can say:

> *S_1 at time t_1 is the basis of a directedness*[11] *towards S_2 at time t_2.*

If S_1 is the basis of a directedness towards S_2 at t_2, then S_1 is also the basis of a directedness towards certain events at times between t_1 and t_2. Two directednesses with the same basis are necessarily pointing in the same direction. We can call S_1's directedness towards S_2 at t_2 'base-identical' to S_1's directedness towards a certain event S' at a time t' that is between t_1 and t_2.

(4.8) Saying that 'the directedness based on S_1 at t_1 towards S_2 at t_2 was *realised*' means that things followed that directedness until S_2 and that thus S_2 occurred. S_1 was then at t_1 the *complete* cause of S_2. S_1 led to S_2. S_2 was a result of the process coming from S_1. S_1 was a complete stage of the process coming from S_1. A causal process is the realization of a directedness.

(4.9) A causal process is stopped if something brings about an event that is incompatible with an event towards which the process was directed. By calling two directions 'conflicting' I mean that they are directed towards incompatible events at some time. That two directions are non-conflicting means that the events towards which they are directed are never incompatible. That two directions are non-conflicting *until time t* means that they are not directed towards incompatible events before t. A *resulting* directedness is one that is for some time not conflicting with another directedness.

---

[11] Instead of the term 'directedness' in earlier publications I used the term 'tendency' and called the theory the 'tendency theory' (Wachter 2009, 2003). I am now using the term 'directedness,' although it is more clumsy, because I found that most readers associate 'tendency' with being probabilistic.

(4.10) As we could take any portion of matter to be a 'thing,' we could take any properties at any locations at one time to be an event. Every physical event is the basis of a directedness. Any number of physical events at the same time can be taken to be one event. The whole event as well as the composite events each have their direction. The directednesses of the composite events constitute the directedness of the whole event. As long as constituting non-overlapping events are non-conflicting, the direction of the whole event is the sum of the directions of the constituting events. If an event is constituted by events whose directions are conflicting, then one of the directions could trump the other, or together they constitute a new direction.

(4.11) Now we can spell out how an effect can be prevented, as mentioned in this sibling of PCO: 'If event S_1 (at time t_1) was the complete physical cause of event S_2 (at time t_2), then it was possible that S_1 occurs but something prevents S_2.' There are at least three ways how it could be that A occurs but B does not:

1. By cooperation: There could exist at t_1 an event which together with S_1 is not directed towards S_2.

2. By intervention through a physical process: It is compatible with the occurrence of S_1 that there is a physical process which is directed towards an event at some time between t_1 and t_2 which is incompatible with an event towards which S_1 was directed. This would prevent S_2.

3. By intervention through a person: S_1 does not exclude that there is a person who, at some time between t_1 and t_2, brings about a choice event that is incompatible with an event towards which S_1 was directed. This would prevent S_2.

(4.12) Thus far I have introduced only non-probabilistic directednesses, which we can also call 'unambiguous directednesses.' A non-probabilistic directedness is one that necessarily will be realised if nothing counteracts. In other words, one for which the following is impossible: it is not realised although nothing counteracts. A probabilistic directedness, on the other hand, is one that can fail to be realised just by chance; that is, one for which it is possible that it is not realised although there is nothing that counteracts.

A probabilistic directedness can have a strength that can be described with a number between 0 and 1, corresponding to the probability of the realisation if nothing else is acting on the object. So that a directedness has strength 1 does not mean that it is necessarily realised, but only that is realised if nothing counteracts. Probabilistic directednesses can be disjunctive, that is, they can be towards either K or, instead, L occurring at t_2. Instead of 'unambiguous' one could use the term 'deterministic,' but philosophers usually use this term in the sense of 'necessitating.'

(4.13) Let me add a hypothesis concerning general facts about directednesses. I suspect that it is impossible that two events are alike but have different directions. Like events (i.e. events that are exactly similar) have like directions. By this I mean, as indicated above (pp. 47–48), that it is not contradictory but impossible that two events are alike but have different directions.

(4.14) With the notion of 'directedness' or 'direction' we can formulate a philosophical theory of laws of nature:

> *A law of nature states that events of a certain kind are the bases of directednesses of a certain kind.*[12]

(4.15) If one accepts the claim that physical events are necessarily necessitated by preceding physical events and adds the thesis that similar events necessitate similar events (events of type x necessarily necessitate events of type y), then one arrives at the thesis, which I call the *regularity thesis*, that there are regularities of succession of the form 'every event of type x is followed by an event of type y' and that every event is an element of such a regularity of succession. Some accept the regularity thesis not because they assume that physical events are necessarily necessitated but because they assume that it is entailed by the laws of nature, or because they believe to observe its truth.

The regularity thesis entails causal closure (at least the nonmodal version), because the physical effect of an immaterial cause would be a physical event instead of which another physical event would have occurred if the

[12] For a detailed defence of this theory of laws, see (Wachter 2009, chap. 6). There I used the term 'tendency' instead of 'directedness.'

intervention had not occurred. More precisely, a physical effect of an immaterial cause would be a physical event z which follows an event x which, had it not been counteracted, would have caused y, while y is incompatible with z. Those x-like events which are not counteracted cause y-like events. So if z is the physical effect of an immaterial cause, then some x-like events are followed by y-like events while others are not.

The directedness theory undermines the regularity thesis. If event S_1 was the complete cause of event S_2, then it is possible that some of the events that are like S_1 do not cause an event like S_2 because something prevents that through cooperation or through intervention; that is, either because there is some event at the same time together with which the S_1-like event does not have a direction towards an S_2-like event, or because something—a physical process or an immaterial cause—intervenes in the process which without the intervention would have led to an S_2-like event. Like events have like directions—in my view even necessarily—but *like events do not need to have like effects*. What is true is only that all S_1-like events that are not counteracted (through cooperation or through intervention) cause S_2-like events.[13]

(4.16) The word 'cause' is useful in ordinary language, but it lumps together quite different phenomena. I have suggested that if event S_1 at t_1 is the basis of a directedness towards event S_2 at t_2 and if that directedness is realised—that is, nothing stops the process—then it is true to say 'A caused B.' But we also make causal statements which are true but not of this type. For example: 'Because there was no dike, the field was flooded,' 'Because the wood was wet, it did not catch fire,' 'John caused the breaking of the window by throwing a stone into it,' 'Because John had not turned off the stove, a fire broke out in the kitchen.' Besides that, we tend to call only the last event which gets a process going the 'cause' of the event towards which the process leads. We call the spark the cause of the explosion. The gas in the room we might not call 'cause'—philosophers sometimes call it a 'condition'—of the event, although it is as much a part of the basis of the process as the spark is. Further, we generally do not call the ball A at

[13] If a complete cause is big in the sense that it involves many things, then the regularity is more likely to exist, because there will be few or no other instances of events that are exactly like this cause.

time t_1 the cause of the ball A at time t_2 (endurance), although, in my view, there is the same kind of process connecting the two stages of the same, enduring thing as there is in the case of the rolling ball. So there are many more cases which we call causation besides causation through processes as I described it.

On the other hand, while directedness is the mechanism of process causation, most directednesses are not realised, because they conflict with other directednesses, and therefore are not cases of causation. Most directednesses contribute to what happens by contributing to some resulting directedness that is realised but they are not realised. For example, the book on my desk is the basis of a directedness towards the moving downwards (i.e. towards being at certain positions at certain later times), but this directedness is not realised because the desk is in the way. 'Cause' is a success term in the sense that of the existing directednesses it singles out only those that are realised. So directedness is more fundamental and more widespread than causation is.

(4.17) Newtonian forces are an example of directednesses:[14] That there is a force acting on a certain body means that there is a directedness towards the body being at certain positions at certain later times. For forces we know precisely how they can be in *superposition* and how they behave when they conflict with each other. Every force can be counteracted by another force. If it is, then it does not accelerate the body in the way in which it would have done had it not been counteracted and had it been the resulting force. Every force can join forces with other forces. If two forces, for example two gravitational forces, are acting on a body then there is a resulting force whose direction and strength can be calculated using vector addition.[15] All the forces that are acting on a body compose the *resulting* force.

[14] This presupposes that component forces are real, as argued by Massin (2017) and Mumford and Anjum (2011, § 2.7). Schrenk (2010, § 5) suggested that 'Maybe dispositional powers can, at least metaphorically, be compared to Newtonian forces: a force pushes an object into a certain direction but it does not necessitate a movement, for other forces might well interfere.'

[15] Mumford and Anjum (2011, chap. 2) argue that causes can be 'modelled' as vectors. They spell out causation in terms of powers and dispositions, but their theory has much in common with my theory of directedness.

(4.18) Sometimes it is said that Newtonian physics is deterministic (in the Hobbesian sense) or that it supported belief in determinism. The opposite is true. Newtonian physics describes superposition, which means that forces can co-operate so that the resulting force acts as if only it and not the component forces existed. This entails that each force is open to co-operation with other forces and thus each physical cause is open to co-opereation with other events. Newtonian physics also implies that processes driven by forces can be interfered with. Thus causes that operate through Newtonian forces are never 'sufficient' in the Hobbesian and Leibnizian sense; no events 'exclude' the operation of further factors. Newtonian physics is precisely formulated in terms of *forces* and not in terms of *actual movements* because, contrary to determinism, it takes into account that all events and all forces can be counteracted and can co-operate with other events or forces. It does not make statements of the form 'Events of kind x are followed by events of kind y,' instead it makes statements of the form 'There are forces of kind x in situations of kind y.' Strictly speaking, it does not entail unconditional predictions but only predictions of the form 'In a state of affairs of type x, if nothing else is acting on the things involved, a state of affairs of type y will follow.'

5. The principle of causal openness re-stated in terms of directedness

(5.1) Now we can state causal openness in terms of directedness. If S_1 was the basis of a directedness towards S_2 at t_2 and that directedness was realised and thus S_1 caused S_2, then something could have prevented S_2 despite the occurrence of S_1, through some event cooperating with S_1, through the intervention of a physical process, or through the intervention of an agent. Thus, S_1 did not necessitate S_2.

That S_2 could have been prevented by a cooperating event means that there could have been an event at t_1 which together with S_1 would have formed the basis of a directedness towards towards an event that does not include S_2.

That S_2 could have been prevented by an intervention of a physical thing means that there could have been a causal process—nonprobabilistic or

probabilistic—that led to an event E between t_1 and t_2 which is incompatible with an event towards which S_1 was directed. If in this scenario that process was nonprobabilistic, then there was at t_1 no resulting directedness towards S_2 but one towards E that is based on the conjunction of S_1 plus the basis of the directedness towards E.

That S_2 could have been prevented by an intervention of a person means that there could have been an agent—for example an animal, a human, or God—who brings about a choice event between t_1 and t_2 which is incompatible with an event towards which S_1 was directed.

(5.2) Hobbes, Kant and many others believed that a physical event can occur only by being necessitated by preceding physical events. This demonstrates that if one believes p strongly, then one can come to believe that p is evident and obvious while in fact p is not evident or even false. When quantum mechanics was discovered and showed that probabilistic processes are at least possible that was a shock for those who believed in determinism. Confronted with quantum mechanics, the belief that a physical event can occur only by being necessitated by preceding physical events was replaced by the belief that an event can occur only by either being necessitated by preceding physical events or by being the result of a probabilistic process. But there is a third way of how an event can occur—and the first way is not a possibility. The most obvious possibility of how an event can occur seems to me to be: *An event can occur as the result of a causal process which is not probabilistic but which can be stopped.*

But we can imagine more ways how an event can come to occur, for example an event can occur as the result of a probabilistic process (that is, through a probabilistic directedness), or an event can occur through the action of a person so that the event has no preceding cause.

(5.3) To summarise: Physical causes are open to cooperation with other physical events and to counteraction by any possible causes, including physical things, souls, and God. Therefore a physical event cannot necessitate a later event, and it cannot exclude that something intervenes and prevents an event that it would have caused had nothing counteracted. Arguments from causal closure against souls acting on the body, against persons stopping causal processes or initiating causal processes through free actions, and against divine interventions require a modal

version of PCC, they need to claim that it is impossible that a physical event is not necessitated by preceding events. Arguments from causal closure fail because it is impossible that an event is necessitated by a preceding event. Whether there are souls, free actions, or divine interventions we can find out not by considering some general principle or modal truth but only by investigating the evidence.

Acknowledgements

I am grateful to the John Templeton Foundation for funding this research in the project 'The Openness of the Universe for Free Will and Special Divine Action, ID 59226.'

References

Bramhall, John. 1655 [1844]. "A Defence of True Liberty from Antecedent and Extrinsical Necessity." In *Works*, vol. 4, 23–196. Oxford.

Crusius, Christian August. 1743. *De usu et limitibus principii rationis determinantis vulgo sufficientis*. Leipzig.

Gibb, Sophie. 2015. "The Causal Closure Principle." *Philosophical Quarterly* 65 (261): 626–47. https://doi.org/10.1093/pq/pqv030

Hobbes, Thomas. 1654. "Of Liberty and Necessity." In *The English Works of Thomas Hobbes of Malmesbury*. Vol. 4, edited by W. Molesworth, 229–78. London.

Hobbes, Thomas. 1655. "Concerning Body." In *The English Works of Thomas Hobbes of Malmesbury*. Vol. 1: *Elements of Philosophy*, edited by W. Molesworth. London.

Kim, Jaegwon. 1996. *Philosophy of Mind*. Westview Press.

Kim, Jaegwon. 1998. *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*. Cambridge, MA: MIT Press.

Kim, Jaegwon. 2005. *Physicalism, or Something near Enough*. Princeton: Princeton University Press.

Lowe, E.J. 2008. *Personal Agency: The Metaphysics of Mind and Action*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199217144.001.0001

Massin, Olivier. 2017. "The Composition of Forces." *The British Journal for the Philosophy of Science* 68 (3): 805–46. https://doi.org/10.1093/bjps/axv048

Menzies, Peter. 2015. "The Causal Closure Argument is No Threat to Non-Reductive Physicalism." *Humana Mente* 8 (29): 21–46.

Montero, Barbara. 2003. "Varieties of Causal Closure." In *Physicalism and Mental Causation: The Metaphysics of Mind and Action*, edited by Sven Walter and Heinz-Dieter Hackmann, 173–87. Exeter: Imprint Academic.

Montero, Barbara Gail, and David Papineau. 2016. "Naturalism and Physicalism." In *The Blackwell Companion to Naturalism*, edited by Kelly James Clark, 173–87. 1st ed. John Wiley & Sons. https://doi.org/10.1002/9781118657775.ch13

Mumford, Stephen, and Rani Lill Anjum. 2011. *Getting Causes from Powers*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199695614.001.0001

Papineau, David. 2009. "The Causal Closure of the Physical and Naturalism." In *The Oxford Handbook of Philosophy of Mind*, edited by Brian McLaughlin, Ansgar Beckermann, and Sven Walter, 53–65. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199262618.003.0003

Schrenk, Markus. 2010. "The Powerlessness of Necessity." *Noûs* 44 (4): 725–39. https://doi.org/10.1111/j.1468-0068.2010.00755.x

Wachter, Daniel von. 2003. "How a Philosophical Theory of Causation May Help in Ontological Engineering." *Comparative and Functional Genomics* 4 (1): 111–14. https://doi.org/10.1002/cfg.258

Wachter, Daniel von. 2009. *Die kausale Struktur der Welt. Eine philosophische Untersuchung über Verursachung, Naturgesetze, freie Handlungen, Möglichkeit und Gottes kausale Rolle in der Welt*. Freiburg: Alber. https://doi.org/10.5282/ubm/epub.1975

Wachter, Daniel von. 2012. "Kein Gehirnereignis kann ein späteres festlegen." *Zeitschrift für philosophische Forschung* 66: 393–408. https://doi.org/10.3196/004433012802834909

Organon F 26 (1) 2019: 62–85
https://doi.org/10.31577/orgf.2019.26105
ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

RESEARCH ARTICLE

Why Determinism in Physics Has No Implications for Free Will

Michael Esfeld*

Received: 17 April 2018 / Accepted: 7 September 2018

Abstract: This paper argues for the following three theses: (1) There is a clear reason to prefer physical theories with deterministic dynamical equations: such theories are maximally rich in information and usually also maximally simple. (2) There is a clear way how to introduce probabilities in a deterministic physical theory, namely as answer to the question of what evolution of a specific system we can reasonably expect under ignorance of its exact initial conditions. This procedure works in the same manner for both classical and quantum physics. (3) There is no cogent reason to take the parameters that enter into the (deterministic) dynamical equations of physics to refer to properties of the physical systems. Granting an ontological status to parameters such as mass, charge, wave functions and the like does not lead to a gain in explanation, but only to artificial problems. Against this background, I argue that there is no conflict between determinism in physics and free will (on whatever conception of free will), and, in general, point out the limits of science when it comes to the central metaphysical issues.

Keywords: Classical mechanics; determinism; free will; functional reduction; Humeanism; physical laws; quantum mechanics.

* University of Lausanne
✎ Department of Philosophy, University of Lausanne, CH-1015 Lausanne, Switzerland
✉ Michael-Andreas.Esfeld@unil.ch Website: http://www.michaelesfeld.com/

© The Author. Journal compilation © The Editorial Board, *Organon F*.

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International Public License (CC BY-NC 4.0).

1. Determinism and probabilities in physics

The central claim of this paper is that laws in physics, even deterministic laws, do not pose a threat to human free will. That claim is intended to come out as a consequence of considering the role of laws and probabilities in physics as well as an argument to the effect that a certain version of Humeanism, dubbed Super-Humeanism, offers the best metaphysical account of these laws as they figure in our physical theories. Therefore, the paper first goes into determinism and probabilities (this section), then considers the ontological status of the magnitudes that enter into the laws of physics (section 2) and finally draws conclusions for free will (section 3).

Atomism is the paradigm on which the success of modern science is based. It is the idea that matter is composed of tiny, indivisible particles. In fact, atomism is as old as philosophy, going back to the Presocratics Leucippus and Democritus. The latter is reported as maintaining that

> substances infinite in number and indestructible, and moreover without action or affection, travel scattered about in the void. When they encounter each other, collide, or become entangled, collections of them appear as water or fire, plant or man. (Fragment Diels-Kranz 68 A57, quoted from Graham 2010, 537)

To turn to contemporary physics, Feynman says at the beginning of the famous *Feynman lectures*:

> If, in some cataclysm, all of scientific knowledge were to be destroyed, and only one sentence passed on to the next generations of creatures, what statement would contain the most information in the fewest words? I believe it is the *atomic hypothesis* (or the atomic *fact*, or whatever you wish to call it) that *all things are made of atoms—little particles that move around in perpetual motion, attracting each other when they are a little distance apart, but repelling upon being squeezed into one another*. In that one sentence, you will see, there is an enormous amount of information about the world, if just a little imagination and thinking are applied. (Feynman et al. 1963, chap. 1–2)

What makes atomism attractive is evident from these quotations: on the one hand, it is a proposal for an ontology of nature that is most parsimonious

and most general. On the other hand, it offers a clear and simple explanation of the realm of the objects that are accessible to us in perception. Any such object is composed of a finite number of discrete, pointlike particles. All the differences between these objects—at a time as well as in time—are accounted for in terms of the spatial configuration of these particles and its change. This view is implemented in classical mechanics. It conquered the whole of physics via classical statistical mechanics (e.g. heat as molecular motion), chemistry via the periodic table of elements, biology via molecular biology (e.g. molecular composition of the DNA), and finally neuroscience—neurons are composed of atoms, and neuroscience is applied physics (applied classical mechanics and electrodynamics, or, maybe, quantum mechanics in case quantum effects are proven to be operational in the brain).

What is relevant for the account of the perceptible macroscopic objects are only the relative positions of the point particles—that is to say, how far apart they are from each other, i.e. their distances—and the change of these distances. Let's call the particles "matter points." They can be considered as substances because they persist. However, in contrast to almost all the traditional philosophical accounts of substances, they are not Aristotelian substances in the sense of objects that have an inner form (eidos)—in other words, that are characterized by some intrinsic properties. There is nothing more to the matter points than the way in which they are spatially arranged and the change in their arrangement. But that is sufficient for their individuation: each matter point can be distinguished by—and hence individuated by—the distances it bears to the other matter points in a given configuration.

More precisely, if there is a configuration of N matter points $i, j, k \ldots$, there are $1/2(N–1)$ distance relations. These relations are irreflexive, symmetric and connex (meaning that all matter points in a given configuration are related with one another). They satisfy the triangle inequality. For these relations to individuate the matter points, we have to stipulate that if matter point i is not identical with matter point j, then the two sets that list all the distance relations in which these points stand with respect to all the other points in a configuration must differ in at least one such relation. We thereby exclude *entirely* symmetrical configurations among others. This is a structural individuation of the physical objects by relations in contrast

to intrinsic essences. It has the great advantage that we do not have to endorse the numerical plurality of matter points as a primitive fact, which would imply that the matter points are bare particulars or bare substrata. Instead, they are individuated by the distance relations. To put it in a nutshell, matter, consisting in matter points, is what is individuated by its standing in distance relations to each other (by contrast to minds, angles, or abstract objects, which, if they exist, do not stand in spatial relations).

We can thus sum up the gist of atomism in these two axioms [for a detailed argument, see (Esfeld and Deckert 2017, chap. 2.1), as well as (Esfeld 2017) for a concise metaphysical argument]:

(1) *There are distance relations that individuate simple objects, namely matter points.*

(2) *The matter points are permanent, with the distances between them changing.*

Let's call the ontology of nature defined by these two axioms the *primitive ontology*: matter points individuated by distances and their change are the ultimate referent of our physical theories, the bedrock of nature according to science so to speak.

However, the idea of matter being constituted by atoms in the sense of matter points is not sufficient to fulfil the promise of atomism, namely to explain everything in nature on the basis of the atomic hypothesis. That explanation is not carried out by the hypothesis of the atomic constitution of matter as such, but by showing how the change in the atomic composition of macroscopic objects accounts for their perceptible change. In other words, in order to fulfil the promise of atomism, one has in the first place to provide for laws of motion of the matter points and then to show how from these laws one also gets to an explanation of the motions of the macroscopic objects with which we are familiar. But the conceptual means provided by the primitive ontology—that is, the concepts of matter points, distances and their change admitted as primitive—are not sufficient to formulate a law of motion. Using only these conceptual means, one could not do much better than just listing the change that actually occurs, but not formulate a simple law that captures that change. The reason is that there is nothing

about the distance relations in any given configuration of matter that provides information about the evolution of these relations.

To extract such information from the configuration of matter, we have to embed that configuration in a geometry and a dynamics: we have to conceive the configuration of matter as being embedded in a space with a fully-fledged metric (such as three-dimensional Euclidean space)—although in the ontology, there are only distance relations and their change, but not an absolute space or space-time. Furthermore, we have to attribute parameters to the configuration of matter that are introduced in terms of their functional role for the change in the distance relations. These can be parameters that are attributed to the matter points individually (such as mass, momentum, charge), to their entire configuration (such as total energy, or an initial wave function), or constants of nature (such as the gravitational constant). They can always remain the same (such as mass and charge) or vary as the distance relations among the matter points change (such as momentum, a wave function, etc.). In any case, conceiving the configuration of matter as being embedded in a geometrical space and as being endowed with parameters that are set up in terms of their function for the change in the distance relations then enables the formulation of a physical law. Let us call these parameters and the geometry, providing for physical laws, the *dynamical structure* of a physical theory. In fact, the geometry, the dynamical parameters and the laws come as a package: the precise functional definition of the dynamical parameters involves the law, and the law is formulated by using the dynamical parameters as well as the geometry. But there is no threatening circularity here: roughly speaking, all three are conjectured at once and then made precise together in order to achieve a theory that is simple and rich in informational content.

The claim then is that the primitive ontology remains constant—from Democritos to today's physics—whereas the dynamical structure changes as we make more progress in formulating a theory that describes the evolution of the configuration of matter in a way that is ever more informative, while remaining as simple, general and informative as possible [for details, see (Esfeld and Deckert 2017, chap. 2.2)]. In other words, there is something in a physical theory that serves as the—ultimate—referent of the theory,

what there simply is in nature according to the theory. That something can be specified independently of the theory change in the history of science: atoms in the guise of matter points characterized by their relative positions and the change of these positions are that something. Furthermore, there is something in a physical theory that is introduced in terms of the role that it plays (i.e. its function) for the evolution of what there simply is according to the theory.

Thus, in classical mechanics, point particles characterized by their relative positions are what there simply is according to the theory—they have no further function in the theory apart from filling the place of the candidate for what simply exists in nature—whereas the parameter of mass, for instance, is introduced in terms of what it does for the motion of the particles. As, for instance, Mach (1919, 241) points out when commenting on Newton's *Principia*, "The true definition of mass can be deduced only from the dynamical relations of bodies." That is to say, both inertial and gravitational mass are introduced through their dynamical role, namely as dynamical parameters that couple the motions of the particles to one another. In general, even if attributed to the particles taken individually, mass, charge, etc. express a dynamical relation between the particles instead of describing an intrinsic essence of the basic objects. As Hall (2009, § 5.2) puts it,

> the primary aim of physics—its first order business, as it were—is to account for *motions*, or more generally for change of spatial configurations of things over time. Put another way, there is one Fundamental Why-Question for physics: Why are things located where they are, when they are? In trying to answer this question, physics can of course introduce *new* physical magnitudes—and when it does, new why-questions will come with them.

This, again, alludes to the crucial distinction between primitive ontology and dynamical structure: the fundamental issue is the location of things and its change. The account of this fundamental issue requires the introduction of further parameters that allow us to formulate laws about how the change of location of things occurs.

Fields can with good reason be taken to belong also to the dynamical structure of physical theories instead of being part and parcel of the primitive ontology. In brief, (a) all the evidence for fields derives from evidence

of particle motion. More importantly, (b) if one includes fields on a par with particles in the primitive ontology, the mathematical problem that there is no rigorous formulation of a physical theory of particle-field interactions, neither in classical electrodynamics nor in quantum field theory, becomes a philosophical problem how to conceive the interaction of these entities in the ontology. However, (c) if fields belong to the primitive ontology, their status is not clear: Are they properties of space-time points, albeit not geometrical ones? And why should only some space-time points have these causal properties (i.e. those where the field magnitudes are not zero)? Are they some sort of a bare substratum physical stuff? As Feynman puts it in his Nobel lecture,

> You see, if all charges contribute to making a single common field, and if that common field acts back on all the charges, then each charge must act back on itself. Well, that is where the mistake was, there was no field. It was just that when you shook one charge, another would shake later. There was a direct interaction between charges, albeit with a delay. [...] Now, this has the attractive feature that it solves both problems at once. First, I can say immediately, I don't let the electron act on itself, I just let this act on that, hence, no self-energy! Secondly, there is not an infinite number of degrees of freedom in the field. There is no field at all. (Feynman 1966, 699–700; see Lazarovici 2018 for a detailed exposition of the arguments against a commitment to fields in the ontology of physics).

Also in what is known today as the standard model of particle physics in the framework of quantum field theory, the ontology of this physics can be set up in terms of a particle ontology only and the conceptual problems that this physics raises can thereby be answered [for a detailed account, see (Esfeld and Deckert 2017, chap. 4)].

In general, the benchmark for the dynamical structure of a physical theory is to simplify the representation of the change that takes place in the configuration of matter—by contrast to merely dressing a list of that change—without losing the information about the change that actually occurs. The common way to achieve this benchmark is to specify a dynamical structure such that, for any configuration of matter given as initial condition,

the law fixes how the universe would evolve if that configuration were the actual one. The dynamical structure then goes beyond the actual configuration of matter: it fixes for any possible configuration of matter what the evolution of the universe would be like if that configuration were actual. It thereby supports counterfactual propositions.

Against this background, it is evident why determinism is a virtue of a physical theory: dynamical parameters figuring in laws that fix all the change, given an initial configuration of matter, are the simplest and most informative way to capture change. In other words, in the ideal case, the law is such that given an initial configuration of matter as input, the law yields a description of all the—past and future—change of the configuration as output. The question that remains in this case only is whether that description is empirically correct and whether it can be further simplified without losing informational content. It may turn out that, as a matter of fact, such a law cannot be achieved. It may also be that an indeterministic theory is simpler than a deterministic one and that the gain in simplicity outweighs the gain in informational content that a deterministic theory provides, such that, when seeking for the best balance between these two criteria, the indeterministic theory wins [see (Werndl 2013) for a detailed elaboration on these issues]. That notwithstanding, if there are dynamical parameters that designate only possibilities for how the configuration of matter may evolve, given an initial configuration, there always remains the question open whether one can do better, that is, find dynamical parameters that fix that change.

In any case, a fundamental physical theory is such that it defines a dynamical structure for the configuration of matter of the *whole* universe. For example, in Newtonian gravitation, the acceleration of any particle at any time depends, strictly speaking, on the positions and masses of *all* the other particles in the universe at that time. Even if action at a distance in Newtonian gravitation is replaced with local action in classical field theories, as soon as there are globally conserved quantities (such as total energy), the motion of any one object in the universe then is represented as being correlated with, in the end, the motion of any other object in the universe such that the quantity in question is globally conserved. In quantum physics, again, strictly speaking, due to entanglement, there is only one wave function

for the configuration of matter as a whole at any given time (i.e. the universal wave function).

On the one hand, thus, the dynamical structure of a fundamental physical theory is defined for the universe as a whole. On the other hand, any such dynamical structure is *per se* useless for calculations. We cannot know initial conditions for the configuration of matter as a whole. Furthermore, the evolution of a given configuration of matter points that we can manipulate may be extremely sensitive to perturbations on its initial conditions. Hence, a slight error about the initial conditions may lead to a great error in predicting the evolution of the system. Already this fact makes clear that there is no conceptual link between deterministic laws and our ability to predict with certainty the evolution of a given system. Everything depends on the extent to which we can specify the initial conditions of a system and on how sensitive the evolution of the system is to slight variations of its initial conditions.

By way of consequence, setting out a primitive ontology and a dynamical structure is not sufficient to build up a physical theory. The dynamical structure has to be construed in such a way that it allows us to answer the following question: What evolution of a given system can we typically expect—that is, in the vast majority of situations—under ignorance of its exact initial conditions? For instance, when flipping a coin, it is impossible to predict the individual outcomes and thus to predict the exact sequence of heads and tails, although this sequence is completely determined by the laws of classical mechanics and the initial conditions. Nevertheless, it is possible to derive the proposition that in by far the most cases, the number of heads will be almost equal to the number of tails provided that the number of coin flips is large enough. There are situations in which we can predict individual outcomes, such as when throwing a stone on Earth, but these are the exception rather than the rule. The dynamical structure of a physical theory therefore has to be linked with a typicality or probability measure by means of which we can derive propositions about which evolution of particular systems we can expect to obtain in most cases under ignorance of the exact initial conditions. There hence is a clear reason why even a deterministic physical theory requires probabilities and a detailed procedure how to introduce them on the basis of—fundamental and universal—laws [for details, see (Esfeld and Deckert, chap. 3.4)].

As regards classical mechanics, notably Boltzmann has established how to derive such probabilistic statements from the deterministic laws via a typicality or probability measure [see (Lazarovici and Reichert 2015) for a detailed account]. Classical statistical mechanics then paved the way for developing atomism into precise scientific theories also in chemistry, biology and beyond. As regards quantum mechanics, it is by no means evident that the situation with respect to probabilities is different from the one in classical physics. It is a fact that situations like the classical coin toss are generic in quantum mechanics—that is, situations that are highly sensitive to slight variations of the initial conditions, and we cannot know these initial conditions with arbitrary precision. This fact is brought out by Heisenberg's uncertainty relations. Consequently, we can only make predictions about the statistical distributions of measurement outcomes by using Born's rule, but in general not predictions about individual measurement outcomes.

However, this fact does not imply that probabilities have another status in quantum mechanics than in classical mechanics. The question is what the law of motion for the evolution of the individual quantum systems is that underlies Born's rule for the calculation of measurement outcome statistics. Only if one includes what is introduced in the textbook presentations of quantum mechanics as the postulate of the collapse of the wave function upon measurement into the law does one obtain an indeterministic law in quantum mechanics. Doing so requires amending the Schrödinger equation with parameters that include the collapse of the wave function under certain circumstances. As things stand, these parameters have to be introduced by hand and compromise the simplicity of the law [see (Ghirardi, Rimini, and Weber 1986)]. Furthermore, they lead to predictions that deviate from the textbook ones in certain specific situations. In any case, it is an open issue whether such an indeterministic law is a fundamental or rather a phenomenological one—taking gravitation into account, for example, may turn this law into a deterministic one [see (Penrose 2004, chap. 30)]. The only example of a candidate for an indeterministic law in a fundamental physical theory hence confirms the general statement made above, namely that in the case of an indeterministic law, it remains an open issue whether that law can be turned into a deterministic one by including further parameters.

Apart from the version of quantum mechanics that includes the postulate of the collapse of the wave function in the physical law, there are two other versions that both are deterministic. In brief, the version going back to (Everett 1957) admits only the Schrödinger equation and, in consequence, no unique measurement outcomes. It is therefore known as many worlds quantum mechanics, because, in short, every possible outcome of a measurement becomes real in a branch of the universe [see (Wallace 2012) for details]. The version going back to (Bohm 1952a, 1952b) adds to the (deterministic) Schrödinger equation a further (deterministic) law, known as the guiding equation, that describes, in brief, how the particles move in physical space as guided by the wave function. In the elaboration of this theory known as Bohmian mechanics, it is shown how Born's rule can be deduced from these laws by means of a typicality or probability measure that is linked with these laws in a way that matches the way in which the probability calculus of classical statistical mechanics is deduced from the deterministic laws of classical mechanics [see (Dürr, Goldstein, and Zanghì. 2013, chap. 2)]. The existence of Bohmian mechanics hence refutes any attempt to infer from Born's rule—or the Heisenberg uncertainty relations, or the randomness of individual measurement outcomes—the conclusion that probabilities have a more fundamental status in quantum mechanics than in classical mechanics. The question is what the law is that underlies Born's rule. The standard for assessing the proposals for that law is independent of the issue of determinism vs. indeterminism. The standard is what is the best solution to the quantum measurement problem (as illustrated, for instance, in Schrödinger's cat paradox). There are cogent arguments in favour of the Bohmian solution to this problem [see e.g. (Esfeld 2014)]. The consequence then is that probabilities in quantum physics have the same status as probabilities in classical physics.

To sum this section up:

(1) There is a clear reason to seek for deterministic laws in the formulation of a physical theory, since these maximize informational content and usually also simplicity.

(2) There is a clear procedure available how to get from fundamental deterministic laws to predictions about statistical distributions of measurement outcomes both in classical and in quantum physics.

(3) Apparently random behaviour of investigated systems (including rules stating that randomness, such as the Heisenberg uncertainty relations) never justifies the conclusion to indeterminism. The issue is what the laws underlying this behaviour are. It is true that the determinism in classical mechanics would lose persuasiveness if there were not the clear cut paradigm examples of deterministic predictions in classical gravity (such as throwing a stone on Earth), and it is a fact that there are no such clear cut cases in quantum mechanics. But this is merely a heuristic matter. There is no conceptual link from deterministic laws to deterministic predictions, and, hence, no link from probabilistic predictions to probabilistic laws either.

2. Explanations in physics

The *raison d'être* for laws in physics is that they explain the observed phenomena by subsuming them under a law—in whatever way one then spells out in philosophy of science how bringing phenomena under a law explains them (covering law, causal explanation, unification, just to name the most prominent accounts). This role of the laws raises the issue of their ontological status. In any case, as regards our knowledge, we cannot but make conjectures about what the laws are based on the data that become available to us. The standard for these conjectures is the extent to which they optimize both simplicity and informational content in accounting for the data. According to the stance known as Humeanism in today's metaphysics, this is all there is to the laws: they are nothing more than means of representation that seek to optimize simplicity and informational content. Super-Humeanism goes beyond standard Humeanism [see e.g. (Lewis 1986, introduction)] by putting the geometry and the dynamical parameters—that is, the dynamical structure—also on the side of the laws: the ontology is only the primitive ontology, such as matter points individuated by distance relations and the change in these relations. That change manifests certain patterns. Geometry, dynamical parameters and the laws linked with a probability measure are the package that enables us to achieve a representation of these patterns that is both as simple about the patterns and as

informative about the change as possible [see (Esfeld and Deckert 2017, chap. 2.3), for details].

(Super-)Humeanism is distinct from instrumentalism. It is a scientific realism: the claim is that what there is as far as the ontology of the natural world is concerned is exhausted by the primitive ontology. Dynamical parameters have a nomological role by figuring in the laws of nature. From that nomological role then derives their role in the predictions, as the laws are linked with a procedure to derive probabilities from them as sketched out in the previous section. The claim of Humeanism then is that the laws do not require additional ontological commitments. The claim of Super-Humeanism is that geometry, dynamical parameters and laws form a package that has only a representational purpose and that does hence not call for ontological commitments that reach beyond the primitive ontology. In short, the issue is what the ontology of the natural world is in a scientific realist framework.

Of course, physics explains the motions of bodies by using a geometry and dynamical parameters that appear in laws. However, the argument for an ontological commitment to the geometry and the dynamical parameters cannot simply be that they figure in our best physical theories. Reading the ontology off from the mathematical structure of physical theories would be begging the question of an argument for ontological commitments that go beyond what is minimally sufficient to account for the phenomena in a scientific realist vein, namely the commitment to a primitive ontology as given, for instance, by the two axioms of distance relations individuating matter points and the change in these relations. In a metaphysics based on science, the argument can only be that by subscribing to ontological commitments that go beyond that minimum, one achieves a gain in explanation.

(Super-)Humeanism can accommodate the scientific practice of explanations and its conceptualisation in terms of covering laws, causation or unification. There is no space in this paper to expand on this claim [see notably (Loewer 2012) for details and the ensuing debate with (Lange 2013), (Miller 2015) and (Marshall 2015)]. The core idea of the (Super-)Humean account is this one: the geometry and the dynamical structure of a physical theory explain the phenomena by bringing out the patterns or regularities in the motion of the particles; bringing out these patterns or regularities requires

no ontological commitment beyond particles that move. On (Super-)Humeanism, first comes the particle motion, which as a contingent matter of fact exhibits certain patterns or regularities, then come the laws, including the geometry. Hence, the laws, the parameters figuring in them and the geometry are not some sort of an agent that forces the particles to move in a certain way. The laws do not constrain the particle motion. It is the particle motion that fixes the laws. Hence, if one asks why there are the patterns in the particle motion that there are in fact, (Super-)Humeanism cannot answer that question. The claim of (Super-)Humeanism is that there is no scientific answer to that question. Our scientific understanding of the world comes to an end once the salient patterns in the change of the elements of the primitive ontology are reached, such as, for instance, attractive particle motion.

The argument for this claim is the one illustrated in Molière's piece *Le malade imaginaire*: one does not explain why people fall asleep after the consumption of opium by endorsing a dormitive virtue of opium, because the dormitive virtue is *defined* in terms of its functional role to make people fall asleep after the consumption of opium. By the same token, one does not obtain a gain in explaining attractive particle motion by subscribing to an ontological commitment to gravitational mass as a property of the particles, because mass is *defined* in terms of its functional role of making objects attract one another as described by the law of gravitation. Of course, mass, charge and the like are fundamental and universal physical magnitudes, by contrast to the dormitive virtue of opium. But the point is that they are defined in terms of the functional role that they exert for the particle motion. Why do objects move as they do? Because they have properties whose function it is to make them move as they do. An ontological commitment to such properties does not yield a gain in explanation. The same holds for forces, fields, wave functions, an ontic structure of entanglement in quantum physics, laws conceived as primitive, etc. It also applies to geometry: it is no gain in explanation to trace the characteristic features of the distance relation back to the geometry of an absolute space, because that geometry is defined such that it allows for the conception of distances in that space.

It is true that by tracing the distance relations back to an absolute space, or the change in the distance relations back to properties of the particles that are dispositions for that very change, the characteristic features

of the distance relations as well as those of the patterns in the change in them come out as necessary instead of contingent. However, merely shifting the status of something from contingent to necessary does not amount to a gain in explanation. Quite to the contrary, one only faces drawbacks that come with the commitment to a surplus structure in the ontology in the guise of an absolute space, fundamental dispositional properties of the particles, ontic dynamical structures of entanglement, etc.: differences with respect to absolute space that do not make a difference in the configuration of matter, questions such as how an object can influence the motion of other objects across space in virtue of properties that are intrinsic to it, how a wave function defined on configuration space can pilot the motion of matter in physical space, etc. [see (Esfeld and Deckert 2017, chap. 2.3)].

To sum this section up:

(1) The business of physics is to achieve on the basis of the available evidence a theory that is as simple and as informative as possible in accounting for that evidence and in predicting new evidence, with such a theory being characterized by the three features outlined in the previous section.

(2) Given the primitive ontology in terms of the notions of distances individuating matter points and the change of these distances, one can then define any further notion that one needs in one's theory of the natural world in terms of its functional role in the represenation of that change, without thereby subscribing to an additional ontological commitment. An ontology that is limited to a primitive ontology of matter points individuated by distance relations and the change in these relations is a scientific realism that is sufficient to accommodate scientific explanations.

(3) Subscribing to an ontological commitment that goes beyond what is minimally sufficient to account for the evidence (i.e. the primitive ontology) is not implied by the physics: one cannot read off the ontology from the mathematical structure of a physical theory. The issue can only be whether granting that structure an ontological status over and above the primitive ontology yields an explanatory gain.

However, far from doing so, such an enriched ontology leads only to drawbacks stemming from a commitment to surplus structure.

3. Free will and the limits of physics

Minimizing the ontological commitments of physics as outlined in the two preceding sections, while fully respecting scientific realism, not only prevents artificial problems from arising in the philosophy of nature, but also has repercussions for metaphysics in general. In particular, against the background set out here, one can establish the conclusion that there is no conflict between physical determinism and free will—although, at first glance, there obviously seems to be such a conflict.

Suppose that classical mechanics were the correct physical theory of the universe. Then, given an initial state of the particle motion throughout the whole universe (which includes the attribution of values of mass to the particles over and above initial positions and velocities) and the laws of classical mechanics, the entire evolution of the universe is fixed by the laws—that is, the entire *future* evolution from that state on as well as the entire *past* evolution leading to that state; that is why this can be an initial state at an arbitrary time. Of course, already in classical mechanics, as pointed out at the end of section 1, nobody within the universe could know its initial state at any time with enough precision to turn the determinism implemented in the laws into predictions.

If one considers physical determinism to be troublesome when it comes to human free will, a little reflection shows that the determinism implemented in the dynamical structure of classical mechanics is not the reason for the trouble. The reason is the very fact of there being universal physical laws. Suppose that a version of quantum mechanics that includes what is known as the collapse of the wave function in the fundamental dynamical law were the correct physical theory of the universe [such as the theory of (Ghirardi, Rimini, and Weber 1986) mentioned at the end of section 1] and that the collapse of the wave function is an irreducibly stochastic process. Nevertheless, the dynamical law then fixes objective probabilities for wave function collapse to occur such that, given an initial quantum state of the universe at an arbitrary time that includes an initial wave function of the

universe, several possible future evolutions of the universe are fixed with objective probabilities attached to them. If the decisions of human beings concerning the motions of their bodies can influence neither the initial state of the universe nor the deterministic laws of classical mechanics (on the supposition that they are the correct laws of the universe), they cannot influence the objective probabilities implemented in a fundamental stochastic law and an initial wave function either (on the supposition that wave function collapse is stochastic and included in the fundamental dynamical law of the universe) [see (Loewer 1996)]. Hence, supposing that there is a conflict between deterministic physical laws and free will, one could not draw any profit for free will if the laws were indeterministic. If there is such a conflict, it concerns the very fact of there being universal physical laws, be they deterministic or not.

The common formulation that, in the case of determinism, the laws plus the initial conditions fix the entire evolution of the objects to which they apply may suggest that the laws somehow bring about the evolution of these objects. However, if this were so, the laws would bring about the *past* evolution of the objects from an arbitrary initial state back as well as the *future* evolution of the objects from an arbitrary initial state on. But no one thinks that the fact that given an initial state and a deterministic dynamical law, the past evolution leading to that state is fixed implies that the law brings about the past evolution by retrocausation. Hence, the mere statement of determinism contains no reason to think that the law brings about the future evolution either. A better formulation of determinism that avoids any ontological connotation of the verb "fix" therefore is this one: the propositions stating the laws of nature and the propositions describing the state of the world at an arbitrary time t (i.e. the propositions describing the initial conditions) entail the propositions describing the state of the world at any other time. Thus formulated, it is clear that determinism in science is—only—about entailment relations among propositions. The question then is, supposing that determinism is true, what it is in the world that makes these propositions true, that is, in virtue of what in the ontology these entailment relations among propositions hold.

On the conception of physical laws sketched out in the preceding section, there can be no clash between laws of nature and free will (in whatever way

one may conceive free will). The reason is, in brief, that the motion comes first, including the motion of our bodies that is the expression of our intentions, then come the laws. In other words, what makes the propositions that state the laws true is the *entire* motion of the objects in the universe, that is, the change that actually occurs throughout the entire evolution of the configuration of matter of the universe. If the laws are "mere patterns in the phenomena", as Hall (2009, 1) puts it, they do not govern or constrain those phenomena, let alone bring them about. Hence, in this case, there is no clash between laws of nature and human free will possible, since the bodily movements that humans choose to make are part of the phenomena. The laws are there to achieve an account of the motions that actually occur that is both maximally simple and maximally informative. Consequently, the laws do not predetermine our actions, they only represent what happens in nature [see (Beebee and Mele 2002)].

Only if one loads the laws of physics with some sort of necessitation—such as by conceiving them as modal primitives, tracing them back to fundamental dispositions, powers or modal ontic structures instantiated by the physical objects—can a conflict with free will ensue (at least on an incompatibilist conception of free will); there then is something in the world that is independent of our decisions and that makes our decisions necessary. However, as far as the ontology of physics is concerned, there is no need to subscribe to any such commitment, and doing so leads only to drawbacks, as argued in the previous section.

Consider the famous consequence argument by means of which van Inwagen seeks to establish a conflict between free will and determinism:

> If determinism is true, then our acts are the consequences of the laws of nature and events in the remote past. But it is not up to us what went on before we were born, and neither is it up to us what the laws of nature are. Therefore the consequences of these things (including our present acts) are not up to us. (van Inwagen 1983, 16)

On the view defended in this paper, the statement "it is not up to us what went on before we were born" is ambiguous if it refers to the initial state of the universe and the statement "neither is it up to us what the laws of nature are" is, strictly speaking, not quite correct. The latter statement is

not quite correct on any version of Humeanism about laws, for the just mentioned reason that what goes on in the universe comes first and then come the laws. However, it would be implausible to take this to imply that if a person had chosen to do otherwise, the laws of nature would have been different.

Here again the virtues of Super-Humeanism show up: it is not only the laws, but the entire dynamical structure of the correct physical theory of the universe that depend on the change in the universe as a whole. As argued in the first section, all the dynamical parameters that are introduced in terms of their functional role for the change in the primitive ontology—that is, the particle motion—are there to simplify, that is, to achieve a representation of the particle motion that is as simple and as informative as possible. Thus, they are not intrinsic to the particles or their configuration at any time. That is to say: the state of the universe at any given time, which enters as initial condition into the laws, contains elements that are not intrinsic to what there is at that time, but depend on the overall change in the universe. These are notably the initial values of parameters such as mass, fields, the universal wave function, etc. In order for these parameters to play their role to simplify the account of the motion that actually occurs in the universe, what role these parameters play and, notably, what their initial values are, depends on the change that actually occurs in the universe—that is, to stress again, the correct value of these parameters that enters into the state of the universe *at any given time* depends not only on what motion happens in the universe earlier than that time, but also on what happens *later* than that time. To put it in a nutshell, we do not know the initial wave function of the universe not only because of a principled limit on our knowledge, but also because what is the *initial* wave function of the universe is only fixed at the *end* of the universe so to speak (because it depends on what the particle motion during the evolution of the universe turns out to be like).

Consequently, if human beings chose to do otherwise, in the first place, slightly different initial values for the dynamical parameter at the initial state of the universe would have to be figured out in order to achieve a system that maximizes both simplicity and informational content about the change that actually occurs in the universe. For the sake of illustration,

assume that quantum physics is the correct theory of the universe. Then, what would be slightly different if humans chose to do otherwise than they actually did were not the Schrödinger equation and the Bohmian guiding equation or the GRW collapse law in the first place, but the universal wave function, that is, the values that this wave function takes as initial condition. In that way, van Inwagen's consequence argument turns out to be invalid without the Humean being committed to saying that it is up to us what the laws of nature are. Instead, there is an ambiguity in the phrase "it is not up to us what went on before we were born." If that phrase is to include reference to an initial state of the universe before we were born, then that initial state, insofar as it enters into a law of nature, includes values of parameters that are not intrinsic to that state, but that depend on what happens later in the universe, including the particle movements that are expressions of human free will.

Hence, this paper is directed against a certain sort of a scientific worldview, namely one that implies a misconception of the enlightenment that comes with science: it is not that science teaches us that if there are deterministic laws in physics—or, for the sake of the argument, deterministic laws in genetics or evolutionary biology—, our decisions are necessitated by factors that are outside of our control. In general, this paper is about limits of science when it comes to the central metaphysical questions. In contrast to other attempts in that sense that argue for a limitation of the range of physical laws within the physical domain itself [see e.g. (von Wachter 2015) according to whom physical laws, even when they are deterministic, indicate only tendencies for what happens in nature], the argument of this paper takes universal physical laws, also when they are deterministic, at face value as encompassing all the motions of bodies in the universe in a simple and general equation (or at least as striving for that ideal, as illustrated by the Newtonian law of gravitation). The argument then is that attributing a modal status to these laws is not justified by the physics, even if scientific realism is taken for granted. From that then follow certain limits of science, in particular that there is no clash between the scientific representation of the motions of bodies in terms of universal and deterministic laws and some such motions being the manifestation of human free will.

Once one has identified a primitive ontology of the natural world and thus settled for the concepts admitted as primitive that characterize that ontology, it is possible to define every further concept that enters into one's theory of the world in terms of the function for the primitive ontology. This applies not only to the parameters that appear in physical theories, but to any concept, including the ones describing the mind. It is at least since (Lewis 1972) well known how to provide a scheme for the functional definition also of mental concepts in terms of, in the last resort, changes of the physical configuration of the body and its environment. Such functional definitions are undisputed in the natural sciences: it would be odd, for instance, to postulate a heat stuff to account for thermodynamical phenomena, since these can be defined functionally in terms of changes in molecular motion. By the same token, it would be odd to postulate an *élan vital* to account for organisms and their reproduction. Again, since the advent of molecular biology, the evolution of organisms and their reproduction can be accounted for in terms of molecular biology. There is no explanatory gap here.

However, when it comes to consciousness as well as rationality and the normativity and free will that are linked to rationality, one may maintain that there is an explanatory gap in the sense that functional definitions in terms of, in the last resort, changes in the configuration of matter do not capture what is characteristic of mental phenomena [see (Levine 1983)]. Once one has understood the science, it is obvious how a functional definition of, for instance, water in terms of the effects on the interaction of H_2O molecules captures and explains the phenomenal features of water and how a functional definition of organisms captures and explains their reproduction, including the link from genotypes to phenotypes. However, it is not obvious—at least not obvious in the sense of these paradigmatic examples—what the qualitative character of conscious experience, or the normativity that comes with rationality have to do with molecular motion in the brain.

The argument of this paper implies the following: in case the mental cannot be functionally defined on the basis of a primitive ontology of matter in motion, then an ontological commitment to the mental is called for over and above the ontological commitment to a primitive physical ontology.

Moreover, such an ontological commitment then is as fundamental as the commitment to a primitive physical ontology, although the mental may only become manifest in certain systems in the universe and only at a certain period of time in the evolution of the cosmos. In general, whatever does not come in as being entailed by the primitive ontology via functional definitions is itself a further fundamental ingredient of the ontology [cf. e.g. (Jackson 1994), or (Chalmers 2012), although the argument of this paper is not committed to *a priori* entailment]. This makes (again) evident the price that comes with any position whose ontological commitments go beyond a primitive physical ontology.

In the case of the additional parameters figuring in scientific theories, there is no reason to pay that price, as argued in the previous section. But the case of the mental is different. Positions that seek to avoid paying that price for instance by putting their stakes on emergence do not cut the ontological ice: if what emerges can be functionally defined on the basis of the ontology that is admitted as primitive, then there is no emergence in the sense of something that calls for new ontological commitments. If what emerges cannot be thus defined, then one is committed to more in the ontology than the ontology originally admitted as primitive. Consequently, there then are further primitives that hence have the same ontological status as the original primitives.

This is the core metaphysical debate, about the cosmos and about our place in it. Science can be understood on the basis of a primitive ontology that, even if the dynamics for that primitive ontology is deterministic, has no implications for what is right or wrong about these core metaphysical issues. Elaborating on the primitive ontology of science makes, however, clear the price that one has to pay for any further ontological commitments that then would have to come in as further primitives. The credibility of any such commitments hinges upon working them out into an overall metaphysical position that matches the paradigm of science in its clarity and precision as well as the concrete explanations that it provides.

References

Beebee, Helen, and Alfred Mele. 2002. "Humean Compatibilism." *Mind* 111 (442): 201–23. https://doi.org/10.1093/mind/111.442.201

Bohm, David. 1952a. "A Suggested Interpretation of the Quantum Theory in Terms of 'Hidden' Variables. I." *Physical Review* 85 (166): 166–79. https://doi.org/10.1103/PhysRev.85.166

Bohm, David. 1952b. "A Suggested Interpretation of the Quantum Theory in Terms of 'Hidden' Variables. II." *Physical Review* 85 (180): 180–93. https://doi.org/10.1103/PhysRev.85.180

Chalmers, David. 2012. *Constructing the World*. Oxford: Oxford University Press.

Dürr, Detlef, Sheldon Goldstein, and Nino Zanghì. 2013. *Quantum Physics without Quantum Philosophy*. Berlin: Springer. https://doi.org/10.1007/978-3-642-30690-7

Esfeld, Michael. 2014. "The Primitive Ontology of Quantum Physics: Guidelines for an Assessment of the Proposals." *Studies in History and Philosophy of Modern Physics* 47: 99–106. https://doi.org/10.1016/j.shpsb.2014.06.003

Esfeld, Michael. 2017. "A Proposal for a Minimalist Ontology." Forthcoming in *Synthese*. https://doi.org/10.1007/s11229-017-1426-8

Esfeld, Michael, and Dirk-André Deckert. 2017. *A Minimalist Ontology of the Natural World*. New York: Routledge.

Everett, Hugh. 1957. "'Relative State' Formulation of Quantum Mechanics." *Reviews of Modern Physics* 29 (454): 454–62. https://doi.org/10.1103/RevModPhys.29.454

Feynman, Richard P. 1966. "The Development of the Space-Time View of Quantum Electrodynamics." Nobel Lecture, December 11, 1965. *Science* 153 (3737): 699–708. https://doi.org/10.1126/science.153.3737.699

Feynman, Richard P., Robert B. Leighton, and Matthew Sands. 1963. *The Feynman Lectures on Physics. Volume 1*. Reading MA: Addison-Wesley.

Ghirardi, Giancarlo, Alberto Rimini, and Thomas Weber. 1986. "Unified Dynamics for Microscopic and Macroscopic Systems." *Physical Review D* 34 (470): 470–91. https://doi.org/10.1103/PhysRevD.34.470

Graham, Daniel W., ed. 2010. *The Texts of Early Greek Philosophy. The Complete Fragments and Selected Testimonies of the Major Presocratics*. Translated by Daniel W. Graham. Cambridge: Cambridge University Press.

Hall, Ned. 2009. "Humean Reductionism about Laws of Nature." Unpublished manuscript. Available at http://philpapers.org/rec/HALHRA

Jackson, Frank. 1994. "Armchair Metaphysics." In *Philosophy in Mind. The Place of Philosophy in the Study of Mind*, edited by Michaelis Michael and John O'Leary-Hawthorne, 23–42. Dordrecht: Kluwer.

Lange, Marc. 2013. "Grounding, Scientific Explanation, and Humean Laws." *Philosophical Studies* 164 (1): 255–61. https://doi.org/10.1007/s11098-012-0001-x

Lazarovici, Dustin. 2018. "Against Fields." *European Journal for Philosophy of Science* 8 (2): 145–70. https://doi.org/10.1007/s13194-017-0179-z

Lazarovici, Dustin, and Paula Reichert. 2015. "Typicality, Irreversibility and the Status of Macroscopic Laws." *Erkenntnis* 80 (4): 689–716. https://doi.org/10.1007/s10670-014-9668-z

Levine, Joseph. 1983. "Materialism and Qualia: The Explanatory Gap." *Pacific Philosophical Quarterly* 64 (4): 354–61. https://doi.org/10.1111/j.1468-0114.1983.tb00207.x

Lewis, David. 1972. "Psychophysical and Theoretical Identifications." *Australasian Journal of Philosophy* 50 (3): 249–58. https://doi.org/10.1080/00048407212341301

Lewis, David. 1986. *Philosophical Papers. Volume 2*. Oxford: Oxford University Press.

Loewer, Barry. 1996. "Freedom from Physics: Quantum Mechanics and Free Will." *Philosophical Topics* 24 (2): 91–112.

Loewer, Barry. 2012. "Two Accounts of Laws and Time." *Philosophical Studies* 160 (1): 115–37. https://doi.org/10.1007/s11098-012-9911-x

Mach, Ernst. 1919. *The Science of Mechanics: A Critical and Historical Account of Its Development*. Fourth edition. Translated by Thomas J. McCormack. Chicago: Open Court.

Marshall, Dan. 2015. "Humean Laws and Explanation." *Philosophical Studies* 172 (12): 3145–65. https://doi.org/10.1007/s11098-015-0462-9

Miller, Elizabeth. 2015. "Humean Scientific Explanation." *Philosophical Studies* 172 (5): 1311–32. https://doi.org/10.1007/s11098-014-0351-7

Penrose, Roger. 2004. *The Road to Reality: A Complete Guide to the Laws of the Universe*. London: Jonathan Cape.

van Inwagen, Peter. 1983. *An Essay on Free Will*. Oxford: Oxford University Press.

von Wachter, Daniel. 2015. "Miracles Are Not Violations of the Laws of Nature Because the Laws Do Not Entail Regularities." *European Journal for Philosophy of Religion* 7 (4): 37–60. https://doi.org/10.24204/ejpr.v7i4.86

Wallace, David. 2012. *The Emergent Multiverse. Quantum Theory According to the Everett Interpretation*. Oxford: Oxford University Press.

Werndl, Charlotte. 2013. "On Choosing between Deterministic and Indeterministic Models: Underdetermination and Indirect Evidence." *Synthese* 190 (12): 2243–65. https://doi.org/10.1007/s11229-011-9966-9

Organon F 26 (1) 2019: 86–103
https://doi.org/10.31577/orgf.2019.26106
ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

RESEARCH ARTICLE

Does Divine Intervention Violate Laws of Nature?

Ralf B. Bergmann*

Received: 13 July 2018 / Accepted: 29 November 2018

Abstract: The present paper discusses processes in our world embedded in a dense net of causes and effects. The role of laws of nature is discussed using classical mechanics as an example to demonstrate that the behavior of objects is not determined by laws of nature on their own. Natural processes rather depend on the environment that implies forces and conditions under which these processes occur. These conditions are usually described in mathematical terms using suitable initial or boundary conditions or further constrains. This analysis may seem trivial but has an important consequence. Divine intervention appears not to violate natural laws, rather, new causes are introduced that lead to results that otherwise would not have occurred. A further consequence arises from the complexity of causal nets and the presence of probabilistic processes in our world as one may not be able to determine the causes that led to a certain effect.

Keywords: Causation; causal processes; probabilistic processes; laws of nature; equations of motion; divine action; miracles.

* University of Bremen
Faculty of Physics and Electrical Engineering, University of Bremen, Otto Hahn Allee NW1, 28359 Bremen, Germany
✉ Ralf.Bergmann@uni-bremen.de

© The Author. Journal compilation © The Editorial Board, *Organon F*.

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International Public License (CC BY-NC 4.0).

1. Introduction: the claim of causal closure

Many contemporary physicists, philosophers and theologians would argue that a God or other transcendent entities, given such entities exist, would not be able to intervene in the physical world, because every event is determined by physical processes that are governed by the laws of nature. Although this thinking has a long tradition, the meaning, the basis and the consequences of this statement, often referred to as *causal closure*, are often rather diffuse and widely differ amongst and within various disciplines. If this reasoning was correct, there would be no room for divine intervention of *whatever* kind. Most obviously, there would be no room for miracles.

The protestant theologian Friedrich Schleiermacher (1768–1834) redefined the term miracle as a mere description of events that have a somehow religious relevance, however common the event may appear.[1] Ludwig Feuerbach (1804–1872) well known for his critique on religion, describes the term miracle as something that is unthinkable and is in the end a meaningless concept (Feuerbach 2016, 125). Feuerbach's statement is, however, an a priori assumption rather than an argument. Perhaps one of the most influential modern protestant theologians arguing on the matter is Rudolf Bultmann (1884–1976). Although his viewpoint and that of his followers was never undisputed,[2] his views are still quite influential and widespread and offer a reasonably clear, though weak reasoning for his case (Bultmann 1960). In his famous essay entitled "Neues Testament und Mythologie" (New Testament and Mythology) he expresses the idea, that the miracles described in the New Testament cannot be taken as reports of historic events, "because experience and application of science and technology have progressed so far, that no man can and does seriously stick to the world view of the New Testament."[3] He concludes that by the knowledge of the powers

[1] For a compact source see (Schleiermacher 1969, 79), as reference for theologians (von Meckenstock 2011) is usually employed.

[2] See e.g. (C. S. Lewis 1947) for a popular version of criticism or (Wilckens 2017) for a theologically based criticism of Bultmann's approach.

[3] Translation of the author from (Bultmann 1960). A slightly shortened but easier accessible source of his paper is (Härle 2007, 174–88).

and laws of nature there is no room left for transcendent intervention and that humans are not open to the influence of "foreign" powers. Due to the philosophical and theological significance of causal closure implied by the statements of Bultmann and others and the claim of the impossibility of divine intervention, there was and is a longstanding and controversial debate [see e.g. (Mackie 1982)] that presently seems to undergo a revival [see e.g. (Keener 2011), (Larmer 2014) and (von Wachter 2015)]. For a philosophic discussion of various views on divine agency and definitions of miracles see e.g. (Larmer 2014, 7–52).

In this paper I will argue, that causal closure does not follow from physics but rather is a metaphysical principle. I will then investigate the nature of laws of physics and argue, that these laws in themselves do not determine physical processes. In what follows, classical mechanics serves as an example to demonstrate that the equations of motion describing the actual movement of a body are determined not only by natural laws but also by initial and boundary conditions or further constraints. This observation, although clear to every physicist, has a profound impact on the question, whether divine intervention violates laws of nature.

2. Causal networks and laws of nature

2.1. The clockwork universe

Before the arrival of quantum mechanics in the beginning of the 20$^{\text{th}}$ century, the universe was frequently conceived as of a kind of clockwork: The history of the universe and all its parts correspond to the functioning of a perfect clockwork. The development of the world is fully determined and can thus—at least in principle—be fully calculated once one knows the laws of nature and the state of the universe at one time. Past and future of the universe can be exactly calculated. For an imaginary being, the so called "Laplacian Demon" (Laplace 1814), capable of knowing and calculating all these details, nothing would come as a surprise.

Figure 1 shows a simple scheme of cause and effect. For the sake of simplicity, I will restrict the discussion at this point to processes related to classical mechanics: The movement or the status of an object (circle) is

changed by some effects (squares) at various points in time. The movement or change of the object can be calculated if one knows the initial conditions and the forces acting on the object, see Fig. 1a and b. Up to this point, the scheme only shows the basic building blocks of the clockwork universe. But there is already a problem at the first cause, see Fig. 1c: Every description in terms of physics needs a starting point, from which the following process can be calculated. The "first cause" is outside the realm of physics, see the bright square at the far left of Fig. 1 c. There is, contrary to frequently made claims even by physicists [see e.g. (Hawking 2010) and (Krauss 2012)], no way to explain "something" from "nothing" [see e.g. (Craig 2002) or (Bussey 2013)].

Figure 2 sketches a scenario that much more resembles the real world than Figure 1: In the real world many processes happen simultaneously and influence each other as indicated in Fig. 2a. Apart from the first starting point (bright squares), we now have the deterministic picture of a clockwork universe. But in the real world other factors enter in as depicted in Fig. 2b: Besides deterministic processes, there are processes that bring in an element of chance (indicated by dices). These processes are a consequence of quantum mechanics[4] and deterministic chaos.[5] Quantum mechanics only describes the probabilities that can be expected for the outcome of an experiment and probabilistic behavior appears to be immanent in nature.[6] Deterministic chaos, on the other hand, describes the behavior of systems based on classical mechanics[7] that reacts strongly on

[4] For a basic introduction to quantum mechanics see e.g. (Giancoli 2010, chap. 39).

[5] For an introduction on deterministic chaos see e.g. (Alligood 2009).

[6] Although the formalism of quantum mechanics is extremely successful in precisely calculating observable phenomena, its very nature is still under heavy debate, for details see e.g. (Ney 2013).

[7] Apart from deterministic chaos, it has been debated that determinism fails in classical mechanics anyway as argued e.g. by the example of Norton's dome, see e.g. (Norton 2008). Norton shows that a point mass located on the apex of a dome-like surface in a gravitational field may either stay at rest or spontaneously move down the surface after an undetermined amount of time. Due to its' several idealizations—especially that of a point mass—this case represents by far no realistic scenario and the validity of the case has been heavily questioned and lead critics to the question

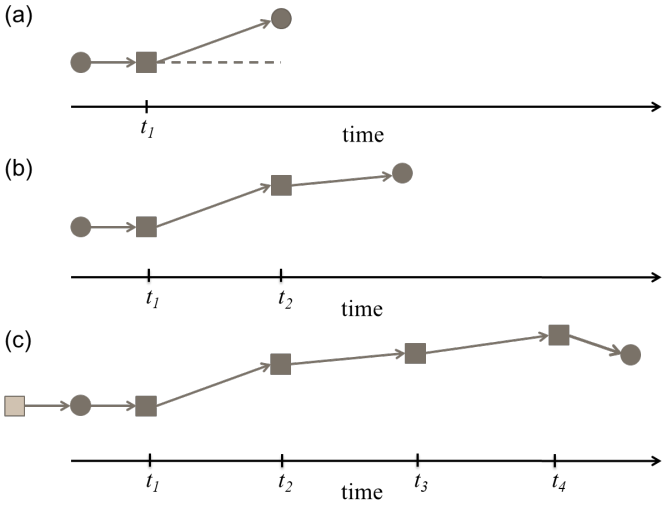

Figure 1: An object (sphere) moves from left to right on the axis of time. a) At time t_1 there is a cause (square) that changes the direction of the object. b) At time t_2 another cause results in another change of direction. c) At further points in time t_3, t_4 etc. there are further effects causing further changes of the movement of the object. The first cause (bright square to the left), however, cannot be explained in terms of physics, since a physical description requires initial and / or boundary conditions and the existence of suitable laws to start with. © R. B. Bergmann, Bremen, 2018.

minute changes of initial or boundary conditions. A well-known example is the daily weather pattern. The fact that weather cannot be predicted reliably over longer periods of time (say more than two weeks) is due to

of the metaphysical implications of such idealizations (Laraudogoitia 2013). Dantas just recently suggested a description that tries to adjust to a more realistic scenario. However, she leaves open the question of "what constitutes the fundamental (not only operational) conditions for establishing valid Newtonian systems, if there are any such conditions at all" (Dantas 2018). Although Newtonian mechanics is in many cases a good approximation for settings in the real world, the example makes abundantly clear that it is only an idealization!

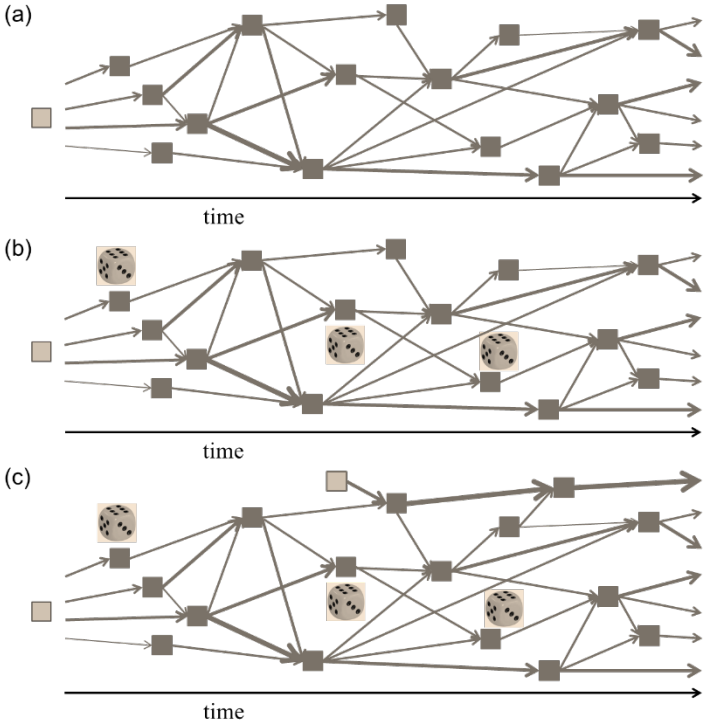

Figure 2: In the real world many processes happen simultaneously and interact with each other (for simplification, the objects which are influenced are omitted in contrast to Fig. 1.) a) The clockwork universe: Dark squares depict a network of processes interacting in a fully deterministic way upon each other. Bright square: First cause. b) The dices indicate the action of processes that incorporate probabilities c) Bright square at the top in the center of the drawing indicates divine intervention that changes or extends the existing net of causes and effects. © R. B. Bergmann, Bremen, 2018.

just this type of behavior and cannot be overcome even by the most powerful computers (or more powerful computers to come). I leave a discussion on the conditions and details of these phenomena to the literature, as the mechanisms described above suffice for the purpose of the present

paper. Consequently, the clockwork universe is a fiction and the Laplacian demon is dead![8]

Finally, I will extend the view developed so far by taking into account divine intervention, see Fig. 2c. In this simple picture, divine intervention introduces another cause into the causal network that alters the previously existing flow of cause and effects or brings about new effects (see the bright square in the upper middle).

2.2. Natural laws and equations of motion

Is divine intervention in contradiction with laws of nature? In order to answer this question, we will have to have a look as to what we call a law of nature and what these laws can predict. It is obvious that we would like a law of nature to be valid in a large variety of situations. As we do not have a "theory of everything" with the one law that describes all aspects of physics, we must content ourselves with the most general laws valid for a broad range of situations. In addition, there is usually not just one single law but a set of laws that is needed to describe a certain physical context, e.g. Newton's laws for classical mechanics or Maxwell's equations for classical electrodynamics.[9]

For the sake of intuitive clarity and simplicity, I will use classical mechanics as it was established by Isaac Newton and others as an example here. Take the law that states the gravitational force between two masses M and m (say the mass of the earth M and the mass of a satellite m) given by

$$F_G = \frac{G}{r^2} Mm \qquad \text{Eq. (1)}$$

[8] It wouldn't help trying to keep the Laplacian Demon alive by arguing that he would know all the initial conditions exactly and has unlimited computational power, since the accuracy of the knowledge of initial conditions is finally limited by Heisenberg's uncertainty principle.

[9] The basics of classical mechanics are described e.g. in (Giancoli 2010, chaps. 4–8) and the Maxwell equations as the basis for classical electrodynamics are e.g. described in (Giancoli 2010, chap. 32).

with the gravitational constant G and the distance r between them.[10] Whether we obtain a movement of the satellite e.g. revolving around the earth or falling on the earth, depends on the initial condition of the problem. Here, the initial condition describes the final velocity and direction the satellite is given by his propulsion system at the end of its launch. For other problems, imagine a vibrating string of a musical instrument, we further need boundary conditions, e.g. given by the fact that the string is tightly fastened at both ends.

To calculate the movement of a body with mass m under the influence of a force F one uses the well-known equation

$$F = ma \qquad \text{Eq. (2)}$$

from classical mechanics. The first step to determine the acceleration a of the body is to determine, which force F acts on it. It may be the gravitational force of Eq. (1), but it may also be another force or a superposition of several forces. Once we know all forces, and thus the resulting force F, we can write down the equation of motion. In a second step we need to know the initial conditions and for problems with two or three dimensions we also need to know the boundary conditions employed. Thus, natural laws don't come on their own but always need a context!

As a first result, we can state that the question whether a natural law on its own determines the motion of a body is "no." This result is in accord for example with much of the reasoning of von Wachter (2015) and will be further developed here.

To illustrate the situation, let us look at a very simple situation: If you drop an apple from a certain height one can calculate how it moves towards the ground and what speed it will reach, when it touches the ground. To solve this problem, we first have to determine the equation of motion. For

[10] Equation (1) is, however, only an idealized case for two point-masses and their gravitational attraction. For maximum simplification, I have also only introduced the scalar expression here in order to avoid further complications by using a vector notation. I also assume the masses not to be too small or two large and moving with a relative speed much smaller than the speed of light in order to avoid quantum mechanical or relativistic effects.

this purpose, we need to know the force that acts on the apple. This force is, close to the surface of the earth, given by

$$F_g = -mg \qquad \text{Eq. (3)}$$

with the earth's acceleration constant $g = 9.81\ \mathrm{m/s^2}$. (The minus sign is only a technical detail arising from the force being directed towards the ground.) To obtain the equation of motion, we equate the forces in Eq. (2) and Eq. (3) and obtain $ma = -mg$ or simply

$$a = -g \qquad \text{Eq. (4)}$$

which results in the equation of motion represented by the ordinary differential equation

$$\frac{d^2x(t)}{dt^2} = -g \qquad \text{Eq. (5)}$$

with the height of the apple $x(t)$ as function of time t.

In the second step, we must apply the initial conditions of the movement to Eq. (5) in order to calculate a specific solution and thus to obtain the trajectory of the apple. Assume you drop the apple from a height h then this is one of the initial conditions necessary to calculate the actual movement. If you just dropped the apple without giving it an extra initial velocity, then the second initial condition is the starting velocity of $v_0 = 0$. If you would, however, choose to throw the apple upwards with an initial velocity of $v_0 > 0$, the movement of the apple will obviously be different.

Solving Eq. (5) for the first case with the starting velocity of $v_0 = 0$ at the time $t_0 = 0$, the movement of the apple follows

$$x(t) = h - {}^1\!/_2\ g\ t^2 \qquad \text{Eq. (6)}$$

until it reaches the ground at $x = 0$. If you decide not just to drop the apple, but toss it up with a velocity $v_0 > 0$, the solution is given by

$$x(t) = h - {}^1\!/_2\ g\ t^2 + v_0\ t \qquad \text{Eq. (7)}$$

until it touches the ground at $x = 0$.

The way the apple moves in a specific situation is therefore not only determined by the corresponding natural laws, but also by the special conditions applied.

We now see that influencing natural (mechanical) processes does not violate laws of nature but only change the corresponding equation of motion or the solution of the corresponding equation of motion. At a later stage of this paper, I will apply this concept—that appears to be quite natural in the context of human action—to divine intervention.

2.3. Conservation of energy, invariance and symmetry

Before we come to the question of divine intervention, I want to discuss a commonly raised objection: Does divine intervention contradict the principle of energy conservation? In order to clarify this question, we need to look a little bit deeper into the idea of energy conservation. As e.g. Larmer (2009) points out, energy conservation can be defined in different ways. One way is to state that "Energy can neither be created nor destroyed," the second way is to state that "in an isolated system the total amount of energy remains constant." These two statements are, however, not equivalent, as the second can be deduced from the first, but the first cannot be deduced from the second. The second statement can be empirically tested and appears to be experimentally well justified. The first one cannot be tested and appears to be a metaphysical principle. Therefore, it appears reasonable to stick to the empirically well-established principle that the total amount of energy in an isolated system remains constant.

How can we then deal with a situation in which an entity from outside interacts with the system? In this case the system is obviously not isolated (or closed) any more. Usually, one deals with non-isolated systems by equating the change of the energy within the system with the difference of the flow of energy into or out of the system. Who or whatever interacts with a system has somehow to supply the energy required to bring about the observed effect. For the system or object to be influenced, it does not matter where the energy is coming from. Nature just follows its laws in combination with the applied forces, initial and boundary conditions or whatever constraints involved.[11]

[11] It has frequently been speculated that biological systems and especially the human brain is susceptible to quantum mechanical influences involving only a minute amount of energy or even zero energy. One of the earlier ideas were described by Popper and Eccles, see e.g. (Eccles 1994). Such mechanisms may allow mind-body

In order to gain a deeper understanding of the concept of energy conservation, I want to point out that energy conservation is not only an empirically well-established concept, but is related to certain types of symmetries and invariances. To discover theses relations, one must use an alternative approach to Newtonian mechanics, the so-called Hamilton-Lagrange-formalism.[12] The interesting aspect of this approach is that it allows in addition to the calculation of equations of motion also to determine which quantities are conserved. Conservation is not limited to energy but may also apply to other quantities and is related to corresponding symmetries and invariances.

Table 1 gives the most prominent examples of the conservation of quantities in the context of classical mechanics and their related symmetries and invariances.[13] The fact that, in the framework of classical mechanics, the same experiment carried out under the same conditions at different times t_1 and t_2 leads to the same result reflects the homogeneity of time and therefore energy conservation. If, however, experimental conditions between times t_1 and t_2 change, e.g. by an external intervention, the situation is not any more invariant against translation in time and the energy involved in the experiment at the two points in time can of course well be different. As can be seen from this simple example, conservation of energy as a consequence of symmetries and invariances cannot safeguard against the possibility of external intervention of whatever kind but rather relies on unchanged experimental conditions between different points in time. This argument of course also holds for the other quantities stated in table 1.[14]

---

interactions, but a discussion of these theories is certainly of interest not only for dualists (Collins 2008). More recently such interactions are discussed in a more general context of top-down causation (Ellis 2016). I will, however, leave this point open, since this discussion is far beyond the scope of this paper.

[12] As the derivation of this approach is too involved for the sake of this paper, I refer to the literature. For the underlying concepts see e.g. (Lanczos 1986).

[13] There are more of such relations beyond classical mechanics. These are, however, based on symmetries mathematically much more abstract than the ones shown here.

[14] The situation is further complicated by the fact that the concept of a well-defined energy cannot be naively transferred beyond classical mechanics. Due to Heisenberg's

Table 1: Conserved quantities and related symmetries and invariances required for their conservation.

conserved quantity	symmetry	invariance against
energy	homogeneity of time	translation in time
linear momentum	homogeneity of space	translation in space
angular momentum	isotropy in space	rotation

3. How could divine intervention work?

3.1. Divine intervention

Looking at stories in the Old and New Testament of the Bible, there are a great variety of examples for divine intervention. In order to relate to the characteristics of mechanical systems described above, I will choose a very simple example of transcendent intervention stated in the New Testament. In the Gospel of Mathew there is a passage that relates to the situation after the crucifixion of Jesus. It is reported that two women wanted to look for the grave of Jesus and experienced an earthquake. This earthquake is reported to be a result of an angel descending from heaven, who moved the stone that was covering the entrance of the grave and sat on it, see Mathew 28, 2. The result of the intervention of the angel is that the women could see that the grave was empty.

uncertainty principle there is an uncertainty in the energy involved in certain quantum mechanical processes. As an example, consider a Laser that emits ultrashort light pulses (e.g. of a duration of only several femtoseconds). In this case, an individual photon can have a substantial uncertainty of its energy and thus wavelength [see e.g. (Giancoli 2010, chap. 39)]. We also encounter a problem in general relativity. Here the concept of energy must be extended and there seems to be no general energy-momentum conservation equivalent to classical mechanics [see e.g. (Bondi 1990) and (Hoefer 2000)]. These findings may suffice to indicate that the naïve picture of classical mechanics that deeply shapes our intuition about the world we live in only represents a small fraction of how the world around and in us functions.

At this point, I do not want to enter a theological debate about the details or trustworthiness of this particular story, but rather consider how such an event may have proceeded. If we assume that transcendent intervention follows the same path as described above in chapter 2, then the stone is moved by a force according to the laws of classical mechanics, no matter where the force came from. Therefore, if God or angels intervene in the world by applying forces or by changing certain conditions, it appears not to be necessary to violate natural laws. Rather, in the sense described in Fig. 2c, new causes are introduced that lead to results that otherwise would not have happened.

In miracles exceeding effects such as those described above, God may create new entities which he introduces into the world. The very act of creating new entities appears to be outside the scope of natural science. However, once such entities are introduced, everything follows natural processes.[15] Other kinds of miracles may also concern other domains of physics and it would be interesting to see, whether there are typical patterns accompanying certain types of miracles. The fact that we can only see the effect of divine intervention may feel uncomfortable or threatening. Nevertheless, claiming divine intervention to be unthinkable or impossible seems to be unjustified.

In addition to the discussion above there are a huge number of observations that indicate that such interventions have really happened and are still happening. Craig S. Keener published an extensive systematic investigation of miracles throughout ancient and recent history in and outside Christianity in his two volume book „*Miracles*“ (Keener 2011) and also discusses various explanations and criticism on reports of miracles. He concludes: “Many healing claims involve blindness, inability to walk and even raisings from the dead; other claims involve sudden changes in nature after prayer. Despite some debatable instances, some other cases are fairly clearly extraordinary. It seems to me that to dispute that such phenomena have sometimes occurred is not really possible to open minded people” (Keener 2011, 599). His viewpoint is supported by other contemporary writers, see e.g. Metaxas (2015), who points out, that miracles always have an element

[15] The classical example from the New Testament would be the pregnancy of Mary, see Math. 2, 18f.

of message or relation to the individuals experiencing them. As a physicist I would like to add that observations have to be taken seriously and should not be discarded without good reason. This does not mean to discard careful investigation in the trustworthiness of reports on miracles or not to investigate the risk of delusion or fraud. However, a general denial of the occurrence of miracles appears unjustified and unscientific and impedes the openness for new scientific hypotheses and the cross-fertilization of different disciplines of science.

3.2. The role of probabilistic processes

Until now I have made no use of probabilistic processes such as quantum mechanics or deterministic chaos with respect to divine intervention. I do not assume that God "hides" behind or depends on quantum indeterminacy or the like. However, chance plays an important role with respect to another aspect: The picture I have developed in the last chapter has significant consequences for the possibility to determine possible causes that led to certain observations. In a purely deterministic world the Laplacian demon described above must be able to derive any previous or future state of affairs. That is, however, not the case in our real world. Take an everyday example: There are many board games that depend on a dice to determine the progress of a player. Suppose you have a documentation of the sequence of each dice thrown, you will be able to reconstruct the course of the game (given that everybody sticks to the rules). If you, however, only see the status of the game at a certain point, you will, in many situations, not be able to reconstruct the sequence of events that led to the current situation. As we do not live in a clockwork universe, it is in many cases not possible, to determine what sequence of events has brought about the result that we now see as an observer. This does not mean that the outcome of processes that include probabilistic mechanisms is always open. The interplay of deterministic and probabilistic mechanisms may well lead to a definite result, as you may observe by playing certain board games right to the end.

As a consequence it may be difficult or even impossible to determine whether a certain event was caused by divine intervention or just by natural processes! Here, the influence of probabilistic processes as described above

indeed become important. A general discussion of this problem, e.g. including theological arguments, would be far beyond the scope of this paper. However, sometimes an answer may be at hand, as can be seen from the following illustration: Imagine, you believe in God and believe that he can intervene in the world and have (foolish enough) engaged yourself in Russian roulette. If you win, you'll have no financial trouble any more, if you lose, the same is true, but under fairly different circumstances. The other players in the game are atheists who only believe in blind chance. You desperately pray that you will survive, and you do survive and attribute your survival to God's intervention. The others around you believe that this outcome is just good luck. Are you able to decide who is right? But imagine, you prayed—say in the presence of the people who joined you for Russian roulette—that a man who just fell victim to Russian roulette rises from the dead. If this were indeed to happen, could one reasonably assume that this just happened by chance?

As a further result, I conclude therefore that based on the complexity of the causal net described above and the presence of probabilistic processes, the causes that led to a certain effect can often not be uniquely determined. A clear distinction between the effect of transcendent intervention and mere chance may therefore not be possible in many cases of interest.

4. Conclusions

The discussion of laws of nature using the framework of classical mechanics shows that the motion of objects is not determined by these laws on their own. To derive equations of motion and their solutions, the forces involved as well as initial and boundary conditions need to be known. Influencing natural (mechanical) processes therefore does not violate laws of nature but only changes the corresponding equations of motion and their solutions by introducing further forces or altering initial or boundary conditions or introducing new entities in the process. Therefore, if God intervenes in the world, new causes are introduced that lead to results that otherwise would not have happened. Further, the objection of energy conservation against divine intervention is not valid, since the total amount of

energy remains constant only in an isolated system, not in a system that is subject to external influences.

Processes in the world are embedded in a dense net of causes and effects that incorporates deterministic as well as probabilistic mechanisms, the latter being influenced by quantum mechanics and deterministic chaos. In many cases, it may not be possible to uniquely determine the events that caused a certain effect or to identify divine intervention.

Acknowledgements

I thank Daniel von Wachter, head of the International Academy of Philosophy (IAP) in the Principality of Liechtenstein, for inviting me as a physicist to the conference "Causality, Free Will, Divine Action" in September 2017 in Vienna sponsored by the Templeton Foundation in the framework of the Project "The Openness of the Universe for Free Will and Special Divine Action" and giving me the opportunity to present the ideas discussed here to a philosophical audience.

References

Alligood, Kathleen T., and James A. Yorke. 2009. *Chaos: An Introduction to Dynamical Systems*. New York: Springer. https://doi.org/10.1007/b97589

Bondi, Sir Hermann. 1990. "Conservation and Non-Conservation in General Relativity." *Proc. R. Soc. Lond. A* 427 (1873): 249–58. https://doi.org/10.1098/rspa.1990.0011

Bultmann, Rudolf. 1960. *Neues Testament und Mythologie*. Hrsg. von Hans-Werner Bartsch. 4. Aufl. Hamburg: Herbert Reich, Evangelischer Verlag GmbH.

Bussey, Peter J. 2013. "God as a First Cause—a Review of the Kalam Argument." *Science and Christian Belief* 25 (1): 17–35.

Collins, Robin. 2008. "Modern Physics and the Energy-Conservation Objection to Mind Body Dualism." *American Philosophical Quarterly* 45 (1): 31–42.

Craig, William Lane. 2002. "The Kalam Cosmological Argument." In *Philosophy of Religion. A Reader and Guide*, edited by William Lane Craig, 92–113. Edinburgh: Edinburgh University Press.

Dantas, Christine C. 2018. A Note on Norton's Dome. arXiv:1801.01719v2 [physics.class-ph] 29. Jan. 2018.

Eccles, John C. 1994. *How the Self Controls its Brain*. Berlin, Heidelberg, New York: Springer. https://doi.org/10.1007/978-3-642-49224-2

Ellis, George. 2016. *How Can Physics Underlie the Mind? Top-Down Causation in the Human Context*. Berlin Heidelberg: Springer. https://doi.org/10.1007/978-3-662-49809-5

Feuerbach, Ludwig. 2016. *Das Wesen des Christentums*. Altenmünster: Yazzybee Verlag Jürgen Beck. (First published in Leipzig: Otto Wigand 1841.)

Giancoli, Douglas C. 2010. *Physics for Scientists and Engineers with Modern Physics*. 3rd Ed. München, San Franzisco et al.: Pearson.

Härle, Wilfred, ed. 2007. *Grundtexte der neueren evangelischen Theologie*. 3. Aufl. Leipzig: Evangelische Verlagsanstalt.

Hawking, Stephen, and Leonard Mlodinow. 2010. *The Grand Design—New Answers to the Ultimate Questions of Life*. London: Transworld Publishers.

Hoefer, Carl. 2000. "Energy Conservation in GRT." *Stud. Hist. Phil. Mod. Phys.* 31 (2): 187–99.

Keener, Craige S. 2011. *Miracles. The Credibility of the New Testament Accounts*. Grand Rapids: Baker Academic.

Lanczos, Cornelius. 1986. *The Variational Principles of Mechanics*. 4th Ed., Dover Books on Physics and Chemistry. New York: Dover Publishing.

Laplace, Pierre Simon. 1814. *Essai philosophique sur les probabilities*. Paris. (Newly printed Brussel 1967.)

Laraudogoitia, Jon Pérez. 2013. "On Norton's Dome." *Synthese* 190 (14): 2925–41. https://doi.org/10.1007/s11229-012-0105-z

Larmer, Robert A. 2009. "Divine Agency and the Principle of the Conservation of Energy." *Zygon* 44 (3): 543–57. https://doi.org/10.1111/j.1467-9744.2009.01015.x

Larmer, Robert A. 2014. *The Legitimacy of Miracle*. Maryland: Lexington Books.

Lewis, Clive Staples. 1947 [2016]. *Miracles*. London: William Collins.

Krauss, Lawrence M. 2012. *A Universe from Nothing*. New York: Free Press, Simon and Schuster Inc.

Mackie, John Leslie. 1982. *The Miracle of Theism: Arguments for and against the Existence of God*. Oxford: Clarendon Press.

Metaxas, Eric. 2015. *Miracles*. New York: Dutton.

Ney, Alyssa, and David Z. Albert, eds. 2013. *The Wave Function—Essays on the Metaphysics of Quantum Mechanics*. Oxford: Oxford University Press.

Norton, John D. 2008. "The Dome: An Unexpected Simple Failure of Determinism." *Philosophy of Science* 75 (5): 786–98. https://doi.org/10.1086/594524

Schleiermacher, Friedrich. 1969. *Über die Religion. Reden an die Gebildeten unter ihren Verächtern*. Stuttgart: Reclam. (First published in Berlin: Unger 1799.)

von Meckenstock, Günther, Andreas Arndt, Jörg Dierken, Lutz Käppel, and Notger Slenczka, eds. 2011. *Friedrich Schleiermacher: Kritische Gesamtausgabe*. Berlin: De Gruyter.

von Wachter, Daniel. 2015. "Miracles Are Not Violations of the Laws of Nature Because the Laws Do Not Entail Regularities." *European Journal for Philosophy of Religion* 7 (4): 37–60. https://doi.org/10.24204/ejpr.v7i4.86

Wilckens, Ulrich. 2017. *Theologie des Neuen Testaments. Band III Historische Kritik der historisch-kritischen Exegese*. Göttingen: Vandenhoeck und Ruprecht.

Organon F 26 (1) 2019: 104–121
https://doi.org/10.31577/orgf.2019.26107

ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

RESEARCH ARTICLE

Elements of a Theory of Nonphysical Agents in the Physical World

Uwe Meixner*

Received: 17 May 2018 / Accepted: 3 December 2018

Abstract: This paper shows that there is a quantum-physical and evolution-biological perspective for (libertarian) free will, and that the so-called scientific arguments against it are in reality metaphysical arguments and insufficient. The paper also develops the idea of a nonphysical organ of higher organisms: the *Domindar* (Detector of macroscopic indetermination, and restrictor).

Keywords: Brain; chance; choice; consciousness; determinism; Domindar; indeterminism; nonphysical cause; quantum physics; principle of causal closure; principle of sufficient cause; subject of consciousness and of physical action.

1. The present situation, and a plan of action

In recent years things have not looked good for human 'free will.' Based on neurophysiological facts, its existence has been denied, or in other words: it has been denied that we ever are in the full and proper sense originators of our own physical actions. In this paper, I propose to show that the scientific (quantum-physical and evolution-biological) facts do not offer a sufficient

* University of Augsburg
Department of Philosophy, University of Augsburg, Universitätsstr. 10, D-86159 Augsburg, Germany
uwe.meixner@phil.uni-augsburg.de

© The Author. Journal compilation © The Editorial Board, *Organon F*.

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International Public License (CC BY-NC 4.0).

reason for a denial of 'free will.' On the contrary, those facts strongly suggest that we (and the subjects of other higher animals) do at least sometimes *act*, in the full and proper sense, in the realm of the physical, that is: they strongly suggest that we make genuine (hence also truly decisive) decisions between alternatives of being, between ways for the physical world to be. We have something to decide, and we exist for deciding something.

2. An important distinction: indeterminism and resolution of indeterminism

I begin with a fundamentally important distinction. Physical indeterminism and *the resolution of* physical indeterminism are two different things. *Physical indeterminism* is given at a time t if, and only if, beginning with time t there are—the completed physical past up to t notwithstanding—several physically possible further courses of the physical world (at least two, if not more). That physical indeterminism occurs at some times, perhaps at all times, is today a widely accepted lesson from quantum physics. This lesson, however, is not a necessary or logical consequence of quantum physics. It certainly is a matter of its interpretation—of an interpretation of quantum physics, however, which is natural and plausible and which has almost completely prevailed against Einsteinian dreams of a restoration of physical determinism, and against the Bohmian alternative interpretation which safeguards the old determinism. In what follows, I presuppose the existence of physical indeterminism.

Now, *resolution of physical indeterminism* always occurs when physical indeterminism occurs; for always the physical world continues in a determinate way: in precisely one of the ways physically possible in the indeterministic situation.[1] Always the physical indeterminacy is replaced by determinacy. There are two possible manners of a resolution of physical

[1] Some philosophers favor the so-called *many-worlds interpretation of quantum physics* [see (DeWitt and Graham 1973)], according to which *all* the ways which are physically possible in a given indeterministic situation are actualized. I do not believe that this view has enough metaphysical, let alone empirical, warrant to be credible. It brings along with it altogether too many unverifiable actualities, and mainly, it

indeterminism, of the replacement of physical indeterminacy by determinacy: *one*, by ontic chance, hence without any sufficient cause; *two*, by a sufficient nonphysical cause. Let us take note: The existence of physical indeterminism does not logically entail the existence of ontic chance in the physical world; for physical indeterminism and the resolution of physical indeterminism are, on the one hand, two different things, and on the other hand, the resolution of physical indeterminism need not always, or even sometimes, come about by ontic chance.

3. Two different metaphysical principles – two different outcomes

It is *not* a logical consequence of quantum physics that physical indeterminism occurs; it is, even more emphatically, *not* a logical consequence of quantum physics that the resolution of physical indeterminism always occurs via ontic chance. This latter consequence follows only if one adds a principle to quantum physics which is not a principle of physics but a principle of a specific kind of metaphysics: *the principle of the causal closure of the physical world*, which principle says that every physical event that has a sufficient cause *at all* also has a sufficient physical cause.[2] Assume there is

seems, in order to get rid of the wonder that is provoked when one sticks to the (truly) empirical facts, the facts of appearance: As far as we know from appearances, *just one* of the physically possible ways is actualized out of every indeterministic situation.

[2] This is a (logically) weak version of the closure principle; it follows logically from (logically) stronger versions of that principle. A stronger version which is so strong as to be directly refuted by the admittance of ontic chance into the physical world is this: *Every physical event has a sufficient physical cause.* Another stronger version, however, is not refuted in this manner: *Every sufficient cause of a physical event is itself physical.* In (Kim 1993, 280) the following version of the "causal closure of the physical domain" can be found: *Any physical event that has a cause at time t has a physical cause at t*, which can be interpreted in such a manner as to be roughly equivalent with the above-presented weak version. Its many versions notwithstanding, in (Popper and Eccles 1977, 51) what needs to be said about the principle of causal closure *in the first place* has already been said: "[T]he physicalist principle of the closedness of the physical [world] [...] is of decisive importance, and I take it as the

physical indeterminism at t; hence beginning with time t there are *several* physically possible *further courses* of the physical world, the entirety of the physical past up to t notwithstanding. Precisely one of these further courses becomes actual. How does this come about? The answer is, prima facie, unclear. What is clear, however, is this: in whichever manner the resolution of the physical indeterminism at hands comes about, there is, in any case, a physical event beginning at time t without sufficient physical cause—since even the entire physical past up to t is not a sufficient physical cause of it. On the basis of the principle of the causal closure of the physical world, it follows for such an event that it has no sufficient cause at all, hence that it happens by ontic chance.

Unfortunately—or fortunately (depending on one's metaphysical point of view)—this conclusion is by no means indisputable. For if, instead of the principle of the causal closure of the physical world, one assumes *the principle of sufficient cause—Every event has a sufficient cause*[3]—as a metaphysical addition to quantum physics, then it follows that the event in question—the event beginning at time t without sufficient *physical* cause—does have a sufficient cause, but, of course, a *nonphysical* one.

4. A brief assessment of conflicting principles

The principle of the causal closure of the physical world has no greater right to be believed in than the principle of sufficient cause, neither systematically nor historically. On the contrary, the principle of sufficient

characteristic principle of physicalism or materialism." This principle belongs to *metaphysics*, not to physics.

[3] The *principle of sufficient cause* is not the—better-known—*principle of sufficient reason*: *For every existing entity there is a sufficient reason of its existence*. The principle of sufficient cause is, however, a logical consequence of the principle of sufficient reason; this is so in virtue of the fact that every event is an existing entity, and that if there is a sufficient reason for the existence of x, that then there also is a sufficient cause for the existence of x (for this to be true, it *need not* be true that every reason is a cause). A near relative of the principle of sufficient cause is relied on in William Craig's modernizations of the cosmological argument; see, for example, (Craig and Sinclair 2012).

cause is metaphysically more neutral, hence in a certain sense more rational than the closure principle (not long ago, the principle of sufficient cause still enjoyed the dignity of a quasi-logical principle), and historically it is, of course, by far the more respected postulate. Mainly one reason seems to speak for the principle of the causal closure of the physical world, and it is not a rational reason: those who believe in it have an *intellectual fear* of nonphysical causes of physical events, they even fear a nonphysical *influence* on the physical. Even theologians, nowadays, seem to fear a causality of the nonphysical that touches the physical world; such is the modern mentality. One cannot well say—although many do say it—that there is no empirical evidence for nonphysical causes of physical events; for if one assumes the principle of sufficient cause instead of the principle of causal closure, then the interpretation of the physical facts—of the quantum-physical facts—will be quite different than it was before: where previously one saw physical events without any sufficient cause, one now sees physical events with a sufficient nonphysical cause. Of course, strictly speaking, one does not 'see' either the one or the other; strictly speaking, one judges the empirical situation differently by adhering to different metaphysical interpretations of it.

The principle of sufficient cause requires that the resolution of physical indeterminism *always* occur by a nonphysical sufficient cause; the principle of the causal closure of the physical world requires that this resolution *always* be by ontic chance, without any sufficient cause. In what follows, I proceed on the assumption that some, even many, resolutions of physical indeterminism occur by nonphysical causes; in doing so, I position myself against the principle of causal closure, consider it false—without, however, wishing to wholly exclude ontic chance as a means of the resolution of physical indeterminism. The principle of sufficient cause, too, is, perhaps, false, and what follows below does not rely on the truth of the principle of sufficient cause.[4]

[4] What follows below is, moreover, essentially different from the proposal of the physicist Eugene Wigner. Wigner did assume an influence of consciousness on the physical; he, however, did not assume an influence of *the will* on the physical. On Wigner's proposal, see (Barrett 2014, 67–70); according to Wigner, "conscious apprehension causes collapses [of the wave function; of the quantum-physical

5. Indeterminism and the agency of natural living beings

What is the import of all this for the agency of natural living beings? The agency of natural living beings is their agency in the physical macro-world. The presupposition for the existence of such agency is not only the existence of physical indeterminism *simpliciter*, but also the existence of physical indeterminism *in the physical macro-world*, to boot, of such indeterminism as is *detectable* by natural living beings. Only if there are situations with alternative possibilities—indeterministic situations—in the physical macro-world and only if they can be detected by natural living beings—only then can such beings contribute to the resolution of such situations (for example, in order to obtain a biological advantage). In this, precisely, does the agency of natural living beings consist: in contributing to the resolution, the ontic deciding, of *macrophysical* indeterministic situations—normally, with the aim of obtaining a biological advantage.

What are the indications that there are indeterministic situations in the physical macro-world which can be detected by natural living beings? What are, in other words, the indications that for some times t the following is true: there are several physically possible further courses of the physical *macro*-world beginning with t, and in such a manner that this situation can be found out by a natural living being?

One indication is that the physical macro-world strikes us—and perhaps not only us—as being to a considerable extent contingent. In every reflective second of our waking life we have, looking back at the past of the physical macro-world, the consciousness that things could have been otherwise than they actually were, that we, in particular, could have acted differently than we did; looking ahead at the future of the physical macro-world, we have the consciousness of being able to act *like this, or otherwise*; more generally speaking, we have the consciousness that things can be *like this* in the future*, or otherwise*. Those who assume determinism for the physical macro-world must consider the apparent contingency of the physical macro-world as one gigantic, incessant illusion—an illusion which cannot be eliminated; which can only be uncovered. But how plausible, really, is this stance? Why

state]" (Barrett 2014, 70). As will become clear below, this is not the view here proposed.

the consciousness of contingency, why even consciousness at all, if the physical macro-world is ruled by determinism? If it is ruled by determinism, consciousness is pure luxury, extravagancy. Sometimes, indeed, it does happen that biological evolution treats itself to a luxury, but not to a luxury which has such a high price: the large amount of energy that the central nervous system consumes in the production of *consciousness*, including the insistent, imperturbable *consciousness of contingency*.

Another indication of the existence of indeterministic situations in the physical macro-world is this: there seem to exist in the physical macro-world goings-on that can be correctly described as 'competition,' 'struggle,' 'fight.' Who believes in determinism in the physical macro-world must, however, deny the existence of such goings-on. True fights cannot exist for a determinist. Two soccer teams, it is said, fight for victory. Not so; if determinism rules in the physical macro-world, then who will win is determined even before the two teams begin to play—and, rightly considered, the verb 'to win' must here be put in scare quotes. Two males, it is said, fight over a female. Not so; if determinism rules in the physical macro-world, then what happens in this latter case is, as in the case of the soccer game, only the masquerade of a fight, only the masquerade of a competition. Everything is already determined, everything has already been decided beforehand, nothing is open anymore—if, *if* indeed, it is really the case that determinism rules in the physical macro-world. There are only 'rigged games' then, and also the very serious game of life, the so-called 'struggle for existence,' is a 'rigged game' then, not a true struggle; it is, then, a bad, not to say *evil*, joke, considering that, for the antagonists in that game, it feels exactly as if it weren't 'rigged,' exactly as if it were an open struggle and they were *fighting*, *really fighting* for a good outcome for themselves.

6. Micro- and macro-indeterminism, and the brain

It is not only morality which is rendered absurd by determinism in the physical macro-world, it is biology, too (and, by the way, philosophy as well; for what is the point in discussing anything at all if all our voiced opinions, whichever they may be, are already determined to be voiced by us even before we began to exist?). And this is nothing less than a *reductio*

ad absurdum of physical macro-determinism—albeit, as we all know, not a generally accepted *reductio*. Indeterminism in the physical *micro*-world, and hence also, as a logical consequence, physical indeterminism *simpliciter*, do indeed enjoy widespread credence—due, firstly, to the almost general acceptance of quantum physics and, secondly, to the prevalence of its standard interpretation. But indeterminism in the physical *macro*-world is still widely rejected. With respect to the physical *macro*-world, determinism of the Newtonian-Laplacian stamp is still widely accepted; the explanation usually given is this: microphysical indeterminism cannot play any macrophysical role, cannot magnify itself into the macrophysical realm because of the massive and unavoidable disturbances produced by the natural environment.

It is, therefore, for many people as if there were indeed *two parallel* physical worlds: *one* microphysical, in which indeterministic situations occur, as can be found out by complicated procedures; and *one* macrophysical, in which no indeterministic situations occur; in which, in any case, none—it is said—are scientifically detectable. In fact, this latter position is nowadays almost invariably assumed by *one* of the metaphysically interested sides: by those who have a materialist-naturalistic metaphysical outlook. Invariably they hold that indeterministic situations in the physical macroworld are scientifically undetectable, and invariably this is taken as evidence for there being no such situations. However, assuming for the sake of the argument that there are such situations, how *could* they be scientifically detected? They would have to be detected after they are already over, after they have already been resolved into determinacy: by proving—that is, by ascertaining beyond scientific doubt—that certain physical macro-events have no sufficient physical cause. Now, one encounters physical macroevents in huge numbers which, in fact, *do not seem* to have a sufficient physical cause—chaotic goings-on of the most various kinds, so-called chance processes, among them the familiar die-throws and coin-throws, with their concluding events that serve, for the purposes of everyday life, as chance events. Yet, the step from 'seems not to have a sufficient physical cause' to 'does not have a sufficient physical cause' is *never ever*, under no circumstances, taken by the metaphysically prejudiced—because they suppose *a priori* that there must be a sufficient physical cause even though

none is apparent. Thus, endeavoring to prove scientifically that there are physical macro-events without sufficient physical cause—in order to prove that there are indeterministic situations in the physical macro-world—encounters in the minds of many people from the start an insurmountable obstacle: *a priori*, and quite unscientifically, such events and such situations are not accorded a fair chance of existing.

They should be given such a chance. Is it not to be expected that a large microphysical indeterminism—one that involves large numbers of elementary particles—will issue into a macrophysical indeterminism? Is it not to be expected that on intricate and involved paths even a small microphysical indeterminism will lead to a macrophysical one? We are, after all, *not* dealing with two parallel physical worlds, we are dealing only with *one physical world*; the distinction between *microphysical* and *macrophysical* has no separating significance. This *one physical world* is ruled by the laws of quantum physics, not by those of nineteenth-century physics. Brain scientists like to emphasize that brain processes obey without exception the laws of physics. This is, of course, *true*; for brain processes are physical processes. But the laws they obey are the laws of quantum physics, not the laws of the old, deterministic physics.

Therefore, also in the brain we must expect to encounter physically spontaneous physical events: physical events without a sufficient physical cause; and that some such events—*as brain processes* (not already as processes in the brain: most *processes in the brain* are microphysical)—would be physical macro-events, would *still* be physical macro-events. Physically spontaneous physical macro-events in the brain would, however, be the unmistakable indicators of the occurrence of brain-centered *macrophysical* indeterministic situations.

7. The inconclusiveness of the Libet-experiment

Some time ago, the result of the so-called Libet-experiment produced considerable reverberations. The readiness potential for a physical action, which is given in the brain already some time before one is conscious of deciding 'to do the deed,' was regarded by some as a refutation of free will. 'Look here,' they rejoiced (the verb 'to rejoice' is not as inappropriate as it

may seem at first sight), 'before you believe that you are deciding, the brain has already decided. You didn't decide anything, or in any case you did not *genuinely* decide anything, for your decision, coming too late, was, in any case, not the truly decisive one.' What seems to have been deigned attention only seldom, if at all, was the question of what produces the readiness potential in the first place. So eager was one to consider the brain a deterministic automaton that hardly anybody—perhaps nobody, I don't know—considered the possibility that the brain processes which contribute to the readiness potential are physically spontaneous processes—that is, physical events without a sufficient physical cause. One has not excluded this possibility, and I, for one, wouldn't know how it could be excluded. If the brain processes in question are in fact physical events without a sufficient physical cause, then there are, concerning their being caused or not (their being *made to come about* or not), only two basic options: either they do not have any sufficient cause, or they do have a nonphysical sufficient cause. In both cases, the occurrence of a brain-centered macrophysical indeterministic situation is indicated. As long as physics remains *pure natural science*, physics excludes neither one of the two mentioned options. Physics, as long as it stays free of metaphysics, gives both ontic chance and nonphysical causality a chance: in general, and in particular with respect to brain processes. Nonphysical causality, however, is what must interest us when considering *agency*; for agency is not a matter of chance.

8. Subjects of consciousness and of physical action

Let us consider a hypothesis, or rather, a network of hypotheses.—The waking subject of consciousness of a natural consciousness-endowed living being is, at the same time, the subject of physical action of that living being. Consciousness—among other things it offers in the line of service—points out to the waking subject, directly in perception (against the background of past experience), and by and large correctly, macrophysical indeterministic situations, especially such on the resolution of which the subject can exercise some influence. Moreover, consciousness proposes to the waking subject 'guidelines' (usually called 'motives') for the exercise of its influence, in case

that influence is actually exercised: appetitive or aversive emotions and sensations, (conscious) needs and drives. But these 'guidelines,' even if they are considered in their entirety, are at least sometimes *non-determinative*: the subject remains—at least sometimes—up to a certain degree the sovereign of its physical actions. If it were not so, subjects of consciousness and of physical action would be biological superfluities. However, a subject of consciousness and of physical action that belongs to a natural living being (in what follows, I exclusively consider subjects of consciousness and of physical action *that belong to a natural living being*) is so far from being a biological superfluity that it is a nonphysical organ of its organism: an organ which serves, like every organ of the organism, the organism's life, especially its preservation. Consider what happens if this nonphysical organ goes temporarily out of service, or does not fulfill its tasks satisfactorily: the organism is rendered less protected, or is left entirely unprotected.

The service, then, that a subject of consciousness and of physical action renders to the life of its organism consists in this: to the advantage of its organism, it contributes to the resolution of some of the macrophysical indeterministic situations it detects and on the resolution of which it can exercise an influence. The agency of a subject of consciousness and of physical action is, precisely, the restriction of macrophysical indeterminism. Consider that every macrophysical indeterministic situation consists in a set of several possibilities: the physically possible further courses of the macrophysical world, beginning with a time t. From some such sets a subject of consciousness and of physical action selects a non-empty proper subset. And if a subject of consciousness and of physical action resolves a macrophysical indeterministic situation already *all by itself*, then it is precisely a singleton set which that subject selects from the pool of possibilities. In any case, the subject of consciousness and of physical action exercises a causal influence, a greater or smaller one. Its choice has an ontic effect, and it is not a blind choice: the subject of consciousness and of physical action chooses in the light of its consciousness, in the light of its outer and inner perceptions, in the light of its 'guidelines of the will,' in the light of its rational considerations (if there are such). Often its choice is, nevertheless, to a certain degree arbitrary, sometimes entirely arbitrary. But arbitrariness does not turn choice into ontic chance. Ontic chance and choice exclude each other—even

if the choosing is arbitrary. If what I do is my choice, I am choosing what I do; if what I do is an ontic chance, *nobody* is choosing what I do. And if my choice is completely non-arbitrary (because consciousness clearly and distinctly indicates *what is best* for the best of reasons, and because I, a *rational* subject of consciousness and of physical action, follow this indication unhesitatingly and unwaveringly), then the perfect rationality of my choice does not turn my choice into a case of passive determination. Passive determination, too, and choice exclude each other. If what I do is my choice, I am choosing what I do; if what I do is due to passive determination, I am not choosing what I do; for under passive determination—not determining, but being determined—I do not have a choice.

The agency of the subject of consciousness and of physical action for the organism does not take place beside or behind the organism; it is not as if such a subject were a sort of guardian angel for the organism. Rather, the subject of consciousness and of physical action emerges together with its consciousness and its powers from the nervous system of the organism, mainly from the brain, and it is in its existence and in many details of its existence *nomologically bound*—bound by the laws of nature—to neuronal functions. It is, however, not determined by the nervous system in every respect; for that would mean that it is an epiphenomenon of the neuronal and hence a biological superfluity; but nature is not fond of superfluities. Rather, the subject of consciousness and of physical action is a highly useful—biologically useful, hence evolution-favored—**d**etector **o**f **m**acroscopic **ind**etermination, **a**nd **r**estrictor, in short: a *Domindar*.[5]

[5] The idea of Domindars is developed and justified in (Meixner 2006) and (Meixner 2008). That Domindars—subjects of consciousness and of physical action—emerge from organisms, in particular, from their nervous systems, is far from providing an ultimate metaphysical explanation of them. The main issue in their ultimate metaphysical explanation would be whether the emergence of Domindars is natural *and without a supernatural grounding*: is effected on the basis of *uncreated* natural laws and circumstances; or is natural *and with a supernatural grounding*: is effected on the basis of *God-created* natural laws and circumstances. My sympathies are with the latter, but one certainly needn't be an orthodox theist to believe in the emergence of Domindars. A wholesale rejection of psycho-physical emergentism—"The physical cannot beget the nonphysical! In any case, it is incomprehensible

The true relationship between the organism's subject of consciousness and of physical action—the Domindar—and the organism's nervous system can be illustrated *cum grano salis* by an analogy. A pianist is not able to play the piano without a piano, and if she plays the piano, then much in her playing is determined by the piano she plays on. But not everything in the pianist's piano-playing is determined by the piano played on: not which pieces of music the pianist plays, and not which interpretation she accords to them; also not how well the pianist plays—insofar as this is *up to her*. The piano is a necessary instrument of the pianist qua pianist; the nervous system—with the brain at its center—is a necessary instrument of the Domindar qua Domindar. Without their respective instruments they cannot—for necessary, compelling reasons—do what is their calling. However, this leaves entirely undiminished the fact that the pianist cannot be reduced to a piano, and that the Domindar is irreducible to a nervous system.

At one point, especially, this analogy is not perfect: a pianist is not a product—let alone a nonphysical product—of the instrument she plays on, whereas a Domindar, regarding its existence and the range of its powers, is indeed a nonphysical product of the very thing that the Domindar 'plays on' (so to speak), *after* having been 'installed.' The relationship between Domindar and nervous system is much closer than the relationship between pianist and piano—so close that the nervous system, especially the brain, can well be called an *instrumental Domindar*.

How does this instrument function? At bottom, it functions not fundamentally unlike the way a piano functions. The pianist reads the musical score, and on the basis of this cognizance, she presses selectively the piano keys as she wishes and thinks fit; the internal mechanism of the piano translates the resulting patterns of key-pressings into the sequences of sounds intended by the pianist. *Analogously*: The subject of consciousness and of physical action (the Domindar) reads in its consciousness, and on the basis of this cognizance it contributes by informed choice to the resolution of an

how it could"—is not recommendable: The facts of mind-brain-interaction undeniably show that consciousness—something non-abstract and nonphysical—nomologically depends in many important ways on the physical. Why should it not also *come from it* (whether with the help of God or without), without being determined by it?

extra-neuronal macrophysical indeterministic situation detected by it (for example, 'Left? Right? Or straight on?'); the internal mechanism of the nervous system translates the resulting physically spontaneous (but Domindar-determined) brain process—the immediate effect of the subject's choosing—into the subject-intended (as such, *extra-neuronal*) restriction of precisely the macrophysical indeterministic situation in question.

9. Domindar vis-à-vis extra-neuronal macrophysical indeterministic situation

An extra-neuronal macrophysical indeterministic situation has—in its relation to a cerebrated natural living being 'which can do something about it'—a cerebral, a neuronal, and an extra-neuronal aspect. Corresponding to this, the contribution of the living being's subject of consciousness and of physical action to the resolution of that situation has, likewise, a cerebral, a neuronal, and an extra-neuronal aspect. The cerebral aspect—the physically causeless brain event *caused* by the nonphysical subject of consciousness and of physical action—is that subject's *originative* (and *properly own*) causal influence; the rest is a (sometimes misfiring) *causal projection* into the larger macrophysical environment by means of 'automatic' neuronal and muscular electro-chemistry and mechanics. However, *what is important* to the subject of consciousness and of physical action is precisely this projection, and mainly the extra-neuronal part of it (because that part is *what is intended* by the subject): the bodily movement (whether in flight, fight, or other life-relevant situation).

The mechanisms of the event-causal connection between the cerebral and the extra-neuronal aspect—via the neuronal aspect—are nowadays well understood. I need not emphasize how precarious, how endangered this connection is in every cerebrated natural living being, and in particular in every human being. I also need not emphasize that it seems mysterious *how* the nonphysical subject of consciousness and of physical action—the Domindar—manages to decide a macrophysical indeterministic situation at least to such an extent *partially* that a physically spontaneous brain event comes about, an event which is not an event of ontic chance

(i.e., without any sufficient cause at all) but caused by the Domindar, and *truly decisively* caused by the Domindar: without superfluity of the Domindar's causality.[6] *How* this is effected seems mysterious to everybody.[7] It should be noted, however, that there is little rationality in seeking to turn the mysterious *how* of the originative, seminal agency of the subject of consciousness and of physical action into, so to speak, a rope wherewith to strangle that subject philosophically. Generally speaking, *the mysteriousness* of the state of affairs that X makes E happen—with regard to *how* X makes E happen—neither entails the non-obtaining of that state of affairs, nor the nonexistence of X. There is, moreover, simply no doubt about the existence of subjects of consciousness and of physical action, or about the existence of their agency. *We* are subjects of consciousness and of physical action ourselves, and we act physically, with our brains, with and for our organisms. There is only a question about the initial, or *first*, agency of a subject of consciousness and of physical action: *How* does *that* agency come about?

Counter-question: Must every instance of causality have a specified way in which it *works*? The answer is: Not every instance of causality *can* have a specified way in which it works. The situation is as follows: If a causal nexus is *mediate*, then *how it works* can, in principle, to some extent be elucidated. Some understanding of *how it works* can be achieved by showing

[6] Does 'without superfluity of the Domindar's causality' mean that the brain event would not have happened if it had not been caused by the Domindar? Not necessarily; for even if it was caused without superfluity of the Domindar's causality (without unnecessary causal overdetermination coming from the Domindar's side), the brain event might perhaps have happened *anyway*, perhaps without any sufficient cause at all: 'by chance.' (The latter possibility is, however, excluded if the principle of sufficient cause is true.)

[7] Not to everybody but to some people, it seems mysterious *when* the Domindar caused the brain event. But this is not mysterious at all: the Domindar caused the brain event at the time it happened, not earlier, not later. In turn, not to everybody but to some people, it seems mysterious *what* the function of the Domindar's causing of the brain event is; but, again, this is not mysterious at all: the Domindar's causing of the brain event is the initiation of the implementation of a choice between macroscopic physical possibilities, a choice regarding *which* of these possibilities will be actual. [Details can be found in (Meixner 2014).]

that the nexus consists of several steps of cause and effect, where each cause and each effect is characterized in some detail; in other words, some functional understanding of a *mediate* causal nexus can be achieved by showing that the nexus is a causal chain, with such and such—specifically characterized—members. The finer the division and specification, the better is the explanatory result with respect to *how* the causal nexus works. However, due to human cognitive limitations, this procedure of discovery must inevitably come to an end. We must stop the procedure after a finite number of rounds (perhaps because we cannot see any further, perhaps because we simply have to go on with our lives), and thus the discovery and characterization of intermediate causes and effects terminates inevitably with causal nexuses where *how they work* is not—at least not yet—understood: the nexuses between items X_i and X_{i+1} in the causal chain that has so far—so far as one has come—been disclosed in scrutiny. If, however, a causal nexus is *immediate*—if it is without intermediate, mediating causes and effects—then a division of the causal nexus into causal steps is impossible; *here*, one cannot discover any (proper) causal chains,[8] and no description of *the given cause and effect*, be it ever so detailed, will produce an understanding of *how their causal connection works* (that is, an understanding in addition to the, so to speak, trivial understanding which is already provided by the very definition of their relationship as one of *cause and effect*[9]). *Here*, one is confronted from the start with a causal nexus which cannot be *functionally* understood, not even to some extent. Now, the causality in which a nonphysical subject of consciousness and of physical action causes brain events which have no sufficient physical cause—this causality is a *purely*

[8] Proper causal chains have N members, where $N \geq 3$: $X_1 \rightarrow X_2 \rightarrow X_3$; $X_1 \rightarrow X_2 \rightarrow X_3 \rightarrow X_4$; ... .

[9] Certain powers, or, alternatively, 'covering laws,' or, more esoterically, certain comparative similarities between possible worlds, may sometimes be sufficient for constituting a connection of cause and effect, but, by themselves, they do not provide any information about *how it works*: they may constitute the connection, but do not begin to make it *functionally transparent*. A thoroughly agent-causal conception of the causal connection can be found in (Meixner 2017); but of course this conception, too, arrives at its limits when the question is *how* the causal connection works.

immediate one. It is, as a purely immediate one, also a *purely first-causal* one.[10] This much, at least, can be perfectly understood about purely immediate causality. But there is nothing explanatory to be said about *how it works*, *how it functions*, *what its mechanism is*. It is an ontically effective choice. That's all.

References

Barrett, Jeffrey A. 2014. "Quantum Mechanics and Dualism." In *Quantum Physics Meets the Philosophy of Mind. New Essays on the Mind-Body Relation in Quantum-Theoretical Perspective*, edited by Antonella Corradini and Uwe Meixner, 65–81. Berlin: De Gruyter. https://doi.org/10.1515/9783110351064.65

Craig, William Lane, and James D. Sinclair. 2012. "The *Kalam* Cosmological Argument." In *The Blackwell Companion to Natural Theology*, edited by William Lane Craig and James Porter Moreland, 101–201. Malden, MA, and Oxford: Wiley-Blackwell. https://doi.org/10.1002/9781444308334.ch3

DeWitt, Bryce S.M., and R. Neill Graham, eds. 1973. *The Many-Worlds Interpretation of Quantum Mechanics*. Princeton: Princeton University Press.

Eccles, John C., and Karl R. Popper. 1977. *The Self and Its Brain*. Berlin: Springer. https://doi.org/10.1007/978-3-642-61891-8

[10] Purely immediate and purely first-causal causation are logically equivalent: Let '$\mathrm{C}ab$' stand for 'a causes b,' 'a is a sufficient cause of b.' *y is purely first-causally caused, for example by x*: $\mathrm{C}xy \wedge \forall z(\mathrm{C}zy \supset \neg\exists z'\mathrm{C}z'z)$; *$y$ is purely immediately caused, for example by x*: $\mathrm{C}xy \wedge \forall z(\mathrm{C}zy \supset \neg\exists z'(\mathrm{C}zz' \wedge \mathrm{C}z'y))$. Given these two definitions, it is true for all x and y: *if y is purely first-causally caused, for example by x, then y is purely immediately caused, for example by x.* [Assume: $\mathrm{C}xy \wedge \forall z(\mathrm{C}zy \supset \neg\exists z'\mathrm{C}z'z)$, and assume for *reductio*: $\neg(\mathrm{C}xy \wedge \forall z(\mathrm{C}zy \supset \neg\exists z'(\mathrm{C}zz' \wedge \mathrm{C}z'y)))$; hence $\exists z(\mathrm{C}zy \wedge \exists z'(\mathrm{C}zz' \wedge \mathrm{C}z'y))$; hence $\exists z'(\mathrm{C}z'y \wedge \exists z\mathrm{C}zz')$—contradicting the first assumption.] It is an obvious corollary that it is true for all y: *if y is purely first-causally caused* $[\exists x\mathrm{C}xy \wedge \forall z(\mathrm{C}zy \supset \neg\exists z'\mathrm{C}z'z)]$*, then y is purely immediately caused* $[\exists x\mathrm{C}xy \wedge \forall z(\mathrm{C}zy \supset \neg\exists z'(\mathrm{C}zz' \wedge \mathrm{C}z'y))]$. And also the converse of this is true for all y: *if y is purely immediately caused, then y is purely first-causally caused.* [Assume: $\exists x\mathrm{C}xy \wedge \forall z(\mathrm{C}zy \supset \neg\exists z'(\mathrm{C}zz' \wedge \mathrm{C}z'y))$, and assume for *reductio*: $\mathrm{C}zy \wedge \exists z'\mathrm{C}z'z$; hence $\mathrm{C}zy \wedge \exists z'(\mathrm{C}z'z \wedge \mathrm{C}zy)$, and from this, by the *transitivity of causation*: $\mathrm{C}zy \wedge \exists z'(\mathrm{C}z'z \wedge \mathrm{C}z'y)$, hence $\exists z'(\mathrm{C}z'y \wedge \mathrm{C}z'z \wedge \mathrm{C}zy)$, hence $\exists z'(\mathrm{C}z'y \wedge \exists z''(\mathrm{C}z'z'' \wedge \mathrm{C}z''y))$—contradicting the first assumption.]

Kim, Jaegwon. 1993. *Supervenience and Mind. Selected Philosophical Essays*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511625220

Meixner, Uwe. 2006. "Consciousness and Freedom." In *Analytic Philosophy without Naturalism*, edited by Antonella Corradini, Sergio Galvan, and E. Jonathan Lowe, 183–96. London: Routledge.

Meixner, Uwe. 2008. "New Perspectives for a Dualistic Conception of Mental Causation." *Journal of Consciousness Studies* 15 (1): 17–38.

Meixner, Uwe. 2014. "Of Quantum Physics and DOMINDARs." In *Quantum Physics Meets the Philosophy of Mind. New Essays on the Mind-Body Relation in Quantum-Theoretical Perspective*, edited by Antonella Corradini and Uwe Meixner, 17–34. Berlin: De Gruyter. https://doi.org/10.1515/9783110351064.17

Meixner, Uwe. 2017. "Agent-Causation—Neither Upward Nor Downward." In *Philosophical and Scientific Perspectives on Downward Causation*, edited by Michele Paolini Paoletti and Francesco Orilia, 278–95. New York and London: Routledge (Taylor & Francis).

Organon F 26 (1) 2019: 122–140
https://doi.org/10.31577/orgf.2019.26108

ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

RESEARCH ARTICLE

Active Doings and the Principle of the Causal Closure of the Physical World

Ansgar Beckermann*

Received: 13 July 2018 / Accepted: 4 December 2018

Abstract: Some philosophers hold that it would be impossible for us to do something actively if the physical world were causally closed, i.e., if in the physical world all events were caused by other physical events if they are caused at all. The reason for this view is that these philosophers adhere to what I call the traditional picture of action. Recently, Martine Nida-Rümelin tried to defend this picture by phenomenological considerations. According to the traditional picture a behavior can only count as something an agent does actively if it is ultimately caused by the agent in an agent-causal way. In this paper I adduce three arguments against agent causation: (1) We do not really understand what agent causation is. (2) If agent causation were real, we would be confronted with the strange fact that human agents can only cause certain tiny events in the brain. (3) There is no empirical evidence that agent causation is real. In the last part of my paper I present an alternative account of the difference between what agents do actively and what is done to them.

Keywords: Active doings; agent causation; causal closure of the physical world; free will; Martine Nida-Rümelin; subject causation.

* Bielefeld University
✎ Department of Philosophy, Bielefeld University, Postfach 10 01 31, D-33501 Bielefeld, Germany
✉ ansgar.beckermann@uni-bielefeld.de

© The Author. Journal compilation © The Editorial Board, *Organon F*.

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International Public License (CC BY-NC 4.0).

1. Kant's dilemma

According to Onora O'Neill, Kant held that we have to adopt two standpoints in leading our lives—standpoints which we cannot help to adopt despite their apparent incompatibility.[1] On the one hand, there is the theoretical standpoint which is naturalistic, "from it we see the world and human life as subject to natural law and causal inquiry" (O'Neill 1997, 273). On the other hand, there is the practical standpoint, the standpoint of human freedom, "from it we see ourselves as agents who intervene in limited ways in that natural order. Only the theoretical standpoint can accommodate science; only the practical standpoint can accommodate morality" (O'Neill 1997, 273). Moreover, in acting we are often guided by reasons—and reasons are, as many believe, totally different from causes which govern the natural world.

But although we cannot go without any of those two standpoints, it is unclear how they can be true at the same time.

> The predicament in which we find ourselves is not that of having to lead our lives in two distinct ontological orders, but that of having to adopt two mutually irreducible standpoints in leading our lives. [...] We are unavoidably, deeply, and thoroughly committed both to the naturalistic standpoint and to the standpoint of freedom. We can dispense with neither standpoint, since neither makes sense without the other. If we do not see ourselves as free we can give no account of activity, hence none of the activities of judging and understanding by which we establish the claims of knowledge; if we do not see ourselves as parts of a causally ordered world we can give no account of the effective implementation of human projects, including moral action, in the world. Our lives would be impossible without commitment to freedom *and* to causality in the robust sense in which Kant understands these terms: neither can stand alone. Yet we do not understand, let alone know, what makes them compatible. (O'Neill 1997, 272f.)

[1] Cf. (Beckermann 2016) concerning this paragraph and the following one.

I cannot judge to what extent this assessment does justice to Kant's philosophy. However, it seems to me that it captures quite aptly the basic controversy of the German debate on the problem of free will that has taken place in the last decades [cf., for example, (Geyer 2004)]. On the one hand, many scientists insist that the natural sciences tell us that the physical world is a causal net of physical events, which, if they have any causes at all, are caused by other physical events. In this net, therefore, there does not seem to be any room for the intervention of personal agents. On the other hand, many philosophers underscore the indispensability of the practical standpoint. For, so the argument goes, first, daily experience confirms this standpoint a hundred times a day and, second, we cannot but view ourselves and our fellow humans as acting persons; even natural scientists have to understand themselves as acting persons, or else they would undermine the very foundations of their work. And, of course, in particular scientists need to act on the assumption that what they do is guided by reasons. Science is simply defined as the search for those hypotheses that are justified by the best reasons [see (Janich 2008), (Heidelberger 2005)]. To put it in a nutshell, a central claim of the theoretical standpoint is that the physical world is causally closed, that in the physical world, all events, including all movements of the limbs of animals, are caused by other physical events if they are caused at all. A central claim of the practical standpoint in contrast is that some animals and we humans are at least sometimes able to do something actively, we are able to intervene in the course of physical events.

Many scientists and philosophers agree on that the two claims are mutually exclusive. If the theoretical standpoint is true, then the practical one cannot obtain, and if the practical one obtains, then it is impossible at least that the natural sciences yield a complete picture of the world. To some, this incompatibility is so obvious that they wonder how one can even try to overcome it. This, in turn, is due to the fact that they construe especially the practical standpoint in such a way that it is indeed precluded that this standpoint obtains if the theoretical one is true. I call this interpretation "the traditional picture of active doings."

When someone does something actively, this usually means that she intervenes in the course of the physical world. Had she done nothing, this

world would have developed in a certain way; by her action, however, the course of the physical world changed. Therefore, the idea of active doings seems to imply that, on the one hand, there is the physical world, which, left alone, develops in a certain way. On the other hand, there is the agent, who is somehow located *beside* or *outside* this world. She can observe the course of the physical world in a way that does not change anything,[2] but she can also, within the limits of her possibilities, intervene by action and change the direction of the course of the physical world. It is crucial to see that, according to the traditional picture, doing something actively means to intervene in the course of the physical world *from outside*. As I see it, in the history of occidental philosophy this basic idea has been spelled out in two different ways—in *Cartesian dualism* and in the idea of *agent causation*. Sometimes, these two perspectives are even merged. Both, interactive Cartesian dualism and the idea of agent causation are incompatible with the claim that the physical world is causally closed. If some movements of my limbs are caused by mental events in the sense of Cartesian dualism then some physical events have nonphysical causes. And if some movements of my limbs are caused by me in the sense of agent causation it is certainly not true that these movements have only physical events as causes.

2. Martine Nida-Rümelin on active doings

Ten years ago, Martine Nida-Rümelin offered a view of active doings which combines Cartesian dualism with the idea of agent causation. She argues that to say that an animal or a human being is doing something actively is to say that something immaterial caused bodily movements of the animal or the human being in an agent-causal way. Martine Nida-Rümelin develops her considerations against a phenomenological background—based on an analysis of what we experience when we are active or what we experience when we observe, for example, animals that are actively doing something, instead of suffering something that merely happens to them. Think of a squirrel which runs around and gathers nuts, which climbs

[2] Provided the assumption that observation alone does not already change the world.

a tree and hops from branch to branch. Most people certainly hold that the squirrel has experiences and that it perceives the world in a certain way. According to Nida-Rümelin, however, most people do not only believe this, "[they] *perceive it as such*" (Nida-Rümelin 2007, 257; emphasis in the original). But, following Nida-Rümelin, to view animals as subjects of experiences implies much more:

> But the experience of an animal as a subject of experience normally not only involves being aware of the fact that there is 'someone' who has experiences, it also normally involves awareness of something we might call spontaneity. Seeing, for example, a squirrel as a subject of experience involves, in addition to seeing it as having experiences, seeing it as being active. Seeing a squirrel as a subject of experience involves seeing a great part of its bodily moves as genuine activities. A squirrel's jump from one branch to another does not look to someone who sees the squirrel as a subject of experience like the mere result of some inner mechanical process. It looks like something done by the squirrel, by the subject at issue itself. (Nida-Rümelin 2007, 257)

This similarly applies to the experiences we have when we ourselves do something actively:

> An analogous claim applies to the way we perceive ourselves in our own doings. In doing something we are at least normally phenomenally aware of doing something. To be phenomenally aware of doing something involves the experience of oneself as being active. In doing something we experience our own spontaneity. We are aware of the difference between those cases where something merely happens to us and those cases where we directly bring about what happens. (Nida-Rümelin 2007, 258)

Nida-Rümelin's phenomenological thesis therefore is:

> [I]t is normally part of the content of our experience when we do something that we are active in that doing; and it is part of the content of our perceptual experience when we observe others in their doings that they are active in what they do. (Nida-Rümelin 2007, 258)

However, this is not decisive. What is crucial is the transition from a phenomenological to the ontological thesis:

> It seems quite obvious to me that to experience oneself as active in one's doing can be described equally well by saying that we experience ourselves as *the cause* or *a cause* of what happens. The same applies to the perception of others: to see another animal as being active in its doing is to see it (the other experiencing subject) as a causal origin of its moves. But what is required for the experience to be veridical? Do we need to assume that these experiences can be veridical only if the subject is itself a cause in a sense similar to the one discussed under the heading 'agent causation'? It seems quite clear to me that we have to admit this further step. (Nida-Rümelin 2007, 262; emphasis in the original)

Thus, "to be active," according to Nida-Rümelin, means "to cause something." A being is active only in case it itself causes the bodily movements that constitute its behavior—in the sense of subject causation. Nida-Rümelin prefers the term "subject causation" to the term "agent causation," because she is concerned not only with actions, but with all activities of which it can be said that humans or other living beings do them on their own (so her concern is for all "active doings").

> In this paper I will be primarily concerned with doings or activities rather than with the more specific case of actions. Many things we do in our daily lives are not actions. A person caught in her thoughts might smile, scratch her head, stand up and walk around. Under normal circumstances these activities are not actions, they are only doings. Doings are often done without reason. The author of a doing need not notice what she is doing. Most bodily movements of human and non-human animals while awake are doings. Doings need not be controlled: the spontaneous laughter after hearing a joke, the crying of a baby that longs for the presence of a parent, or the happy smile of surprise when meeting a friend unexpectedly are examples of doings without control. Nonetheless the person is active in her laughing, crying, or smiling. (Nida-Rümelin 2007, 245f.)

Thus, Martine Nida-Rümelin is concerned not only with actions but with everything that can be considered an active doing. Let us return to the issue of subject or agent causation.[3] Nida-Rümelin asks whether it makes sense to assume that the subject or the agent causing a behavior is a material being.

> Can the claim of subject causation be combined with the idea that the subject of experience is a material thing (the whole body, the brain or a part of the brain)? [...] Suppose that subjects cause events in their brain *and* that subjects are identical to their brain. It follows that the brain as a whole causes events happing within the brain. (Nida-Rümelin 2007, 264; emphasis in the original)

Is this a reasonable idea? What is it supposed to mean that a physical system causes a process *within* that system? Does it, for example, make sense to say that a computer causes processes that run within it? "It seems plain that to talk in this way is to talk nonsense. All causation happening in the case of a computer is event causation" (Nida-Rümelin 2007, 264). According to Nida-Rümelin, our brains are not different in this respect. Even if we concede that brains are subjects of experiences, the following applies here as well:

> A philosopher who accepts that material objects as a whole can cause events happening inside of them should have a way to render this supposed causal relation conceivable which is independent of the assumption that the object belongs to the special class of experiencing subjects and which is applicable to nonconscious material things as well. If this reasoning is correct then we are confined to a dualist subject causation theory. (Nida-Rümelin 2007, 264)

Nida-Rümelin does not seek to set out a traditional dualism, but that does not matter here. What is important is that she holds that to do something actively means that the subject causes the corresponding physical movements in the sense of agent causation and that the causing subject itself cannot be anything physical, neither a body nor a physical part of a body.

[3] In the following I will make no difference between subject and agent causation.

3. Objections to agent causation

I think that Nida-Rümelin is surely right in assuming that there is a difference between what some beings do actively and what only happens to them, what they suffer passively. However, I object to her equating *being active* with *the agent being the cause of bodily movements, in the sense of agent causation* without much ado. First, I think, she is wrong about the phenomenology. At least, when I observe a squirrel jumping from branch to branch I do not at all have the phenomenological impression that the squirrel's movements are caused by the squirrel itself or that they are not "the mere result of some inner mechanical process." But what is more important is that Nida-Rümelin so easily pushes aside the many problems that come with the idea of agent causation. In the following I shall raise essentially three objections against agent causation.

A common objection to agent causation has been that it cannot explain the exact time at which the caused event takes place. To me, however, a more general objection, my first objection, seems even more severe. The very notion of agent causation is hardly intelligible. According to the generally acknowledged view, causality involves two relata—cause and effect. In the case of event causation both relata are *events*. The stone's hitting the glass pane (cause *C*) brings about the breaking of the pane (effect *E*). With agent causation, however, things are different. There is only one event—the effect *E*. The cause is not an event, but an agent *A*. Agent causation consists, so we are told, in the agent *A* bringing about the event *E*—straightforwardly, without doing anything else that is the real cause *E*. Of course, I can scare away a fly by moving my hand; but this is just a case of event causation—the movement of my hand causes the fly's disappearance. But the movement of my hand, the exponents of agent causation reply, is not caused by another event, it is caused solely by me. I do not do anything else which causes my hand to move. One central problem that arises from this assumption can be illustrated as follows. In the case of event causation one must answer the question: what distinguishes the case that event *E* merely followed event *C* from the case that *C* indeed *caused E*. This is the question of what constitutes the difference between *post hoc* and *propter hoc*. The answer to this question is contentious. Many would,

however, agree with the following two claims. (a) If *C* causes *E*, there must be a nomological connection between *C* and *E* based on appropriate natural laws. And (b) if *C* causes *E*, *E* would not have occurred if *C* had not taken place.[4]

What now of agent causation? Here the question is what distinguishes the case that *E* occurs *simply in the presence of A* from the case that *A causally brings about E*. To my knowledge, there is not even a remotely satisfactory answer to this question. Often it is said that in the latter case *A* just caused *E*; however, this is, of course, not an answer, but merely a repetition of the thesis. Therefore, it is all but surprising that traditionally movements caused by an agent are traced back to *acts of will*; in my view, this is an attempt to model agent causation in analogy to event causation. Still, the question remains what causes those acts of will.

One can also give these questions an epistemic twist. In the case of event causation, the question then is how can I *find out* whether event *E* merely follows *C* or whether *E* was caused by *C*. In the case of agent causation, the question is how can I find out whether *E* merely occurred in the presence of *A* or whether *E* was caused by *A*. Again, there is, to my knowledge, no satisfying answer to the second question. The only possibility seems to be to ask *A*. But this is not possible with beings incapable of language. Moreover: is information given by agents truly reliable? There is, however, an answer to the first question—by means of experiments. In many experiments researchers try to hold one factor constant—e.g., *C*—and to systematically vary as many other factors as possible. If in all of these variations *E* always follows *C*, then this indicates that *C* caused *E*. This approach is based on the fact that we can actively intervene in the course of the world; because this allows us to isolate the factors with which we can reliably bring *E* to the fore. *Bring to the fore* is a decisive term here. If we can reliably make sure that *E* occurs, by making sure that *C* occurs, this indicates that *C* causes *E*. With respect to agent causation, there are then again only open questions. To the best of my knowledge, there is no approach with which we can test whether an agent *A* has caused an event *E*.

[4] Both points are already present in (Hume 1993, Section vii, part ii).

A second objection is not so fundamental, but should give proponents of the idea of agent causation something to think about. Already Descartes was aware of the fact that the soul cannot directly cause changes in the periphery of our body and that, reversely, things in our environment or states of affairs of our feet and hands cannot directly cause changes in the soul. As soon as the afferent nerves leading from the hand to the brain are interrupted, we no longer feel any more pain from our injured hand; and as soon as the efferent nerves leading from the brain to the hand are interrupted, we can no longer move our hand. The spot of interaction must hence lie in the brain, which for Descartes meant in the *pineal gland*. Even proponents of agent causation concede that, in general, agents—human beings or animals—can only cause changes in the brain. This leads first and foremost to the question why we do *not* experience precisely this. When I lift my arm, I may have the experience of causing the lifting of the arm, but certainly not the experience of causing a certain process in my brain. But putting that aside: If it is true that agents can only cause changes in their brains, why is that so? In principle, it should not be more difficult to directly cause certain muscle contractions than to make certain motor neurons fire. And why can't agents directly cause changes in the world, in the way telekinesis is supposed to work? Why do I have to hit the keyboard with my fingers? If there is agent causation, it should be no problem to operate the keys by causing their movements directly. The fact that I cannot do all these things is in my view evidence for the conclusion that something is basically wrong with the idea of agent causation.

A third objection is for those who are not yet convinced by the arguments presented so far. Even if the idea of agent causation were not faced with the problems just mentioned, there is no *empirical* evidence for the occurrence of such causation. First of all, I already mentioned that agent causation is incompatible with the thesis of the causal closure of the physical world, i.e., the thesis that all events have physical events as causes, if they have any cause at all. Therefore, if agent causation really does occur in this world, two conditions have to be met: 1. There are physical events which have no (sufficient) physical causes. 2. It can be made plausible that these events are nevertheless caused, namely, by certain agents. Is that feasible? Perhaps. Let us again consider the phenomenon of telekinesis. Suppose

a person is asked to make a bottle, standing five meters away from her on a table, fall over by "mere power of thinking." Let us suppose further that the person succeeds in eight of ten attempts and that with later repetitions the success rate is approximately the same. This certainly would be a convincing result. I assume that the scientists who conducted the experiment have checked the experimental arrangement carefully. There are no hints with respect to manipulation or technical tricks of any kind. In other words: There is no evidence of physical causes for the bottle's tilting over. Apparently, it is the subject alone who succeeds by mere power of thinking or willing to make the bottle fall over. This is possible; but, actually *there is no empirical evidence for such cases*. Here and there we find scattered reports of the kind in question, but telekinetic phenomena could so far not be proved in systematic examinations.

In this context, ever since Descartes, interactionistic dualists have applied a remarkable trick. They assume that the physical effects that can be brought about by the soul or by agents are very small and therefore hardly observable. Descartes, for example, held that the soul cannot do more than cause certain movements (twists) of the pineal gland.[5] But even if this were the case: The interaction of the soul with the physical world would have to be demonstrated by first ascertaining that certain movements of the pineal gland have no sufficient physical causes and secondly, by making plausible that those movements nevertheless have a cause—the soul.

Nowadays, obviously there are only few proponents of agent causation (if any) who explicate in detail which brain events agents can cause. The firing of motor neurons, however, as well as the firing of other neurons is essentially dependent on which and how many neurotransmitters are being released from the synaptic terminals of precedent neurons. Hence, those

[5] John Ecces holds that the mind can only interact with a certain part of the cortex (the liaison brain). The mind scans small functional units (modules) in that area which are accessible to it due to their "openness" and "slightly modifies them, whereupon the modules react collectively to these slight alterations and forward this common reaction via the associational and commissural fibres" (Eccles and Zeier 1980, 173; my translation). In his more recent work, Eccles conjectures the location of the mind/matter-interaction in dendrones, units even smaller than the modules.

authors have to assert something like the following: In some cases, the release of neurotransmitters (1) has no sufficient physical cause, but (2) it has, as can be made plausible, nevertheless a cause—the agent herself. Or, to take another example: In some cases, the generation of an action potential in a motor neuron (1) has no sufficient physical cause, but (2) it has, as can be made plausible, nevertheless a cause—the agent herself. Undoubtedly, it is far from easy to examine such assertions empirically. However, to my knowledge we don't yet have the slightest evidence speaking in favor of the truth of those claims.

4. A naturalistic alternative

I agree that there is indeed a difference between active and passive. With regard to the movements of a human being, we can justifiably claim that some of these movements are due to the person herself—she lifts her arm; she sings a song; she scratches her head. On the other hand, there also are movements which are not doings of the person—somebody takes my arm and pulls it up; someone pokes me to the ground. In both cases, these movements of my arm or my body *cannot* be ascribed to myself. In my view, however, it inevitably leads to a dead end if one tries to spell out the difference between active and passive by means of the idea of agent causation. But what could an alternative account look like?

To begin with, at least one fundamental difference between actively moving and passively being-moved is that the latter is often brought about by *external forces*.[6] A blast of wind blows me over, Fido the dog is pulled away from his favorite tree—in both cases, external forces affect me and Fido, respectively, causally bringing about the corresponding movements. It is something altogether different if Fido gets up from his basket and runs to the door. Of course, this too can have (amongst others) external causes, e.g., the noise being caused by the unlocking of the door. But this cause works in a completely different way than the blast of wind or the dog's master pulling the leash. Fido's running consists in the movement of his

[6] Cf. (Beckermann 2008, § 2.3) and (Beckermann 2011) with respect to this paragraph and the following ones.

legs, but the noise at the door doesn't exert any forces on those legs. In fact, the movement of Fido's legs is based solely on the contraction and relaxation of certain muscles. External forces do not play any role here. That animals move by themselves thus means at least in part that the *energy* which is necessary for the execution of such movements stems *from within them*.

However, this aspect alone does not suffice. Consider the patellar reflex.

> Striking the patellar ligament with a reflex hammer just below the patella stretches the muscle spindle in the quadriceps muscle. This produces a signal which travels back to the spinal cord and synapses (without interneurons) at the level of L3 in the spinal cord, completely independent of higher centres. From there, an alpha motor neuron conducts an efferent impulse back to the quadriceps femoris muscle, triggering contraction. This [...] causes the leg to kick. (<https://en.wikipedia.org/wiki/Patellar_reflex>)

In this case, too, no external forces affect the leg; the energy needed for its movement stems completely from the patient himself. Still, in this case we don't say that the person herself has lifted her leg. In general, the following applies: Internally induced movements of a person's limbs rest on muscle contractions and relaxations. Those contractions and relaxations are themselves brought about by the firing of motor neurons whose cell bodies are located in the anterior horn of the spinal cord and whose axons reach down to neuromuscular endplates directly neighboring muscle cells. The lower motor neurons can in turn be activated by upper motor neurons, the cell bodies of which lie in the motor cortex of the brain and axons of which reach to the cell bodies of the lower motor neurons located in the anterior horn.

With regard to the kicking of the leg due to the patellar reflex, no higher regions of the CNS are involved. This is important. For in the neurobiological literature one can often read that it is the *upper* motor neurons which are responsible for intentional movements. In other words: With regard to intentional movements, the neuronal impulse leading to the corresponding contractions and relaxations must come from the motor cortex. This suggests the

following way to characterize active doings: A human being moves one of his limbs actively if the corresponding muscle contractions trace back to neuronal impulses from his motor cortex. But that is still not the whole story.

With regard to human beings, there is an option we don't have concerning animals—we can ask them whether they did something by themselves, i.e., whether they ascribe a certain movement of their limbs to themselves or to outward causes. This was taken advantage of by Roger Penfield, who during surgeries on open brains in the middle of the 20th century induced complete movements of limbs by stimulating the premotor cortex and supplementary motor area. However, as Gerhard Roth reports in his book *Fühlen, Denken, Handeln*, the patients experienced these movements as imposed; they did not say: "I did that."[7,8] Thus the fact that the neuronal impulse that led to a movement of a limb originated from the motor cortex of a person does not suffice to his ascribing the move to himself. Rather, it seems that this is only the case if the impulse came about in a specific way. In the first edition of the book mentioned, Roth presented the following hypothesis:

> The impression of *our movements being instigated by ourselves in an act of will* rests on a completely different reason. This impression is a sign for our brain that before the movement started the dorsal and ventral cortico-limbic loop has been passed through and that the executive centers of the cerebral cortex together with the limbic system have 'sufficiently dealt' with it. In this case the symmetric and then the lateralized readiness potential build up, and the latter makes the 'starting shot' for the execution of the

[7] German original: "Ich habe das getan".

[8] Cf. (Roth 2003, 515). On the other hand, Roth refers to José Delgado who "reported that under conditions similar to those of Penfield, the stimulation of the rostral part of the so-called internal capsule (i.e., the fibers that run from the thalamus through the basal ganglia to the cortex) led to movements of the patient which he ascribed to himself" (Roth 2003, 516; my translation). In (2009), Winfried Löffler conducted a critical and to my mind convincing analysis of Roth's description of Delgado's findings. Löffler's conclusion: There are so far *no* empiricial findings that suggest it is possible to externally stimulate real actions (actions that agents ascribe to themselves "in the truest phenomenological sense").

> intended movement. The impression of *fiat*!, of *I want this now* is therefore the conscious notification of this neurophysiological process. (Roth 2001, 446; my translation)[9]

In the second edition of his book, Roth did not repeat this hypothesis. Perhaps, however, the exact details are not decisive. Maybe it is enough to notice that people seem to ascribe just those movements to themselves which rest on neuronal impulses from the motor cortex that in turn have been neuronally produced in a certain way.

At any rate, these findings fit in very well with the following idea: Humans (and some animals) are in a certain way automata—entities that move by themselves; but they are not automata that always act mechanically, knee-jerk or thoughtlessly, i.e., automatically. Rather they are *autonomous systems*. Firstly, this means that they have a repertoire of diverse behavioral patterns which allows them to act differently even in situations of the same kind. And, second, it means that they have the ability to choose between those different behaviors according to the situation. This ability contains two sub-capabilities—the ability to analyze the situation at hand appropriately (Which objects are where in relation to me? Are those objects dangerous or useful? etc.) and second, the ability to find a course of action which in the given situation serves the attainment of one's own goals. A dog is being attacked; it needs to choose between fight and flight. This decision rests, if I see this correctly, on a neuronal decision-making process. The dog itself (or maybe even the dog's self) plays no role in the causation of the dog's behavior.

In other words: Humans (and many animals) have a decision-making or control system that allows them to find their way in very diverse situations

[9] German original: "Das Gefühl der *Selbstveranlassung unserer Bewegungen im Willensakt* haben wir aus einem ganz anderen Grund. Dieses Gefühl ist für das Gehirn ein Zeichen, dass vor dem Starten der Bewegung die dorsale und ventrale cortical-limbische Schleife durchlaufen wurde und die exekutiven Zentren der Großhirnrinde zusammen mit dem limbischen System sich damit 'ausreichend befasst' haben. In diesem Falle baut sich das symmetrische und dann das lateralisierte Bereitschaftspotential auf, und letzteres gibt den 'Startschuss' für die Ausführung der intendierten Bewegung. Das Gefühl des *fiat*!, des *ich will das jetzt* ist demnach die bewusste Meldung dieses neurophysiologischen Vorgangs."

and to choose the course of action that seems most beneficial. This control system is, from what we know, realized in the CNS. Therefore, my hypothesis is that the neuronal subsystems of the CNS, which play a central role in Roth's considerations, are precisely those parts of the brain in which the control system is realized that makes us autonomous systems. That is to say, humans and animals carry out those movements themselves that are based on neuronal impulses from those parts of the motor cortex that are in turn under the control of their central neuronal control system.

However, one restriction must be added: perhaps it is possible to induce movements by stimulating certain cerebral regions, which the respective persons then ascribe to themselves.[10] In this case, one would have to say that these persons were wrong. In fact, these movements were none of their doings. Therefore, one should finally say: humans and animals carry out those movements themselves that are based on neuronal impulses from the motor cortex, which in turn are being controlled by their central neuronal control system, as long as this control system is not manipulated from outside, i.e., as long as the humans and animals are not being "remote-controlled." This point is very important, too. *Remote-controlled* model airplanes or toy robots resemble animals in many ways. But they don't act on their own as long as they're being controlled from outside. This would be different only if by implementation of *internal* control-systems the model airplanes or robots became autonomous systems themselves.

5. Summary

I think, just like Martine Nida-Rümelin, that there is a real distinction between active and passive, between what humans and other living beings do actively and what merely happens to them. I take it to be a mistake, however, to spell out this difference in terms of agent causation. It is in general a mistake to construe agents as entities *beside* or *outside* the physical world. Humans and animals are as much part of the physical world as stones and rivers. They differ from stones and rivers only in that they are

[10] Once again: Delgado's findings do not indicate this. But that doesn't mean that it isn't possible.

much more complex and contain many physical mechanisms which make them autonomous systems—but not in that they contain additional non-physical components. When they intervene in the course of the physical world, they do so from within, as part of this world.

But is Nida-Rümelin not right in claiming that we often have the *phenomenal impression* that we as agents cause some of our movements? Here I would like to make three comments. First, phenomenal impressions may be wrong. I approach the door of a supermarket and the door opens. I may have the impression that my approaching the door was the cause of the opening of the door. But I may be wrong; the opening of the door may have quite another cause. Second, is it really true that I often have the phenomenal impression that I as an agent *cause* some of my movements? I doubt that. I agree that I often have the phenomenal impression that one of my movements is something I do actively, not something that happens to me. But in my view this impression is not directly about causation. Remember how Nida-Rümelin initially expresses herself:

> To be phenomenally aware of doing something involves the experience of oneself as being active. In doing something we experience our own spontaneity. (Nida-Rümelin 2007, 258)

Initially she does not say that in doing something actively we experience ourselves as *causing* our behavior. Third, scientific progress sometimes forces us to adjust our phenomenal experiences or at least to interpret them in a new way. I have the phenomenal impression that the table in front of me is a solid object, i.e., that the space it covers is *entirely* filled with a hardly penetrable material. But science tells us that the table is made up of billions of atoms at a large distance from each other and that there is plenty of empty space between the atoms. Nonetheless, our phenomenal impression that the table is solid can be saved, since other solid objects like a cup that I put on the table will not fall through it. The reason is that the forces holding together the atoms of the table and the cup respectively are so big that the atoms cannot easily be separated from each other. In much the same way the problems with which the idea of agent causation is confronted and the fact that scientists have found no evidence whatsoever that agent causation is real should motivate us to interpret *doing something actively* in a way that does not depend on the idea of agent causation.

I have tried to sketch such an alternative account. Besides the fact that this account of active doings avoids the problems of the idea of agent causation, it has another advantage. It demonstrates that doing something actively is compatible with the principle of the causal closure of the physical world. If a behavior is an active doing in fact it is caused by a very specific kind of neuronal processes then even active doings may have only physical events as causes. Thereby, the analysis also opens up a solution of the Kantian dilemma. We can very well understand ourselves *simultaneously* as part of the physical world which is causally closed *and* as active beings.[11]

References

Beckermann, Ansgar. 2008. *Gehirn, Ich, Freiheit. Neurowissenschaften und Menschenbild*. Paderborn: mentis.

Beckermann, Ansgar. 2011. "Ich sehe den blauen Himmel, ich hebe meinen Arm." In *Geist und Moral*, edited by Christoph Lumer and Uwe Meyer, 19–34. Paderborn: mentis.

Beckermann, Ansgar. 2016. "Willensfreiheit. Die Agenda des Naturalisten." In *Wollen. Seine Bedeutung, seine Grenzen*, edited by Neil Roughley and Julius Schälike, 341–67. Paderborn: mentis.

Eccles, John C., and Hans Zeier. 1980. *Gehirn und Geist*. München: Kindler.

Geyer, Christian, ed. 2004. *Hirnforschung und Willensfreiheit*. Frankfurt am Main: Suhrkamp.

Janich, Peter. 2008. "Naturwissenschaft vom Menschen versus Philosophie." In *Deutsches Jahrbuch Philosophie Band I: Naturalismus und Menschenbild*, edited by Peter Janich, 30–51. Hamburg: Meiner. https://doi.org/10.1524/dzph.2007.55.6.893

Heidelberger, Michael. 2005. "Freiheit und Wissenschaft! Metaphysische Zumutungen von Verächtern der Willensfreiheit." In *Neurowissenschaften und Menschenbild*, edited by Eve-Marie Engels and Elisabeth Hildt, 195–219. Paderborn: mentis.

Hume, David. 1993. *Untersuchung über den menschlichen Verstand*. Übersetzt von R. Richter, mit einer Einleitung herausgegeben von J. Kulenkampff. Hamburg: Meiner.

[11] Translated from a German version with the help of Alin Cucu.

Löffler, Winfried. 2009. "Neuromythologie, oder: Wie aus empirischen Mücken narrative Elefanten werden." In *Gehirne und Personen,* edited by Martina Fürst, Wolfgang Gombocz, and Christian Hiebaum, 113-26. Frankfurt am Main: Ontos.

Nida-Rümelin, Martine. 2007. "Doings and Subject Causation." *Erkenntnis* 67 (2): 255–72. https://doi.org/10.1007/s10670-007-9064-z

O'Neill, Onora. 1997. "Kant on Reason and Religion." *Tanner Lectures on Human Values* 18: 267–308.

Penfield, Wilder. 1958. *The Excitable Cortex in Conscious Man.* Springfield, Ill.: Liverpool University Press.

Roth, Gerhard. 2001. *Fühlen, Denken, Handeln.* 1nd ed. Frankfurt am Main: Suhrkamp.

Roth, Gerhard. 2003. *Fühlen, Denken, Handeln.* Neue, vollständig überarbeitete Ausgabe, 2nd ed. Frankfurt am Main: Suhrkamp.

Organon F 26 (1) 2019: 141–168
https://doi.org/10.31577/orgf.2019.26109

ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

RESEARCH ARTICLE

Freedom, Power and Causation

Thomas Pink*

Received: 1 July 2018 / Accepted: 29 November 2018

Abstract: Freedom or control of how we act is often and very naturally understood as a kind of power—a power to determine for ourselves how we act. Is freedom conceived as such a power possible, and what kind of power must it be? The paper argues that power takes many forms, of which ordinary causation is only one; and that if freedom is indeed a kind of power, it cannot be ordinary causation. Scepticism about the reality of freedom as a power can take two forms. One, found in Hume, now often referred to as the *Mind* argument, assumes incompatibilism, and concludes from incompatibilism that freedom cannot exist, as indistinguishable from chance. But another scepticism, found in Hobbes, does not assume incompatibilism, but assumes rather that the only possible form of power in nature is ordinary causation, concluding that freedom cannot for this reason exist as a form of power. This scepticism is more profound—it is in fact presupposed by Hume's scepticism—and far more interesting, just because freedom cannot plausibly be modelled as ordinary causation.

Keywords: Causation; chance; compatibilism; freedom; Hobbes; Hume; incompatibilism; law of nature; power; reason; scepticism; Suarez.

* King's College London
Department of Philosophy, King's College London, Strand, London WC2R 2LS, United Kingdom
tom.pink@kcl.ac.uk

© The Author. Journal compilation © The Editorial Board, *Organon F*.

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International Public License (CC BY-NC 4.0).

1. Power

We think that, within limits, we have control over our actions—that it is up to us what actions we perform. A very natural conception of this control is as a kind of power. Our control of our actions is a power to determine for ourselves what we do—a power of self-determination in the form of freedom. But what then is the nature of freedom so conceived—as a power to determine our own actions for ourselves?

Many philosophers think or write as if power were a rather uniform phenomenon. It is often proposed that power is by its very nature a causal phenomenon:

> In the first place, the notions of power or disposition are already causally laden notions and it can thus reasonably be argued that unless one already has a grasp of causation, one cannot have a grasp of power. Powers, indeed, are often called causal powers.[1]

But is all power causal by its very nature? And if control or freedom is a kind of power, must it in particular be a form of causal power?

Of course, the claim that all power is causal could be so understood as to be trivial. 'Causation' could be used as no more than a general label to apply to whatever power turns out to be. But in the passage just cited the claim that all power is causal is presented as a substantial thesis—as something that is not trivial, but to be 'reasonably argued.' In which case the idea might be to inform the theory of power by importing a definite and specific conception of causation. And this has certainly been a project of much metaphysics since Thomas Hobbes. One very intuitive case of power is the very familiar kind that appears to be involved in obvious cases of causation, and to be possessed and exercised, not by causes and effects indifferently, but specifically by causes. This is the power of stones to break windows or the power of fire to melt ice—the power that ordinary causes have to produce their effects. Is all power, then, power in this specific form? The project would then be to understand all forms of power in terms of this particular form of power—the

[1] (Harre and Madden 1975)—as cited and endorsed in (Mumford and Anjum 2011, 7).

power involved in ordinary causation. All genuine power is like the power of stones to break windows and of fire to melt ice.

I shall understand the claim that all power is causal as a claim that is substantial in just this way. The claim is that all power is like the power to produce outcomes exercised by ordinary causes—such as by stones to break windows and fire to melt ice.

At this point it is useful to step back, and raise the question why causation itself is so widely viewed as involving power—and, more specifically, a power possessed and exercised by causes over what they affect?

Power involves a kind of capacity. Causal power constitutes, after all, a capacity to produce effects. But, of course, it is not the mere presence in them of a capacity that makes it true that causes possess power. And that is because the idea of a capacity extends far wider than that of power. For example, there are capacities not to cause and affect, but to be affected. But the capacity to be affected hardly constitutes any kind of power over anything, and the process of being affected is hardly the exercise of power. The contrary is true: to be affected is to be subject to power that is possessed and exercised by something else. Contrast my view with Locke's. In *An Essay Concerning Human Understanding*, Locke divides power into active and passive (Locke 1975, 234). Active is defined as the power to make a change, passive is the power to receive it. As an account of power this is certainly defective. For, of course, Locke ignores powers to prevent change from occurring. But more importantly, Locke's 'passive power' involves the opposite of any exercise of genuine power. It is a form of powerlessness—subjection to the power of another.

Power, then, is a very special capacity. And what, I conjecture, is common to power in all its forms is a capacity to produce or, at the upper limit, to outright determine the occurrence or non-occurrence of outcomes. It is this capacity to determine what happens that causes possess, but which their effects lack. Causes determine the occurrence of their effects, and not vice versa. Furthermore, linked to the general notion of determination, where power is concerned, is an equally general notion of responsibility. Whatever exercises a power to determine outcomes is in some corresponding way responsible for what that exercise of power determines. Causes possess power in so far as they can influence or determine an outcome. And in so

determining that outcome they are responsible, causally responsible, for its happening. Causes are responsible for the occurrence of their effects, and not the other way round, just as anything that possesses power is correspondingly responsible for what that power's exercise determines.

Freedom is a capacity to determine one very important kind of outcome—that involved in the performance of action. So understood, as a capacity to determine, freedom is definitely a case of power. But how far, in discussing freedom, are we concerned with power in specifically causal form—the kind of power that is involved in stones breaking windows and in fire melting ice?

2. Powers causal and non-causal

If a power is a capacity to determine, a causal power must be a capacity to determine causally. And that immediately suggests the possibility, at least at the conceptual level, of power that is not causal. A power that is not causal is going to be a power the exercise of which determines outcomes, but without determining them causally. And so understood, we certainly entertain ideas of other kinds of power besides the causal. We do deploy an understanding of capacities that determine outcomes without determining them causally.

There are, for example, various kinds of moral power, such as powers to impose or release from moral obligations. Consider promising, for example. Promisors have the power to impose an obligation on themselves—an obligation, owed to a promisee, to act as promised. And then promisees can release the promisors from the moral obligation of their promise by declaring them released. The promise determines or produces an obligation; and then the promisee's declaration determines the removal of the obligation and the promisor's release. But in neither case is the obligation or release from it produced as an effect, by virtue of some causal law. Rather the promise constitutes the imposition of the obligation and the promisee's declaration constitutes the promisor's release from it; and each does so by virtue of something very different from causal laws, namely the moral principles governing promising.

As with moral powers, so too there are legal powers. A creditor has the power to release his debtor from a debt—a power exercised through making

some legally valid declaration of release. By making the declaration, such as by declaring 'I release you', the creditor determines that the debtor is released, and so is responsible for the occurrence of that release. But again the power here is not causal. Uttering the declaration does not cause the debtor to be released. Rather, thanks not to some causal law but to rules governing credit, the utterance constitutes that very event of release. The declaration determines the release, but does not determine it causally. But uttering the declaration is no less an exercise of a power or capacity to determine outcomes for that.

Then there are normative powers attaching not to agents, but to things that cannot be causes at all, to which we readily attribute capacities to move and to determine nonetheless. Suppose you entertain in thought a mathematical truth. That truth might determine or strongly incline you to assent. But what about the truth moves you to assent? The everyday answer in such cases is clear: its evident nature—in other words you are moved to believe by the clear justification there is for believing the truth. Perhaps indeed you are not only led to believe the truth, but that there is justification for believing it. Either way, what leads you to form the mathematical belief, and possibly also the belief in the justification, is the truth and the justification for believing it. Now in this case what is described as moving you is not the sort of feature involved in ordinary causation. You are contemplating, not an entity with a location in time and space, but an object of thought—a mathematical truth. And if the object of your thought is true, its truth is plausibly necessary, and certainly not something that functions as an ordinary cause. For ordinary causes and effects are contingent. And what moves you to believe it is the evident nature of the truth, the clear justifiability of believing it. And this is a normative property attaching to the truth—something that, no matter if it moves or even determines your assent, is again not the sort of feature involved in ordinary causation.

As with mathematical truths, so too with desirable options by way of actions and outcomes. The evident goodness or desirability of a possible action might move you to form a desire for its performance or believe its performance desirable, and eventually even to decide on and intend its performance. What led you to want to perform that action or to believe performing it desirable and then to decide to do it? We have just given the

answer: the evident desirability of doing it. And such an answer reveals you as rational animal, responsive to and moved by the good—just as you might also be moved, as rational animal, by the true. The action that is so desirable may never be performed, of course. What moves you is not some contingent event in the world, but, again, an object of your thought—such as the truth that a given kind of action is desirable. You may be moved by the justifiability of desiring and intending the action. But the desirability or justifiability of a given kind of action is again something normative, and not the sort of property involved in ordinary causation.

We readily use a vocabulary of power, influence and determination to pick out normativity as well as causation. We talk of being moved by the *force* of an argument. And we use this vocabulary of force just because we think of normatively forceful arguments as really possessing the capacity to influence or move or even to determine our assent—and to move or determine it through the justifications they present. Our capacity for reason or rationality is a capacity to be moved by argument—and by the normative force of an argument in particular. So reason or rationality involves responsiveness to a kind of power—the power of good argument and genuine justification.

It is tempting to dismiss this talk of normative power as not literal, just because the power envisaged is not causal. At best we have here, it might be alleged, a manner of speaking or a metaphor. But there is an obvious difficulty with this move. Power follows from a genuine capacity to determine or move or influence. And surely we think it true that the very quality of an argument can be what moves people to assent to it. But for that to be true the quality of an argument must actually have a capacity to move. Good arguments must have genuine force—a power that is non-causal because located in the normativity attaching to objects of thought, but which is a genuine capacity to move nonetheless.

Again, consider the normative powers, not of objects of thought, but of agents themselves. Our talk of promisors and promisees as having the capacity to determine moral obligations—as imposing those obligations on themselves or releasing others from them—does not appear to be metaphorical. That is exactly how promising is understood: as an act by which moral obligations to promisees may be imposed, and from which obligations promisees may in turn generally provide release should they so choose. Whatever

moral principles underlie these powers, we very clearly do understand them as genuine powers—genuine capacities, possessed by promisors and promisees, to determine outcomes. But the powers are not causal. For the powers are underpinned by moral principles and not by causal laws.

The idea of power in non-causal form is very controversial—a controversy that, as we shall see, was raised in especially sharp form in the seventeenth century by the work of Hobbes. But two different issues are involved that we must take great care to distinguish. One issue concerns our very concept of power, while the other concerns the metaphysics of power.

Some philosophers would insist that whatever we might ordinarily think or say, all genuine power is causal, at least regarding its fundamental constitution. No outcomes are determined without being determined causally. Now this claim is about metaphysics. It is about the true constitution of power, and the reality of its operation. And this metaphysical claim might turn out to be correct. But this is not the sort of claim proposed by Mumford and Anjum above, which is about our very concept of power. That claim is that our very concept of power is causal. To think of a power operating to determine an outcome just is to think of it as determining the outcome causally. But this conceptual claim is not obviously true; and it should not be assumed to be true just because of the metaphysical conviction that only power in causal form is real. We may have a concept of power that allows for power to take non-causal form, even if it indeed turns out that all the cases of power that do really exist are causal.

Take our psychology as it involves rationality or what we ordinarily take to be our receptiveness to justifications. The metaphysician who believes that all genuine power and determination is causal must claim that if my belief in a mathematical truth or in its justifiability really is produced or determined by anything, it cannot be determined by the normative properties of an object of thought. If anything really is determining me to believe in the truth or in the justifiability of assenting to it, this must be some genuine cause, and the determination must be causal. Perhaps, for example, the immediate cause of my belief in the mathematical truth is a prior psychological event—such as the event of entertaining that truth. It is this psychological entertaining that is the immediate, and causal, determinant of my belief. To describe me as rational is just to describe my beliefs as

susceptible to a particular kind of causal force—the causal force of those entertainings that present good justifications for believing. Similarly with desires: what determines me to desire an outcome is not its desirability, but the psychological event of entertaining that outcome. Agents are rational if their desires are susceptible to the causal power of those entertainings of outcomes that are in fact desirable.

Now it might in the end prove true that causation is the only real case of power in nature. But if so, that is still not what we ordinarily assume. Take the prior entertaining of the object of thought. It is certainly not this event alone that we ordinarily think of as determining belief or desire. For we ordinarily take a rational agent's belief to be determined not simply by the fact that they have entertained a claim, but by the evidence for or clear justification for the truth of what is entertained. Similarly a rational agent's desire is motivated not by the mere fact that they have entertained a given option, but by the desirability of the option entertained. Rational or reasonable agents respond to justifications not because the event of entertaining them just happens to produce that effect, but because the justifications are good ones. The quality of these justifications really is what moves rational agents to respond as justified. That is rationality as ordinarily understood—not susceptibility to causal forces merely, but susceptibility to the force of justification.

Again, consider the moral powers we ordinarily ascribe to promissors and promisees. These are powers to produce moral obligations and remove them. And the principles that base these powers, as we have observed, are not causal laws but moral principles—the moral principles governing promising. These principles are, in particular, distinct from any psychological laws that might cover what we actually, whether rightly or wrongly, do. For moral principles are concerned not with what we actually do, but with what we are under an obligation to do. Now it might be that these moral powers and the obligations they are exercised to produce and remove do not really exist. It may be that moral principles transcending and normatively corrective of actual human practices are entirely imaginary; and so are the moral powers those principles supposedly constitute. But again this claim is highly controversial. If true, it will be made true by metaphysical reality, not by the fact that our concept of power is just a concept of causation. For our entertainment of various

moral and normative powers—powers reflecting moral and normative principles, and not causal laws—seems to suggest that our concept of power is considerably broader than any concept of causation.

We need then to make an important distinction. One question is whether we think of some capacities to determine outcomes as determining those outcomes non-causally. And it seems that we do. It is then a further and importantly different question whether these non-causal powers really do exist and operate—as we so clearly assume them to. That is, we should distinguish analytic claims about our concept of power from metaphysical claims about what powers actually obtain. For it seems plain that whether or not moral and normative powers actually exist, we do at least suppose them to exist, and to operate non-causally.

What seems to distinguish cases of power that, real or not, are at least understood by us to be non-causal? There seems a range of possible differences from ordinary causation, from the breaking of windows by stones and the melting of ice by fire, not all of which need be exemplified together. There may be many different forms of non-causal determination.

In some cases, two things combine together to differentiate the power from ordinary causation. First there is the nature of the determining or moving entity, which is not an entity contingently located and operative in space and time as a stone must be, but an object of thought. And then secondly there is the mode of determination or influence, which is through properties that are normative—that are to do with the justifiability of responding to the object in certain ways, such as by forming a desire or a belief directed at it. It is these normative properties that move us so to respond. It is the evident truth or desirability of what we are thinking of that moves us to believe or want it to be true—or so we suppose. And not everyone is so moved, of course. Our rationality reflects our susceptibility to the power of justifications. We will be moved by justifications, but only to the extent that we are indeed susceptible to their normative power—only to the extent that we are indeed rational.

In other cases the outcome is certainly determined, just as in ordinary causation, by a specific entity located in time and space, such as by an agent doing something. What in these cases establishes the non-causal nature of the power is the mode of determination involved taken together

with the nature of the outcomes immediately determined. In the case of moral and legal powers, determination of the outcome seems to involve the application not of some causal law, but of moral or legal principles to determine specifically moral and legal outcomes, such as obligations. Thanks to these principles, the utterance of certain words constitutes the incurring or removal of an obligation.

What of freedom? As a power freedom belongs not to truths or objects of thought, but to agents—and so to potential bearers also of power in causal form. Moreover, as with causation the outcomes immediately determined need not be legal or moral, and are not determined according to specifically legal or moral principles. Which is no doubt why it is so especially tempting to assimilate freedom to some form of causation. But freedom may yet prove not to be a form of causal power. Though as agents we may also be bearers of causal power, and though the outcomes we control may be outcomes that could also be produced by us through mere causation, the way we determine those outcomes through exercising freedom may prove to be very unlike the way causes such as stones or events involving these would determine them. Free agents may determine outcomes, but quite differently from the way causes determine outcomes.

English-language philosophy has tended to suppose that the causal nature of power in general is somehow a conceptual truth, so that it is not only unproblematic but actually mandatory to assimilate the capacities to determine outcomes that we postulate in ordinary thinking to various cases of causal power. Freedom, especially, has been treated by philosophers in just this way. Even the sceptic has tended to assume that though our belief in freedom may be belief in a power that does not exist, it is still a belief in some non-existent form of causal power. But this approach to understanding the concept of freedom may be misconceived. Whether freedom exists or not, our conception of it cannot be assumed to be of some kind of power in causal form.

3. Thomas Hobbes on power

Thomas Hobbes pursued each of the two philosophical projects that tie power exclusively to causation—the metaphysical project of claiming that

causation is the only power there is, and the analytic project of identifying the very concept of a power or capacity to determine outcomes with the concept of causation. We shall shortly be examining Hobbes's disbelief in the existence of freedom as a power involving alternatives. But Hobbes's attack on freedom was part of a wider scepticism. He denied the very intelligibility of any kind of power or determination beyond ordinary efficient causation. For example, Hobbes's scholastic Aristotelian opponent Bramhall was happy to talk of the desirability of the goal or object of an action as moving or determining the action's performance, but without doing so as an efficient cause. The source of the motivation involved an object of thought; and its mode of determination was characterized by Bramhall not as natural, as in efficient causation, but as normative or moral:

> Secondly, for the manner how the understanding doth determine the will, it is not naturally but morally. The will is moved by the understanding, not as by an efficient, having a causal influence into the effect, but only by proposing and representing the object. (John Bramhall in Hobbes and Bramhall 1656, 55–56)

We can now see what Bramhall has in mind. The understanding moves us by presenting us with a claim or with an option. But to the extent that we are rational, what finally determines our belief or will is not the understanding, or some occurrence within it, operating merely as an efficient cause. What determines our belief is not simply the psychological event of entertaining the claim or option. What determines us to believe a claim or to decide on an option has to do with the object that the understanding presents and, specifically, its normativity—such as its evident truth or desirability.

The acknowledgement of powers or capacities to determine that are normative, and that do not simply involve ordinary causation, was a central feature of scholastic ethical theory—and a feature, in particular, of its theory of action and motivation. Motivation is naturally conceived by us often to involve our subjection to a form of power: something *moves* us to act as we do in pursuit of goals. Modern philosophers, taking it to be a conceptual truth that power is inherently causal, assume that any motivating power must be understood by us in causal terms. Those contemporary philosophers, therefore, who oppose Hobbes and Davidson, and who assume that

motives are not causes, tend to write as if motivation, as ordinarily understood by us, had nothing to do with power at all.[2]

But Hobbes's scholastic opponents were quite different. They took motivation to involve a variety of kinds of power. One sort might be efficient causal, as where the motivation of voluntary actions by prior decisions and intentions so to act was concerned. Intending to do something would indeed move us to do it by causing us to do it. But motivation was provided not simply by attitudes operating as ordinary causes, but also by the objects at which those attitudes were directed—a motivation that explained those attitudes themselves as well as the further voluntary actions that those attitudes caused. We could be moved to want something, as well as to decide on it and to pursue it as our goal, by its clear goodness and desirability. And here some form of determination was again involved—but a determination that was moral rather than efficient causal.

There is a familiar tension in the common-sense psychology of action between two kinds of power. There is the power involved in motivation—a power of motives to get us to perform actions. This is a power to which we as agents are subject, and by which we may be influenced or even determined. We are being moved to act by something else—a motive. And then there is a power that we ourselves exercise—the power of self-determination, our power to determine actions for ourselves. The long-standing debate about freedom or free will between compatibilists and incompatibilists is a general debate about the compatibility of freedom or self-determination with determination of the agent by prior factors, and therefore concerns the tension between these two forms of power in particular. How far is an agent's capacity to determine for themselves what they do compatible with the determination of what the agent does by prior motives? Are some forms of motivating power incompatible with our power to determine for ourselves what we do?

[2] Thus in his *Teleological Realism* Scott Sehon claims that the explanation of action by motivating psychological attitudes is teleological not causal. But, in his account of common-sense psychology, the only power involved is causal—as exercised not by our attitudes themselves, but by the physical states that underlie our attitudes. Our attitudes themselves and their objects appear not strictly to *move* us after all.

Just because scholastic action theory allowed room for more than one kind of motivating power, so for this reason there existed within the scholastic tradition more than one problem about the compatibility of freedom with prior determination, and with motivation in particular. There was of course a problem specifically about causation and causal power—about the compatibility of freedom with the determination of the agent by causes outside his control. But there was also a parallel problem to do with the compatibility of freedom with normative power. If evidences or proofs are sufficiently powerful, can they not outright determine the assent of any rational agent in a way that removes any freedom to believe otherwise? And similarly can there not be outcomes or objects so completely good as, once entertained, to determine the agent's choice in a way that removes the agent's freedom to decide otherwise? God, or the good in infinite and unqualified form was conceived within the Thomist tradition as just such an object. That even now we readily describe proofs or evidence as *compelling* or *overwhelming* shows that we still allow for normative power in a form that can reduce or threaten freedom.

Hobbes, by clear contrast, caustically rejected all such appeals to powers other than ordinary causation:

> Moved not by an efficient, is nonsense. (Thomas Hobbes in Hobbes and Bramhall 1656, 59)

In Hobbes's view, any determination of anything, including any action, must be by an efficient cause. Hobbes therefore turns Bramhall's motivating object of thought into a prior psychological occurrence. Rather than being moved into a decision by a normative property—the desirability of an option—we are moved causally by a psychological event, such as a prior passion for or desire for that option, an occurrence located in the world as is any efficient cause, and of the same metaphysical kind as the action it motivates and causes.

4. Freedom as power over alternatives

What I shall call the *causal theory of freedom* says that, whether freedom actually obtains or not, our concept of freedom is not only of a power,

but of a causal power. It is a conceptual truth, the causal theory says, that to exercise freedom is to exercise power causally, so that any outcome determined through the exercise of freedom is determined causally. What we determine to happen through the exercise of our control occurs as an effect that we cause:

> The exercise of active control is essentially a causal phenomenon. (Clarke 2003, 151)

The causal theory is widely believed. But even if the Hobbesean metaphysical view, that all real power is causal, is true, the causal theory of freedom—an analytic theory of our concept of freedom—may still be false. It may still misrepresent our understanding of what freedom is.

We certainly understand causation as extending freedom. Given control of how I act, I can control what causally depends on my actions. If flicking the switch would cause the lights to go on or off, controlling whether I flick the switch will give me control over whether the lights go on or off. But the fact that freedom is causally extendible does not show freedom itself, as ordinarily conceived, to be a causal power. For other powers besides causation can extend our freedom too. Indeed, any powers attaching to my actions may further extend my freedom, what I have control over, provided I control those actions. If my actions have the power legally to determine a given outcome, such as your release from debt, then my control of how I act can give me further control over that outcome too; I gain control over whether you remain in debt to me. So the power of freedom is legally extendible. But freedom is not shown by this to be itself a legal power. No more does freedom's causal extendibility show it to be a causal power. All that has been established thus far is that freedom is a power that can be extended by a variety of other powers, whether causal or non-causal. What kind of power freedom itself amounts to remains quite open.

Freedom is ordinarily understood by us to be a power to determine alternatives—a power of control over which actions we perform. Our conception of our power of self-determination is as up-to-usness—a conception of self-determination that immediately characterises it as a power over more than option. Freedom is a power that leaves it up to us whether we do A or refrain; it is a power of *control* over which actions we perform. Central to the idea of freedom, then, is power over alternatives. This involvement of

alternatives is picked out by the 'up to me whether' construction, which is completed by specification of alternatives by way of actions and outcomes within my power; freedom is the power to determine for ourselves which alternative occurs.

This constitution as power over alternatives seems to distinguish freedom from other forms of power, and from ordinary causal power in particular. The power under given circumstances to produce more than one outcome seems essential to the character of freedom. It is not obvious that there is anything left of our ordinary understanding of up-to-usness if we subtract this capacity to produce more than one outcome. That is just what the power is: control of how I act. So to have the power, at least in its complete form, there must be more than one outcome that I can determine. How can my action be within my control if I lack the power to refrain as well as to do? Our conception of causal power, on the other hand, is quite different. We have an understanding of causation as commonly a power to produce but one outcome. Heavy bricks hurled at fragile windows may have a causal power then to do but one thing—to break the window. We have no tendency to understand causal power as being always and by nature a power to produce alternatives.

There are other ways in which freedom differs from ordinary causation. One way is especially obvious. Any exercise of the power of freedom has to occur through agency—and specifically through agency that is intentional or deliberate. If I am to exercise my power to determine for myself what happens, then I must do so either through deliberately and intentionally doing something or through intentionally refraining. If it is to be up to me whether the lights are on or off, there must be some action available to me—up to me to perform or not—such as flicking a switch, by which I can affect whether or not the lights are on or off. And actually to be exercising my power I must either be intentionally performing the action—intentionally flicking the switch—or be intentionally refraining from its performance. But ordinary causation carries no such tie to agency. I can produce many effects other than through doing or refraining. I can crush something just through my very weight, independently of any action I may perform or omit performing.

Furthermore, this power over alternatives by way of action seems to matter to moral responsibility as ordinary causation does not. I may, just

through my weight and size and other features of me, produce many effects. But that I have produced effects does not come close to establishing any moral responsibility on my part for their occurrence, unless I had some control over what led to these effects—or so it is very natural to suppose. Whereas that it was up to me whether or not something occurred seems immediately relevant to the question of my moral responsibility for the occurrence.

There is much to be said about freedom's peculiar tie to agency and its distinctive relevance to our moral responsibility. But I wish to concentrate here on freedom's essential character as involving power over alternatives.

Thomas Hobbes saw in freedom's involvement of a power over alternatives a central and very problematic difference between freedom and ordinary causation. The way in which freedom is supposed to involve alternatives violated, in Hobbes's view, central truths about causation. Since in his view causation was the only power in nature, Hobbes concluded that there could not be such a power as freedom. Hobbes was not even a compatibilist about freedom as a power. He denied its very existence outright. Freedom consisted not in a power over alternatives, but in something quite different: namely, in an absence of obstacles to the satisfaction of an ordinary one-way causal power—the power of a motivation to cause its satisfaction. Freedom consists, for example, in the absence of external constraints, such as chains, that might prevent my desires from causing movements by me that might satisfy them:

> Liberty is the absence of all impediments to action, that are not contained in the nature, and in the intrinsecal quality of the agent. (Thomas Hobbes in Hobbes and Bramhall 1656, 285)

Indeed Hobbes not only denied the existence of freedom as a power. He denied its very intelligibility. He claimed that we lacked even the concept of a power to determine things for ourselves. Talk of such a power was mere philosophers' jargon. He mounted his assault on the very intelligibility of self-determination as part of a radical programme to detach ethical and political theory from reliance on the notion. How did Hobbes propose to detach ethics from self-determination? Some of the time Hobbes did what Hume would do later as well—which is to treat moral blame as no more than negative evaluation:

> [Why do we blame people?] I answer because they please us not. I might ask him, whether blaming be any thing else but saying the thing blamed is ill or imperfect [...] I answer, they are to be blamed though their wills be not in their power. Is not good good and evill evill though they be not in our power? And shall I not call them so? And is that not praise and blame? But it seems that the Bishop takes blame not for the dispraise of a thing, but for a praetext and colour of malice and revenge against him that he blameth. (Thomas Hobbes in Hobbes and Bramhall 1656, 40)

In other contexts Hobbes seems to allow for a distinctive responsibility for how we act:

> The nature of sin consisteth in this, that the action done proceed from our will and be against the law. (Thomas Hobbes in Hobbes and Bramhall 1656, 185)

But the responsibility here involves a kind of legal responsibility—according to a view of that responsibility which avoids appeal to self-determination. Holding someone responsible, in Hobbes's view, seems to involve no more than holding them to sanction-backed directives on the voluntary—something that presupposes no more than their rational responsiveness to such directives. To be morally responsible, on this model, we have merely to be legally governable. But, for Hobbes, that only requires that we be capable of performing or avoiding actions on the basis of a desire so to do, as a means to avoiding sanctions. And this presupposes nothing more than what Hobbes termed *voluntariness*—action occurring as an effect, through ordinary causation, of prior desires or appetites, such as desires to avoid sanctions. And this was something that Hobbes thought had nothing to do with self-determination. We were not determining for ourselves what we did. Rather our actions were being determined by our desires and appetites, and not by us.

Hobbes's opponent Bramhall was effectively a spokesman for the ethical and psychological theory of the late scholastic Francisco Suarez. And it is Suarez who is the ultimate target of much of Hobbes's writing in this area. In Suarez the idea of freedom really is the idea of a special kind of power—a power that, though still for Suarez a form of causation, is causation of a quite distinctive kind. Freedom is causal power in what he describes as *contingent* form [see (Suarez 1994, disputation 19)]. As a free agent I am not

a necessary cause as causes in wider nature are—a cause that under any given circumstances can operate in only one way. A massive brick that strikes a window can determine but one outcome—that the window breaks. Whereas, by contrast, I have a power, freedom, by which in one and the same set of circumstances I could equally well determine any one of a range of alternative outcomes. So under a given set of circumstances I have the power, say, to lower my hand or to raise it—and my nature as possessor of the power leaves it contingent how I will exercise it, and so which action I shall perform.

Hobbes denied that such a contingent power is possible, because it is unrecognizable as causal power. For Hobbes's scepticism about freedom is based on a clear view of the only form that power can take in nature. The only possible form that power, the capacity to produce or determine outcomes, can take, in Hobbes's view, is as ordinary causation—the kind of power that bricks, or motions involving them, possess and exercise to break windows. We shall see that Hobbes is right on one point at least. Whether or not the power of freedom is real, our conception of it radically distinguishes freedom and its operation from ordinary causation. In particular, freedom involves modes of determination not to be found in ordinary causation. In exercising freedom we exercise a power to determine that does not determine causally.

It is tempting to think that Hobbes's problem with freedom is mainly with what I have called *multi-wayness*. Freedom or control of what we do involves alternatives. To have control of whether one does A is to be capable of determining either that one does A or that one refrains. And it is very natural to view this control as a single power that could under given circumstances be employed in more than one way—hence multi-wayness—to produce either the outcome that I do A or the outcome that I refrain. That is the nature of control as a power: to leave it up to me which I do, and to be employable in doing either. Hobbes's case, on this reading of him, is simply that there cannot be such a thing as a multi-way power—a power that can, under a given set of circumstances, be used in more than one way, to produce one of a variety of outcomes.

However, we should beware of this tempting assumption. It should not be assumed that freedom, understood as its being up to me to determine a range of alternatives, need involve multi-wayness as just defined—a single

power employable in more than one way, to produce any one of these alternatives. Indeed, I shall suggest, even if freedom did not involve multi-wayness, it would still involve a form of power which Hobbes denied.

Moreover, it seems there could be cases of multi-way power that are not at all like freedom, but much more like (possibly slightly unusual) cases of ordinary causal power. True, much ordinary causation seems not to be multi-way—as the case of the brick hitting the window reminds us. Causation here seems to take one-way form. In a given set of circumstances, when the massive brick hits the window, the brick or its motion can exercise its power to produce but one effect—that the window breaks. But need this be true universally? Can there not be probabilistic causes with a power that could, under certain circumstances, operate in more than one way, to produce a range of outcomes? Perhaps the power of one particle to accelerate another could produce in the other particle, with some probability, one acceleration; or perhaps, with another probability, another slightly different acceleration instead. This would still be recognizable as ordinary causal power. And it would not involve the causing particle's possession of freedom. It would not be up to the particle which acceleration it produced; that would not be something that the particle 'determined for itself.'

Hobbes was, of course, a determinist. Probabilistic causation is not a possibility on his metaphysics of causation. He thought that a cause's power operates, under any given circumstances, to produce but one outcome. But the issue of multi-wayness—the possibility of a causal power's operating under given circumstances in more than one way, to produce more than one possible outcome—is not what was fundamental to Hobbes's scepticism about the very reality of freedom, or indeed of self-determination in any form at all. Hobbes's scepticism has more to do with something that can be detached from multi-wayness, and that radically distinguishes freedom from ordinary causation. I shall call this factor *contingency of determination*; and it has to do with how the possessor of a power, such as a cause, *determines* an outcome when it does.

In Hobbes's view, if an entity has the power to determine a specific outcome, and the conditions required for the successful exercise of the power are all met—then the power must be exercised. The determining entity's

very presence, with its power, must necessitate the occurrence of the outcome it has the power then to determine. It follows on this view that an entity cannot really possess the power to determine, under one and the same set of circumstances, more than one alternative outcome. For an entity really to be capable of determining each outcome, Hobbes argues, it must simultaneously produce each outcome. Referring, abusively, to Suarez's contingent cause as an 'indetermination,' Hobbes writes:

> But that the indetermination can make it happen or not happen is absurd; for indetermination maketh it equally to happen or not to happen; and therefore both; which is a contradiction. Therefore indetermination doth nothing, and whatsoever causes do, is necessary. (Thomas Hobbes in Hobbes and Bramhall 1656, 184)

Suarez was right about one thing. Contingent determination is part of our ordinary understanding of freedom, and distinguishes freedom from ordinary causation. In the case of freedom, the power-bearer may have the power to determine the occurrence of a particular outcome, and all the conditions required for the power's successful exercise may be met—without the power being exercised to produce that outcome. Freedom can involve the power to determine alternatives, only one of which can actually be produced, only because this is so.

Suppose by contrast an ordinary cause has under given circumstances the power to produce a range of possible effects. The cause is probabilistic: any one of these effects might with some probability occur, or it might not. In such a case the cause does not count as determining the effect that it produces. A probabilistic cause at most influences the occurrence of that effect, but without determining it in a way that removes all dependence of the final outcome on simple chance. Whereas we do think of the free agent as determining that he does what he does, but without the action's performance being guaranteed just by his presence as a free agent with the power then to determine it.[3]

[3] I made this distinction between freedom and ordinary causation, and discussed the problem it poses for a view of freedom as a straightforwardly agent-*causal* power in (Pink 2004, 114–15).

Contingency of determination distinguishes a free agent from any cause—including a probabilistic cause. But so too does something else—something which involves not the power's relation to outcomes, but the agent's or power bearer's relation to the power.

Consider again ordinary causes. Either their operation is predetermined by the very nature of the power and the circumstances of its exercise: in those circumstances their power is to determine one particular outcome, an outcome which they will then produce. Or, as in the case of probabilistic causes, how the cause will operate is undetermined, that is, dependent on mere chance. But what seems importantly to distinguish freedom, as ordinarily conceived, is that this is not so. It is neither predetermined nor merely chance and undetermined which way a free agent exercises their power. The agent determines for himself how he exercises his power. And it seems impossible to characterize this relation that the agent has to the power without using the concept of freedom. If the agent can determine for himself how the power is exercised, it must be up to the agent whether he exercises his power to produce this outcome or that. If the power of freedom is indeed multi-way, a power employable in more than one way to produce more than one outcome, then in relation to that power there is what we might term a *freedom of specification*: it is up to the bearer which outcome the power is exercised to produce.

Hobbes was very well aware of this element to our conception of freedom as a power. The idea of the agent's determining his exercise of the power is arguably central to self-determination—to the very idea of determining outcomes *for oneself*. In Hobbes's view, this idea of a determination of how the power is exercised is viciously regressive.

> And if a man determine himself, the question will still remain what determined him to determine himself in that manner. (Thomas Hobbes in Hobbes and Bramhall 1656, 26)

So the very idea of self-determination, for Hobbes, is incoherent. And that is because it viciously involves the idea of an agent's power to determine, the exercise of which that same agent has first to determine.

But it is not obvious that Hobbes is right about the regress. The regress is vicious only if the way in which the exercise of the power is determined—to produce this outcome or that—involves a prior exercise of power distinct

from the exercise of the power determined. But this is not obviously what we ordinarily suppose.

There is in the case of freedom a *conceptual* distinction between (a) the power's relation to outcomes—the power can operate to produce more than one outcome—and (b) the power's relation to me, namely that I determine for myself what way it operates. But we do not suppose there to be any corresponding *ontological* distinction between two distinct exercises of power—an exercise of power to produce outcomes, and then another and distinct exercise of power to determine the operation of that power to produce outcomes. Multi-wayness and determination of the mode of exercise by me are simply conceptually distinct features of a single exercise of control. In exercising control over outcomes I *ipso facto* determine for myself how the control is exercised. That is what control is—a power to produce outcomes the manner of exercise of which I determine for myself. In one and the same exercise of power I produce one outcome rather than another, and I determine how the power is exercised.

This freedom of specification does not involve then any exercise of power over and above that involved in the production of the outcome. But though there need be no vicious regress, we are clearly dealing with a kind of power that is not ordinary causation. In relation to this radically different kind of power the notion of freedom not only conveys a power over alternatives in relation to outcomes, but also the agent's distinctive relation to the power as its bearer.

Freedom, it now appears, brings alternatives into self-determination in two ways, one relating the power to outcomes, the other relating the power to its bearer, the free agent. *Freedom in relation to outcomes* relates self-determination as a power to the outcomes it determines. The power is a power to determine more than one outcome. *Freedom in relation to the power* has to do with the relation of the power to the free agent. This relation again involves alternatives, but this time concerning how the power is exercised or whether it is exercised at all. It might be that there are alternative ways in which the power might be exercised: it is up to the agent how he exercises the power, to produce this outcome or that. This is what we have already termed a freedom of specification. Or it might simply be that it is up to the agent whether he exercises the power at all. This we might term a *freedom of exercise*.

We can separate contingency of determination and multi-wayness. First, we might have multi-wayness but without contingency of determination. Probabilistic causation seems to involve such a possibility, as we have just seen. A probabilistic cause might possess, as we have noted, a power to affect acceleration that could under given circumstances operate in more than one way, to produce acceleration at more than one rate. But this causal power involves no contingency of determination. Given that it is initially chancy how the power will operate, the effect is influenced by the cause but not determined by it.

There might also be, as at least a conceptual possibility (this would not be any kind of power we actually accord ourselves) contingency of determination without multi-wayness. That is, under any given set of circumstances the power can be exercised in only one way—to determine but one outcome. But though the power is outcome-determining, its exercise to produce that outcome is not ensured just by the presence, under the relevant circumstances, of the power's bearer. The agent could possess the power then to determine that outcome, and all the conditions required for that power's successful exercise could be met—and the agent just not exercise it. There could be a power involving contingency of determination that was not multi-way.

Here there would be no possibility of a freedom of specification. It would not be up to the agent how he exercised the power, to produce this outcome or that, as under any given circumstances there would only ever be one way the power could be exercised. But it could still be up to the agent whether he exercised the power at all. In which case we would have something recognisable as a power of self-determination, but involving freedom only in relation to the power itself, as a freedom of exercise, and not freedom in relation to outcomes too.

5. Two scepticisms about freedom

Modern philosophical discussion of free will centres on a debate about causation between incompatibilists and compatibilists. This is a debate about the relation of freedom and causation—and specifically about the implications of causal determinism for the freedom to do otherwise. Is freedom as a power to do otherwise compatible with our being causally

determined to do what we actually do? Much modern scepticism about the very possibility of freedom is then based on the supposed conceptual truth of incompatibilism and centres on what I shall call the randomness problem. This is the worry, famously put by David Hume, that if incompatibilism were true—if freedom did require causal indeterminism—that really would leave us, not with genuine freedom, but with mere chance.

> […] liberty, by removing necessity, removes also causes, and is the very same thing with chance. (Hume 1978, 407)

And that threatened indistinguishability of freedom from chance drives the incompatibilist sceptic into concluding that freedom, as anything more than randomness or mere chance, must be impossible.

But now we see that there is another scepticism about freedom, and one that also involves causation and the freedom to do otherwise, though in quite a different way. This form of scepticism objects to the very idea of freedom as a power over alternatives, on the grounds that causation is the only possible form of power—and that such a power over alternatives would be too radically unlike causation. This second form of scepticism is even more threatening to everyday belief. Incompatibilism is not universally believed, even by ordinary people—witness the intractable nature of the debate about whether incompatibilism is indeed true, a debate that has long interested a public extending well beyond professional philosophy. Whereas our freedom's identity as a capacity to determine more than one outcome seems far more basic. It seems far more central to our ordinary understanding of what freedom is like in itself.

We have begun to examine how freedom as a power over alternatives might differ from ordinary causation. And it has emerged that freedom seems to differ from ordinary causation in a number of ways. First, there is multi-wayness—a single power that might under given circumstances operate in more than one way, to produce more than one outcome. Now it is true, as we have discussed, that at least in some cases ordinary causal power could take multi-way form. What distinguishes freedom from causation, is that multi-wayness seems to be characteristic of the very kind of power that freedom is—control of how we act. Then, and as a presupposition of any power to determine alternatives, freedom involves contingency of determination—a radically different way of determining outcomes from that

involved in causation. And then with multi-wayness comes, as equally essential to control, a freedom of specification. It is not mere chance and undetermined by anything how the power will operate. As control the operation of the power is determined by its possessor—the free agent.

The ideas of freedom that Hobbes attacks are not obviously incompatibilist in themselves. To say that a power involves contingency of determination, is not itself to say anything about the power's compatibility with causal determinism. All that contingency of determination expressly asserts, is that an agent might possess the power to determine an outcome in the circumstances—and yet still not exercise the power to produce that outcome. It is quite another question whether, compatibly with his possession of the power, the agent's exercising or failing to exercise it could itself be causally determined. And if contingency of determination is compatible with causal determinism, so too is multi-wayness. If it can be causally determined that I do not exercise a power then to determine one outcome, a power that I nevertheless possess, but instead exercise a power to determine another outcome, the power involved in relation to each outcome could perfectly well be one and the same. Is the power to produce one outcome distinct from the power to produce another? This question about the individuation of powers seems to have to do with their basis or constitution, not with their compatibility with causal determinism. And again the idea of a freedom of specification with respect to how control as a single multi-way power operates seems to add nothing to the case for an incompatibilist conception of freedom.

Hobbes's scepticism about freedom as a power over alternatives is the expression of a kind of philosophical naturalism. This is the naturalism that refuses to allow that human nature and its capacities involve powers and capacities that are *sui generis*—that are qualitatively different from powers and capacities found in wider nature. And freedom is being attacked by Hobbes precisely as such a *sui generis* power. His is an especially penetrating attack, and a reminder that even prior to any incompatibilist theory of it, freedom as we ordinarily understand it is already vastly unlike ordinary causation. We may reject our ordinary belief in freedom because of its supposed incompatibilist commitment. But we may also reject freedom just because the kind of power envisaged, whether or not consistent with causal

determinism, is too radically unlike any other power we are familiar with and, in particular, too unlike power in causal form. Hobbes's arguments serve to remind us of this radical dissimilarity.

Hobbes's scepticism raises a second issue too. How far are all the problems for freedom that are supposedly raised by incompatibilism really, on closer examination, incompatibilist in origin? Or do some arise as genuine problems, to the extent that they are genuine, from something else: from freedom's identity as a non-causal form of power—a power to determine that operates quite differently from ordinary causation?

Take the randomness problem—the threat that the operation of freedom is left indistinguishable from chance, so that to remove prior necessity is to leave the final outcome to a degree random or dependent on mere chance. Certainly with ordinary causes, if it is not determined in advance what effect a given cause will produce, the outcome must indeed depend, to a degree, on simple chance. If causation is the only power in play, take away prior necessity and you certainly are left with mere chance—chance and nothing else. So to the extent that a cause is merely probabilistic, what effect it will produce depends to a degree on mere chance. But to suppose that in all cases the alternative to necessity is mere chance is to assume that there can be no such power as freedom as we ordinarily understand it—a power involving contingency of determination. For even if the outcome is not already causally predetermined—so that it is initially chancy how the agent will act—freedom, as ordinarily understood, may prevent the final outcome from depending on simple chance. Freedom allows the outcome still to be determined—by the agent. It is arguable, then, that the real target of Hume's scepticism is not freedom conceived in incompatibilist terms, but freedom in a form that involves contingency of determination.

Where freedom is concerned, there are two forms of scepticism. There is scepticism from the supposed conceptual truth of incompatibilism. But there is also scepticism from freedom's basic identity as a power over alternatives distinct from ordinary causation. The second scepticism denies the very possibility of such a power, not because of any incompatibilist theory of it, but because as ordinarily understood, as a power over alternatives, freedom is too radically unlike the causation found in wider nature. It is this second form of scepticism that may prove the most serious. Indeed, it

looks as though, as in Hume's case, some of the first kind of scepticism might really depend on the second. Freedom is indistinguishable from chance only if there can be no such thing as a power that is distinct from ordinary causation—a power to determine alternatives that can operate even in cases where the final outcome is undetermined causally.

6. Conclusion

Our idea of an ordinary cause allows for only two possibilities. Either the cause is powerful enough to determine the outcome it produces—in which case the cause's operation to produce that outcome is fixed by the very nature of its power. The cause has punch—but as a cog within a mechanism has punch. Its presence with the power to operate under given circumstances guarantees its operation when those circumstances arise. Or else, as with a probabilistic cause, the operation of the cause is not fixed but open. The presence of the cause with the power to produce a given outcome does not guarantee that outcome. But then the cause's power is partial. The cause influences what happens, but its operation does not determine the outcome, which remains dependent on mere chance. So either an ordinary cause is powerful, but like a cog within a mechanism, or its operation is reduced to a chance-involving form of weakness.

By contrast freedom, as we ordinarily understand it, is a power that, thanks to contingency of determination, combines the two features, punch and openness, which in ordinary causation always oppose each other. Freedom is a power whose nature never mechanically dictates its exercise. But by contrast to probabilistic causation, this openness does not diminish the power at all. Even if it were initially chancy whether or how the power would operate, the operation of the power can still remove any dependence of the outcome on chance. The operation of the power can still determine without merely influencing.

That we conceive of freedom as involving contingency of determination does not of itself commit us to incompatibilism. Compatibilism remains a possible view. But contingency of determination does explain why incompatibilism remains an intelligible option too—why, on our ordinary understanding of freedom, incompatibilism does not immediately reduce freedom

to nothing more than chance. The intractable conflict in ordinary belief between compatibilist and incompatibilist views of freedom reveals and depends on something often missed today—our underlying conception of freedom as a non-causal form of power.[4]

References

Clarke, Randolph. 2003. *Libertarian Accounts of Free Will*. Oxford: Oxford University Press. https://doi.org/10.1093/019515987X.001.0001

Harre, Rom, and Edward H. Madden. 1975. *Causal Powers: A Theory of Natural Necessity*. Oxford: Blackwell.

Hobbes, Thomas, and John Bramhall. 1656. *The Questions Concerning Liberty, Necessity and Chance, clearly stated between Dr Bramhall Bishop of Derry, and Thomas Hobbes of Malmesbury*. London.

Hume, David. 1978. *A Treatise of Human Nature*. Edited by P.H. Nidditch. Oxford: Clarendon Press.

Locke, John. 1975. *An Essay Concerning Human Understanding*. Edited by P.H. Nidditch. Oxford: Clarendon Press.

Mumford, Stephen, and Rani Lill Anjum. 2011. *Getting Causes from Powers*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199695614.001.0001

Pink, Thomas. 2004. *Free Will: A Very Short Introduction*. Oxford: Oxford University Press. https://doi.org/10.1093/actrade/9780192853585.001.0001

Pink, Thomas. 2017. *Self-Determination* (*The Ethics of Action*, vol. 1). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199272754.001.0001

Sehon, Scott. 2005. *Teleological Realism: Mind, Agency, and Explanation*. Cambridge, MA: MIT Press.

Suarez, Francisco. 1994. *Metaphysical Disputations*, disputation 19: *On causes that act necessarily and causes that act freely or contingently; also, on fate, fortune, and chance* in *Francisco Suarez S.J. on Efficient Causality*. Edited by Alfred Freddoso. New Haven: Yale University Press.

[4] This argument is developed further in my (2017).

Organon F 26 (1) 2019: 169–174
https://doi.org/10.31577/orgf.2019.26110
ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

DISCUSSION NOTE

The Principle of Causal Completeness: Reply to Daniel von Wachter

Michael Esfeld*

This[1] discussion note is motivated by the fact that, to my mind, Daniel von Wachter does not represent the position of the friends of the principle of the causal closure of the physical world (PCC) in a completely fair manner in his paper "The principle of the causal openness of the physical" in this issue. What Kim (1998, chap. 2) and Papineau (2002, chap. 1) among others defend is this principle:

> PCC: For any physical event p, insofar p has a cause, it has a complete physical cause.

PCC is therefore better be called "principle of causal completeness" rather than "closure." PCC is a claim about what holds in the actual universe. Thus, it is a contingent principle. Possible worlds in which PCC is false are not ruled out. PCC is not committed to determinism, since it leaves open to what extent physical events are caused. It only states that if there are causes of physical events, there are complete physical causes.

PCC can be employed in an argument for physicalism in the sense of the metaphysical stance according to which everything that exists—or at least everything that causes something—is physical. To turn PCC into an

[1] This is a reply to (von Wachter 2019).

* University of Lausanne
Department of Philosophy, University of Lausanne, CH-1015 Lausanne, Switzerland
Michael-Andreas.Esfeld@unil.ch Website: http://www.michaelesfeld.com/

© The Author. Journal compilation © The Editorial Board, *Organon F*.

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International Public License (CC BY-NC 4.0).

argument for physicalism, one has to endorse a further principle that excludes systematic overdetermination: if physical events have causes, these causes do not systematically overdetermine these events. Hence, if (by PCC) physical events have physical causes insofar as they have causes at all, there are no additional non-physical causes that systematically cause these events as well. This then yields the exclusion of non-physical causes for physical events. In contrast to PCC, the usual argument to rule out systematic overdetermination is indeed an *a priori* argument against banning causes that do not make a difference. Be that as it may, systematic overdetermination plays no role in the case for the causal openness of the physical that von Wachter seeks to make. It would be of no help for the metaphysical stance according to which non-physical minds (souls) or God can cause physical events: if they do so, they do not overdetermine these events. If they intervene in the physical domain, PCC is false, also in the non-modal version given above.

What, then, is the argument for PCC? The argument is that if the physical laws are true, PCC is true as well. In other words, the argument is that scientific realism with respect to physical laws is committed to PCC. For this argument, it is irrelevant what the physical laws are—that is, whether they are the laws of classical mechanics, of the standard model of elementary particles in quantum field theory, of a future theory of quantum gravity, etc. Furthermore, it is irrelevant whether they are deterministic or probabilistic. The argument is that taking the laws of our physical theories (classical, relativistic, quantum) as guide to the true laws of nature, anyone who is a scientific realist is committed to PCC. Again, scientific realism is not at disposal in von Wachter's paper. The issue hence is whether this argument is correct. In other words, von Wachter has to show us a way how to be a scientific realist and, yet, reject PCC, also in its non-modal version as stated above.

I have reservations about the manner in which von Wachter analyses physical theories with respect to making a case for separating scientific realism from the commitment to PCC. Consider Newtonian physics. According to Newton, every change in the state of motion of a physical object is due to forces influencing the motion of the object. It is true that Newton's second law leaves open what these forces are. Hence, there could in principle also be forces originating in a non-physical agent that influence the motion

of physical objects—provided that these forces respect also the third law, that is, the principle of action and reaction corresponding to each other, so that there also would have to be a reaction on the non-physical agent. The problem with that, obviously, is that a non-physical agent does not have a state of motion. Be that as it may, the central issue is that the three laws of Newton are a framework for the formulation of physical theories. But they are not a physical theory themselves.

The physical theory is gravitation. Von Wachter is right in pointing out that Laplacean determinism follows from Newtonian gravitation only if gravitation is the only force. But this does not open up a way to separate scientific realism from PCC. The point at issue is that Newtonian gravitation sets up the paradigm for a fundamental and universal physical theory, that is, a theory that applies to everything in the universe and that cannot be reduced to another theory. According to Newton's theory of gravity, every physical object interacts with every other physical object in the universe via gravitation (insofar as the objects have a mass). Hence, there is a single law that relates all the objects in the universe with one another. More precisely, the law relates the state of motion of all the objects in the universe at a given time with the state of motion of all the objects in the universe at any other—past or future—time. Formulating the law as a differential equation, on the left hand side figure the positions and velocities of all the objects in the universe and on the right hand side figure the variables that determine the temporal evolution of the positions and velocities in the case of gravitation.

This way of conceiving physical laws applies to modern physics as a whole, whatever the details and differences are in classical, relativistic and quantum physics. Laws are formulated in terms of differential equations such that the variables figuring on the right hand side of the equation indicate what determines the temporal evolution of the objects as they figure on the left hand side of the equation. Over and above gravitation, these are variables for electromagnetism, the weak and the strong interaction. The important issue is that given these variables, the temporal evolution of the objects is fixed. According to our current state of knowledge, the temporal evolution of the physical objects is fixed by these four types of interaction. The fact that we do not have (as yet) one single equation that brings all these four types together is not relevant for present purposes. Furthermore,

it is not relevant for present purposes whether future physics will find more types of interaction of the same kind as gravitation, electromagnetism, the weak and the strong interaction or reduce them to less types of interaction. What is relevant is that one cannot add to this structure of physical theories formulated in terms of differential equations that indicate what determines the temporal evolution of the objects under consideration a clause to the effect that something may intervene from the outside that stops the evolution of the objects as indicated by what figures on the right hand side of the differential equation in question.

Of course, a law that relates the state of motion of all the objects in the universe at one time to the state of motion of all the objects in the universe at other times is as such of no practical use and cannot even be tested. The law has to contain a procedure how to apply it to specific systems in the universe that we can observe or control while abstracting from the rest of the universe. Furthermore, it has to tell us something about the evolution of specific systems that we can expect under ignorance of their exact initial conditions. That is why even a universal, deterministic law has to be linked with a probability measure that enables the derivation of propositions about the evolution of ensembles of systems that are prepared under identical conditions when ignoring their exact initial conditions (such as, for instance, the distribution of the outcomes in a large series of coin tosses). It is clear how to do this both in classical and quantum physics. In brief, these demands do not call the mentioned structure of physical theories in question.

Nonetheless, like von Wachter, I think that the link between scientific realism and PCC is not as firmly established as these considerations may suggest. However, to my mind, the point at issue is not the mentioned structure of physical theories, but the metaphysics of laws. One should be very cautious about using terms such as forces acting on objects. In physics, when objects interact, there is nothing that travels from one object to other objects and that literally pushes them to move in a certain manner. There is no justification to associate forces with agents that literally act on objects. Recall the following statement of Russell in his famous critique of the notion of cause:

> In the motions of mutually gravitating bodies, there is nothing that can be called a cause and nothing that can be called an

> effect; there is merely a formula. Certain differential equations can be found, which hold at every instant for every particle of the system, and which, given the configuration and velocities at one instant, or the configurations at two instants, render the configuration at any other earlier or later instant theoretically calculable. [...] This statement holds throughout physics, and not only in the special case of gravitation. (Russell 1912, 14)

Whatever one thinks about Russell's views on causation and his philosophy in general, what he says about the physics here is certainly correct. This is already evident if one considers the formulation of physical determinism: the propositions stating the laws of nature and the propositions describing the state of the world at an arbitrary time (i.e. the propositions describing the initial conditions) entail the propositions describing the state of the world at any other time. Hence, determinism in physics is—only—about entailment relations among propositions. There is nothing here that suggests that what there is in the state of the world at a given time literally produces or brings about the evolution towards the future. The question for a scientific realist then is what it is in the world that makes these propositions true, that is, in virtue of what in the ontology these entailment relations among propositions hold.

As an example of a parsimonious ontology of physics that is nevertheless rich enough for these propositions to come out true in virtue of what there is in the world, consider Humeanism about laws of nature. On Humeanism, first comes the motion of the matter in the universe, then come the laws. They are theorems of the best system, that is, the system that strikes the best balance between being simple and being informative in representing the motion of the matter in the universe. Hence, the laws are only fixed on the basis of the motion of the matter throughout the entire history of the universe. What is relevant in the present context is that Humeanism about laws is compatible with the idea that some motion of matter in the universe depends on the free will of minds that may be non-physical entities, and, yet, when it comes to representing the motion of the matter in the universe in the best system, the variables that enter into this system are only ones for universal physical interactions [see my paper in this volume (Esfeld 2019)].

On the stance that can be dubbed Super-Humeanism, what the Humean says about laws applies to the dynamical variables that are functionally defined in terms of their role for the particle motion—such as mass, charge, energy, fields, an initial wave function in quantum physics, etc. Their values are not intrinsic to the state of the universe at any given time, but depend on the motion that the matter in the universe actually takes [see (Esfeld and Deckert 2017, chap. 2.3)]. Consequently, again, what these values are depends on the actual motion of matter, and there is the conceptual space open for maintaining that these values are—in part—influenced by motions that humans make out of their (non-physical) free will. To put it in a nutshell, if humans had chosen to do otherwise out of their (non-physical) free will, on Humeanism, the laws of nature would be slightly different [see (Beebee and Mele 2002)]; on Super-Humeanism, the initial values of some of the dynamical variables that enter into the laws of nature would have been slightly different.

In sum, when entering the metaphysics of mind, it seems to me that we should not touch upon the sketched structure of physical theories and scientific realism as regards that structure. The metaphysics of laws of nature is rich enough to leave the conceptual space for a large variety of positions in the metaphysics of the mind open.

References

Beebee, Helen, and Alfred Mele. 2002. "Humean Compatibilism." *Mind* 111 (442): 201–23. https://doi.org/10.1093/mind/111.442.201

Esfeld, Michael. 2019. "Why Determinism in Physics Has No Implications for Free Will." *Organon F* 26 (1): 62–85. https://doi.org/10.31577/orgf.2019.26105

Esfeld, Michael, and Dirk-André Deckert. 2017. *A Minimalist Ontology of the Natural World*. New York: Routledge.

Kim, Jaegwon. 1998. *Mind in a Physical World. An Essay on the Mind-Body Problem and Mental Causation*. Cambridge (Massachusetts): MIT Press.

Papineau, David. 2002. *Thinking about Consciousness*. Oxford: Oxford University Press. https://doi.org/10.1093/0199243824.001.0001

Russell, Bertrand. 1912. "On the Notion of Cause." *Proceedings of the Aristotelian Society* 13: 1–26.

von Wachter, Daniel. 2019. "The Principle of the Causal Openness of the Physical." *Organon F* 26 (1): 40–61. https://doi.org/10.31577/orgf.2019.26104

Organon F 26 (1) 2019: 175–184
https://doi.org/10.31577/orgf.2019.26111
ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

DISCUSSION NOTE

Do the Laws of Nature Entail Causal Closure? Response to Michael Esfeld

Daniel von Wachter*

1. Michael Esfeld's objection

(1.1) In[1] the article 'The Principle of the Causal Openness of the Physical' I claim that for an argument from the principle of the causal closure of the physical (PCC) for physicalism one needs a modal version of PCC: 'There cannot be a physical event that is not necessitated by preceding physical events.' Michael Esfeld objects that Kim's and Papineau's PCC should be understood as the non-modal claim that

> *For any physical event, insofar p has a cause, it has a complete physical cause.*

Therefore, Esfeld rightly suggests, 'causal completeness' would be a more suitable name than 'causal closure.' By a 'complete physical cause' Esfeld means not only a cause that includes all the partial causes and thus all the physical events that contribute to the causing, but one that includes all that is required for the causing, so that any additional causes would constitute causal overdetermination. But the term 'complete physical cause' neither implies that the cause necessitates the effect nor does it imply that the

[1] This is a response to (Esfeld 2019).

* International Academy of Philosophy in the Principality of Liechtenstein
International Academy of Philosophy in the Principality of Liechtenstein, Fürst-Franz-Josef-Str. 19, FL-9493 Mauren, Liechtenstein
epost@von-wachter.de Website: http://von-wachter.de

© The Author. Journal compilation © The Editorial Board, *Organon F*.

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International Public License (CC BY-NC 4.0).

cause is non-probabilistic. Esfeld clarifies this by saying that 'PCC is not committed to determinism, since it leaves open to what extent physical events are caused.'

(1.2) In my article my response to the argument from non-modal PCC for physicalism was that the non-modal PCC can be justified only by evidence against the existence of non-physical objects. It therefore cannot be used in an argument for physicalism. Esfeld replies that the non-modal PCC can be known through the laws of nature: If the laws of nature are true, then non-modal PCC is true, because the laws describe how the universe develops, and they refer only to physical forces as causes. Or rather, the laws describe a correlation between how the universe develops and physical forces. For example, according to determinism, 'the propositions stating the laws of nature and the propositions describing the universe at an arbitrary time entail the propositions describing the state of the world at any other time.' If Esfeld's claim that the laws of nature entail PCC were true, then we could derive PCC from our knowledge of the laws. So the question before us is whether it is true that the laws of nature entail PCC.

(1.3) Esfeld's crucial assumption, which is shared by the majority of contemporary philosophers of science, is that laws of nature are differential equations that describe the evolution of physical systems, they describe what happens when. This entails the assumption that laws entail regularities of succession of the type 'All events of type x are followed by events of type y.'[2] All the prominent philosophical theories of laws of nature share this assumption, they differ only in whether laws are *more* than regularities of succession and in what this 'more' is.

A 'fundamental and universal physical theory' describes the evolution of the whole universe. It would be a differential equation that links states of the physical universe at one time with states of the physical universe at other times. Let me call this the 'Comprehensive Differential Equation' (CDE). If things always happen as descibed by this CDE, then the non-modal PCC is true.

[2] That this assumption is false is argued in (Wachter 2015).

My response to Esfeld is that evolution equations are distinct from the laws of nature from which they are derived and that evolution equations require a no-further-causes clause while laws do not.

2. Laws of nature are not differential equations

(2.1) Laws of nature are different from the predictions and differential equations that we can derive from them. Let me first present the arguments by John Earman and John T. Roberts for this claim. They have argued that the fundamental laws do not contain *ceteris paribus* clauses, and that if a formula does contain *ceteris paribus* clauses, then it is not a fundamental law but part of a 'work-in-progress theory' (Earman and Roberts 1999; Earman, Roberts and Smith 2002).

> 'If laws are needed for some purpose, then we maintain that only laws will do, and if "*ceteris paribus* laws" are the only things on offer, then what is needed is better science, and no amount of logical analysis on the part of philosophers will render the "*ceteris paribus* laws" capable of doing the job of laws.' (Earman and Roberts 1999, 466)

(2.2) While they hold that laws do not require *ceteris paribus* clauses, they say that *applications of a theory* require what Carl Hempel called 'provisos.' As an example of an application of a theory, Hempel considers a description of the motion of two bodies that are 'subject to no influences from within or from outside the system that would affect their motions.' (Hempel 1988, 158) Earman and Roberts call such an application of a theory a 'differential equation of the evolution type' (Earman, Roberts and Smith 2002, 285); we can abbreviate this as 'evolution equation.' Discussing the proviso required for a system description, Hempel touches the issue of miracles:

> The proviso must [...] imply the absence [...] of electric, magnetic, and frictional forces; of radiation pressure; and of any telekinetic, angelic, or diabolic influences. [(Hempel 1988, 158), also quoted in (Earman and Roberts 1999, 444)]

So Hempel recognises the possibility of divine interventions, and he could have added interventions by souls. The proviso must state that the evolution equation does not apply to cases where there are such influences. To achieve this objective, he proposes the proviso, 'the total force acting on each of the two bodies equals the gravitational force exerted upon it by the other body; and the latter force is determined by the law of gravitation.' The expression 'total force' is supposed to exclude telekinetic, angelic, diabolic etc. influences. As a diabolic influence would be an action and, in my view, not a 'force' in the Newtonian sense, Hempel should say instead that 'nothing besides the gravitational force exerted by the two bodies is affecting their motion.' The point is clear: First, while laws of nature do not require *ceteris paribus* clauses or 'provisos,' evolution equations do. Secondly, *ceteris paribus* clauses are not enough, what is required is the no-further-causes clause that implies that there are no other things acting on the objects.

(2.3) What then is a law of nature, in contrast to an evolution equation? According to Earman and Roberts, the law of gravitation, for example, asserts this:

> (Regardless of what other forces may be acting) any two massive bodies exert a gravitational force on one another that is directly proportional to the product of their masses and inversely proportional to the square of the distance between them. (Earman and Roberts 1999, 473, footnote 14)

(2.4) So the evolution equations are not themselves properly called 'laws of nature.' Rather, $F = G\frac{m_1 m_2}{d^2}$ is a laws of nature, and from this and the other laws one can derive differential equations that describe the development of systems in which nothing except the forces described by the laws used are affecting the movement of the two bodies. If there is a fundamental theory of physics, then perhaps a CDE can be derived from it, but CDE will not be the fundamental theory of physics itself, and it will describe the actual universe only if no non-physical objects are acting on it and only if no physical things have agent-causal powers.

> [D]ifferential equations of evolution type are not laws; rather, they represent Hempel's applications of a theory to a specific case.

> They are derived using (unhedged) laws along with non-nomic modelling assumptions that fit (often only approximately) the specific case one is modelling. Because they depend on such non-nomic assumptions, they are not laws. For example, because Kepler's 'law' that planets travel in ellipses is derived from laws together with the assumption that there are only two bodies in the universe, it is not a law in spite of the normal nomenclature. [...] The 'law of free fall' is a consequence of a differential equation that involves the assumption that there is no resistance from the wind. That too is a non-nomic assumption, for it is not a law that there is no resistance from the wind. It seems to us that the role played by idealizations in physics is typically found here, in the derivation of differential equations, rather than within the laws themselves. (Earman, Roberts and Smith 2002, 286)

(2.5) Let me formulate Earman's and Roberts's point with my terminology. A law of mechanics says that there are forces of certain kinds in situations of certain kinds. More generally, using the notion of 'directedness' that I have introduced in the article (§ 4.14), A law of nature states that events of a certain kind are the bases of directednesses of a certain kind.

(2.6) Thus, contrary to Esfeld's view, a law of nature does not say that events of kind *x* are always followed by events of type *y*. Laws do not entail regularities of succession. They speak not about what happens when but about Newtonian forces or, more generally, directednesses. As John Stuart Mill pointed out:

> All laws of causation, in consequence of their liability to be counteracted, require to be stated in words affirmative of tendencies only, and not of actual results. (Mill 1843, book III, chap. 10, § 5)

(2.7) Evolution equations can be derived from laws by applying them to a particular system. From evolution equations we can derive statements of the form 'States of affairs of type *x* on which nothing else is acting than the factors described in the equation are followed by events of type *y*.' Evolution equations describe the evolution of systems on which the factors that are taken into account in the equation and nothing else is acting. They require

the *no-further-causes clause*. Laws of nature do not require such a clause because they describe, in the case of mechanics, forces, not actual movements, and the forces obtain under all circumstances (and in my view even necessarily) as the laws describe them.

3. The universal physical theory

(3.1) A comprehensive differential equation (CDE), which Esfeld calls the 'universal physical theory,' is an evolution equation that takes into account all the physical forces that exist. For the sake of the argument I grant that a CDE that can be applied to a state of the universe might exist, although the fact that we cannot even find a general closed-form solution for three-body systems raises doubts about this.

(3.2) Like all evolution equations, CDE would reqire the no-further-causes clause 'if nothing else is acting on the objects.' One might object that a CDE will not require this, because it will take into account all factors. But first, CDE will not only be applicable to complete stages of the universe, but to all physical states of affairs. If it is applied to some state of affairs that is a part of the universe, then, because other things could affect the development of the system, the no-further-causes clause is required.

(3.3) If CDE is applied to a state of the whole universe, then there are still three ways how CDE could fail to apply. First, there could be physical forces that are not described by laws of nature. That is, there could be forces for which we cannot formulate a law that says that in situations of a certain type there are such forces. It is surprising that we can know any laws of nature and that we can make any predictions at all. Perhaps God created the universe so that all physical forces are governed by laws of nature, but it is a possibility that there are other physical forces, and the no-further-causes clause takes into account that possibility.

Second, perhaps there could be physical objects that can be agent causes. That is, they can by acting bring about choice events, i.e. events that are not the result of a law-governed causal process and thus have no preceding cause but occur through an action. A dualist will hold that choice events are brought about by souls, but somebody who holds the view that

some material objects have consciousness can also hold that some material objects can bring about choice events.

Third, there could be a God or souls that sometimes cause physical events, in which case CDE does not apply. The data which justify CDE and the laws of nature justify only

- the proposition that there are the forces described by the laws, and
- the proposition that consists of CDE and the no-further-causes clause, as well as all other evolution equations with the no-further-causes clause.

They do not justify the assumption that there are no non-physical things acting on the system. That is a metaphysical thesis that cannot be justified by physical experiments. Given my arguments for openness, it can only be justified by defeating the putative evidence for the various possible non-physical objects. I conclude if we shall ever find a CDE, then its application will also require the no-further-causes clause und thus it will not entail PCC.

(3.4) CDE requires the no-further-causes clause for the same reason why all evolution equations require the no-further-causes clause: Physical causes are open to cooperation and to intervention, as I have argued in the article. If a physical state of affairs S_1 caused S_2, then something could have prevented S_2 so that S_1 would have occurred but not S_2. That this is true can be seen considering our experience of causal processes. Rolling billiard balls and tidal waves can be stopped. Also the process that is constituted by the complete physical universe is stoppable. Of course, if there are no non-physical objects, then that process is never stopped. So would not CDE then be true also without the no-further-causes clause? Yes, but the proposition that CDE applies without the no-further-causes clause would not be justified by the experiments and observations through which we know the laws of nature.

(3.5) Let me clarify my position by commenting on some of Esfeld's statements. He writes: 'According to Newton, every change in the state of motion of a physical object is due to forces influencing the motion of the object.'

My response: This assumption entails that Newton's laws of motion exclude the existence of physical choice events and of physical events that

have non-physical causes. Regardless of what Newton's view was, Newton's formulae can and should be used without this assumption. Newtonian physics describes which kinds of forces there are in which kinds of situations and says that if force F and nothing else is acting on an object, then it accelerates with $a = F/m$. The assumption that there are no choice events and no physical events that have non-physical causes is not justified by the observations and experiments which justify Newton's laws.

(3.6) Esfeld denies that evolution equations require the no-further-causes clause: '[O]ne cannot add to this structure of physical theories formulated in terms of differential equations that indicate what determines the temporal evolution of the objects under consideration a clause to the effect that something may intervene from the outside that stops the evolution of the objects as indicated by what figures on the right hand side of the differential equation in question.'

My response: Evolution equations describe the development of a particular system only if no further forces are acting on the objects and only if no non-physical things are acting on the objects. Some might object that in order to do physics, we need to assume that all physical events are the results of causal processes and that we need no no-further-causes clause. This is what is called 'methodological naturalism.' But neither no-further-causes clauses nor divine or human interventions would impede physics or violate the laws.[3] To the contrary, predictions without the clause are not justified through the experiments and may turn out to be false. However, I agree that the fundamental physical theory is not an evolution description with a no-further-causes clause. It is not an evolution description or a differential equation at all, but a formula that is to be interpreted as saying that there are directednesses of certain kinds in situations of certain kinds. Like all laws of nature, it is true without no-further-causes clause and without *ceteris paribus* clause, and in my view, which I have not defended here, even necessarily. But it does not entail PCC.

(3.7) Esfeld suggests that 'the point at issue is [...] the metaphysics of laws.' 'One should be very cautious about using terms such as forces acting on objects. In physics, when objects interact, there is nothing that travels

[3] More on this in (Wachter 2015).

from one object to other objects and that literally pushes them to move in a certain manner. There is no justification to associate forces with agents that literally act on objects.'

My response: Does the view that there are no pushing forces—let us call it the Russellian view, because some have contested that Hume held it—fit better with the view that evolution equations require the no-further-causes clause or with the view that they do not? The denier of pushing forces cannot say, as I do, that laws are not evolution equations but describe that there are forces of certain kinds in situations of certain kinds. He has no plausible alternative to saying that laws are evolution equations. However, evolution equations with the no-further-causes clause lack the universality which we associate with laws. The no-further-causes clause destroys the entailment 'The propositions describing the initial condition of the system and the laws entail the propositions that describe the later states.' Therefore perhaps some deniers of pushing forces are drawn towards accepting evolution equations without the no-further causes clause.

On the other hand, I object to the Russellian and Humean view that the truth of evolution statements *without* the no-further-causes clause would be highly unlikely on the assumption that there are no pushing forces and no causal connections. In my view, the truth of the statement that includes the differential equations and the no-further-causes clause is evidence for the existence of pushing forces which can be counteracted and of directednesses.[4]

(3.8) Esfeld's theory of laws is the majority view in philosophy of science. But why is it the majority view? What reasons do we have for believing it? Look at, for example, the law of gravitation, $F = G\frac{m_1 m_2}{d^2}$, or at some more complicated law. I do not see anything there that suggests the majority view, because the formula says only something about forces. We can derive differential equations, but for systems that include less than the whole universe all agree that the differential equations describe only the evolution of those systems in which no further forces are acting than those taken into account in the equation. What reason do we have for believing that only

[4] That laws *explain* regularities has been argued by Armstrong (1983, chap. 4) and by Foster (2004, chap. 3).

forces and not agent causes can affect the system? Why should the differential equation that describes the universe not require the no-further-causes clause? And why should one think that the differential equations rather than the formulas from which they are derived are the laws and are in some sense more fundamental?

For the reasons given, I suggest that we should question the view that laws are evolution statements and entail regularities of succession and replace it by one that posits directednesses or, to use Mill's term, tendencies instead of regularities of succession. The laws, then, do not entail the principle of causal closure of the physical.

References

Armstrong, David M. 1983. *What Is a Law of Nature?* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781316499030

Earman, John, and John T. Roberts. 1999. "'Ceteris Paribus,' There Is No Problem of Provisos." *Synthese* 118 (3): 439–78. https://doi.org/10.1023/A:1005106917477

Earman, John, John T. Roberts, and Sheldon Smith. 2002. "Ceteris Paribus Lost." *Erkenntnis* 57 (3): 281–301. https://doi.org/10.1007/978-94-017-1009-1_1

Esfeld, Michael. 2019. "The Principle of Causal Completeness: Reply to Daniel von Wachter." *Organon F* 26 (1): 62–85. https://doi.org/10.31577/orgf.2019.26105

Foster, John. 2004. *The Divine Lawmaker: Lectures on Induction, Laws of Nature, and the Existence of God*. Oxford: Clarendon. https://doi.org/10.1093/0199250596.001.0001

Hempel, Carl Gustav. 1988. "Provisos: A Problem Concerning the Inferential Function of Scientific Laws." *Erkenntnis* 28 (2): 147–64.

Mill, John Stuart. 1843. *A System of Logic Ratiocinative and Inductive*. New York: Harper, 1882.

von Wachter, Daniel. 2015. "Miracles Are Not Violations of the Laws of Nature Because the Laws Do Not Entail Regularities." *European Journal for Philosophy of Religion* 7 (4): 37–60. https://doi.org/10.24204/ejpr.v7i4.86