

# Open Research Repository

## Physical Determinism, Zygote- Manipulation and Responsible Agency

Item Type	Journal article
Authors	Huoranszki, Ferenc
DOI	<a href="https://doi.org/10.1007/s11406-020-00307-1">10.1007/s11406-020-00307-1</a>
Publisher	Springer
Rights	CC BY 4.0
Download date	2024-08-06 13:39:12
Item License	<a href="https://creativecommons.org/licenses/by/4.0/">https://creativecommons.org/licenses/by/4.0/</a>
Link to Item	<a href="http://hdl.handle.net/20.500.14018/13896">http://hdl.handle.net/20.500.14018/13896</a>



# Physical Determinism, Zygote-Manipulation and Responsible Agency

Ferenc Huoranszki<sup>1</sup>

Received: 24 August 2020 / Revised: 14 December 2020 / Accepted: 17 December 2020 /  
Published online: 12 January 2021

© The Author(s) 2021

## Abstract

Agents have no control over the formation of their own zygote. Others may do. According to a well-known argument, the so-called Zygote Argument for incompatibilism, these facts, together with a prima facie plausible further assumption, are sufficient to prove that human agents cannot be responsible for their actions if they live in a deterministic universe. This paper argues that the lack of agents' control over the constitution of their own zygote can undermine their responsibility only in exceptional conditions and that the occurrence or non-occurrence of those conditions has nothing to do with the truth or falsity of determinism. What undermines agents' responsibility in the situations described by the Zygote Argument is the occurrence of some specific initial conditions which may render the manipulation of agents' behaviour possible, and not the truth of determinism.

**Keywords** Responsibility · Compatibilism · Physical determinism · Manipulation

By incompatibilism I understand the view that the truth of a type of determinism, which I shall call physical determinism, is sufficient to undermine agents' freedom and responsibility.<sup>1</sup> Physical determinism is a thesis about the nomological structure of the universe. Roughly, it is the thesis that the dynamical laws of physics together with the complete specification of the physical state of the universe at one instant entail its physical states at any other instant. I will offer a more precise characterization of physical determinism later. In order to be able to formulate the claims of the present

---

<sup>1</sup>I borrow the term 'physical determinism' from Ginet 1990. See also van Inwagen 1983. Some incompatibilists hold that responsibility is incompatible with both the truth and the falsity of determinism, and hence freedom and responsibility are unconditionally impossible (see e.g., Pereboom 2001). I am interested here only in the compatibility of responsibility with physical determinism. I shall also assume that the truth of physical determinism does not imply psychological, social, or genetic determinism.

✉ Ferenc Huoranszki  
huoransz@ceu.edu

<sup>1</sup> Philosophy Department, Central European University Private University, Quellenstrasse 51,  
A-1100 Vienna, Austria

paper, however, I need to begin with two preliminary remarks about the relevant notion of determinism.

First, for reasons which shall become clear later, the notion of determinism, as I understand it here, essentially involves *laws*, the knowledge of which makes certain expectations rational and certain predictions possible. However, even if all laws of physics are deterministic, rational expectations and reliable predictions about the physical future might not be possible. Hence, I use here a stronger notion of physical determinism than what deterministic physics usually requires.

Second, it is important to distinguish the notion of *physical* determinism from other forms of determinism. There are biological, psychological and sociological laws which may or may not be deterministic. Correspondingly, we may talk about the biological, psychological or social determination of behaviour. However, I shall assume that such forms of determinism do not entail, neither are they entailed by, the truth of physical determinism. My interest is here confined to the possible consequences of physical determinism.

Compatibilism is the view according to which incompatibilism is false. According to compatibilism, the truth of physical determinism does not entail the impossibility of freedom and responsibility. An influential argument by Alfred Mele, which is called the ‘Zygote Argument’ or ‘Original-Design Argument’, poses a serious challenge to compatibilism so understood. The purpose of this paper is to show that the Zygote Argument cannot undermine rational belief in compatibilism. I do not deny that the possibility of manipulating or creating zygotes with a certain molecular structure can, *if certain conditions are satisfied*, menace the freedom and responsibility of the agents who develop from such zygotes. But I shall argue that the relevant conditions have nothing to do with the truth and falsity of physical determinism.

## 1 The Zygote Argument

The Zygote Argument is based on a thought experiment. Imagine a superhuman being with extraordinary knowledge about physics. This being, like Laplace’s famous demon, knows the precise content of all physical laws that govern the evolution of the universe. She also knows the physical state of the universe at a specific instant down to the last detail. If physical determinism in the sense characterized above is true in the universe, then her knowledge about the laws and the global state of the universe allow her to know the exact physical states of the universe at all future instants as well.<sup>2</sup>

Let us call this superhuman being – following Alfred Mele’s original statement of the Zygote Argument – Diana. Diana does not only have superhuman theoretical knowledge, she has also extraordinary technical knowhow. She can create zygotes. She does then create one and implant it in a woman (called Mary). More specifically, imagine that Diana

---

<sup>2</sup> This assumption is heavily counterfactual. From the knowledge of physical states and the knowledge of dynamical laws it does not follow that anyone can compute the exact future physical states of the universe; and not for contingent but for necessary, arithmetical reasons. It is one thing to know the differential equations which represent the evolution of certain physical states; it is another to have unique and computable solutions for them. So, we need to assume that Diana in the thought experiment has not only superior knowledge about physics, but also a new mathematics that is ‘superior’ to the one known by us. I am not going to question this possibility here.

combines Z's atoms as she does because she wants a certain event E to occur thirty years later. From her knowledge of the state of the universe just prior to her creating Z and the laws of nature of her deterministic universe, she deduces that a zygote with precisely Z's constitution located in Mary will develop into an ideally self-controlled agent (called Ernie) who, in thirty years, will judge, on the basis of rational deliberation, that it is best to A and will A on the basis of that judgment, thereby bringing about E. (Mele 2008, 279).

Suppose, further, that, initially, we would hold agents who are ideally self-controlled and act on the ground of their rational deliberations responsible for their actions. How exactly we should understand self-control and rationality is, of course, a further philosophical problem. However, since self-control and rational deliberation require the exercise of some psychological capacities the possession of which seems to be independent of the truth or falsity of physical determinism, we can, at least *prima facie*, assume that an agent can be ideally self-controlled, and his actions can be based on rational deliberation, even if physical determinism is true. This is, of course, not an uncontroversial assumption. But without this assumption the zygote argument cannot even get off the ground.

Mele claims that (P1) because of the way in which Ernie's zygote was created in the deterministic universe, Ernie cannot be responsible for doing A. However, (P2) from the point of view of agents' freedom and responsibility, there is not any relevant difference between the case as described in the example and a case in which a zygote naturally comes to exist in a deterministic universe. Consequently, (C) if physical determinism is true, then no human being can be free and responsible.<sup>3</sup>

Standard responses to this argument challenge the truth either of its first or of its second premise; or claim that the truth of the first premise is not compatible with the truth of the second.<sup>4</sup> In this paper, I shall argue that there is a more fundamental problem with the argument; a problem that has interesting implications about debates over compatibilism in general.

The fundamental problem with the Zygote Argument is that we can infer from the example to the falsity of compatibilism only if we also grant the following assumption:

( $\alpha$ ) If there is a deterministic universe in which the sort of intervention described by the zygote-example is physically possible, then it *must* be possible in *every* deterministic universe in which human agents develop from a zygote.

In the first part of the paper, I shall explain why  $\alpha$  is necessary for the argument to prove the falsity of compatibilism (or at least to support 'incompatibilist intuitions'). In the

<sup>3</sup> This is how Alfred Mele presents the argument in Mele 2008, 280. He first introduced a variation of this argument in Mele 1995. I need to make two initial remarks about the aim and the presentation of the argument. First, Mele does not claim that the argument proves incompatibilism; but he seems to think that it provides sufficient ground for his agnosticism about the compatibility of responsibility and determinism. The argument is, however, often interpreted as an attempt to refute compatibilism. Second, as a reviewer pointed out, in this formulation the argument's first premise does not only state that Ernie is not responsible because of the way in which his zygote was created by Diana; it also seems to assign already some role to determinism, which is usually meant to be introduced only by premise 2. Nonetheless, this is how the argument was presented by Mele and I decided to follow his own presentation as faithfully as possible, since nothing I will argue for hinges on this.

<sup>4</sup> See respectively Fischer 2011, Schlosser 2015, and Kearns 2012.

second part, I shall argue that  $\alpha$  is false. But if without  $\alpha$  the zygote argument cannot prove the falsity of compatibilism and  $\alpha$  is false, then the argument cannot provide any further support for incompatibilism, neither can it challenge the truth of compatibilism.

I shall argue that  $\alpha$  is false because whether or not the responsibility-undermining manipulation described by the zygote-case is possible *does not depend at all* on considerations about physical determinism. It depends only on the *specific, local causal structure* of the world. And whether or not such local causal structures can obtain is entirely independent of the truth of physical determinism. Thus, physical determinism is neither sufficient nor necessary for the possibility of manipulating an agent's behaviour by creating his or her zygote. I shall end the paper by explaining the significance of this result.

## 2 An Analogue

In order to explain the importance of assumption  $\alpha$  for the argument, let us compare it to another imaginary situation, which requires much less sophisticated knowledge than knowing the complete physical state of the universe at one instant or of the laws of nature does. Imagine that Diana has superhuman knowledge of every randomly selected agents' preferences and she also has the power to manipulate Ernie's brain so that she can ensure what decisions he is going to make. Suppose, further, that Ernie participates in a certain collective decision-making procedure in which the distribution of preferences of the other participants is completely evenly balanced. Diana, given her knowledge about people's preferences, can predict with almost absolute certainty how they will choose about the issue at hand. She also wants that the procedure end with a specific outcome O. Then, she uses her power to manipulate Ernie's decision and thereby achieves that the procedure ends with the collective decision favouring O.

If this type of situation regularly arises, and hence it is always *Ernie's will* that determines which collective decision shall be reached, then Ernie acts as a 'dictator'. This might shake our confidence in the fairness of democratic decision making; or at least, in the fairness of the application of simple majority rule. And our confidence would certainly be shaken even if Ernie's decisions had not been manipulated by Diana. For if, *across various contexts* (from local meetings to parliamentary elections) it is *always* Ernie's will which determines the outcome of collective decisions, then he emerges as a 'dictator' irrespective of whether or not his decisions were brought about by Diana's interference.<sup>5</sup>

But should this imaginary situation really undermine our belief in the fairness of democratic voting procedures *as such*? More generally, should we conclude that democratic procedures are just 'forms of (personal) dictatorship'? I do not think any of us would think so. The most rational response to this imagined case is that Ernie does not become a 'dictator' *because of* the application of the simple majority rule. He has become a 'dictator' only because of the highly improbable recurrence of some very specific structure of the distribution of other agents' preferences. The *mere possibility* of this scenario should not undermine our *overall* belief in the fairness of majority rule.

Diana cannot manipulate the results of the voting procedure in every possible situation. All she can do is to manipulate the results *in very specific circumstances*:

<sup>5</sup> My imagined case is similar to, though much simpler and requires the occurrence of more specific conditions than, Arrow's famous impossibility theorem in social choice theory (see Arrow 1963, Chapter 2).

in the circumstances in which Ernie's vote is *decisive* with regard to the outcome of a procedure. Thus, knowing the distribution of the voters' preferences and having the power to manipulate Ernie's decisions does not entail that Diana also has the power to determine in every case what will happen in the future. In order to have that power the distribution of agents' preferences must also display a special structure. Otherwise, Diana cannot tilt the balance, as it were, by manipulating Ernie's decision. Similarly, we cannot claim about Ernie that he acts as a 'dictator'.

I suggest that the case about majority rule is analogous to the case as described by the Zygote Argument. Diana's knowledge and power are not sufficient for achieving her aims even if determinism is true. As I shall explain in the sequel, the specific state of the physical world the evolution of which she interferes at a point must be extremely 'hospitable' to her purposes too. Otherwise designing Ernie's zygote cannot determine which specific actions he will perform in the future. However, as Marcus Schlosser notes, 'that Diana can control Ernie's entire development and his behavior over a period of 30 years merely by arranging the initial properties of Ernie's zygote in a certain way ... is truly and utterly incredible' (Schlosser 2015, 77).<sup>6</sup>

Incredible or not, it is certainly logically possible. However, I shall argue that *the mere logical possibility* of this kind of manipulation is not sufficient to prove the impossibility of responsibility in physically deterministic worlds. For, just as in the imaginary case about collective decisions, Ernie's unfreedom and non-responsibility is explained in the zygote-case by the 'hospitality' of the circumstances, that is, by the presence of certain peculiar initial conditions of the universe in which the zygote was created, and not by the truth of physical determinism as such.

### 3 Conditions of Responsibility

As even most incompatibilists admit, there are some conditions of responsibility that are independent of the truth or falsity of physical determinism.<sup>7</sup> For instance, an agent can be responsible only if she is not ignorant about what she does; if she is not pathologically compelled; if her behaviour is responsive to her reasons; if she is ideally psychologically self-controlled etc. I shall call these conditions the non-physical conditions of responsibility (NPCR for short).<sup>8</sup>

<sup>6</sup> However, Schlosser concludes from this observation that P2 should be rejected, and I do not see why this should follow. Of course, if we interpret the notion of 'deterministic universe' in P2 to mean *any* deterministic world in which human zygotes are formed, then most of what I shall argue for is in complete agreement with Schlosser's view. But I grant a weaker interpretation of P2 and hence, I hope, my objection to the zygote-argument is stronger (that is, I do not attribute a view to my opponent which is relatively easy to reject). According to my interpretation, P2 claims that if two worlds  $W_1$  and  $W_2$  differ *only to the extent* that in  $W_1$  Ernie's zygote was created by Diana whereas in  $W_2$  it has come to exist 'in the natural way', then Ernie must be *responsible* for his adult behaviour *to the same extent* in both worlds. So interpreted, we can grant the truth of P2: if Ernie is not responsible for A-ing at  $t_3$  in  $W_1$ , then neither is he responsible for A-ing at  $t_3$  in  $W_2$ . Yet, I shall argue, this does not prove the impossibility of agents' responsibility in *all* deterministic worlds in which human beings develop from a zygote.

<sup>7</sup> See for instance Ginet 1990 or Kane 1996.

<sup>8</sup> Some of these conditions are physical, of course, in the sense that they consider an agent's physical abilities and local physical circumstances. What I mean by 'non-physical' here is only that they are independent of what we think about the truth of physical *determinism* as characterized earlier.

The exact list of NPCR and their interpretation is a contentious issue. Some of the mentioned conditions might not even be necessary for responsibility. But this does not matter for our purposes. What incompatibilists need to prove is that NPCR *can never be sufficient* for an agent's freedom and responsibility. And what the Zygote Argument aims to establish is that *even if* Ernie's action satisfies all non-physical conditions of responsibility, he is not free and responsible if he lives in a physically deterministic world.

Let  $t_1$  be any time in the history of the universe that precedes Diana's intervention. Let  $t_2$  be the time when she intervenes; and  $t_3$  the time when Ernie will do, or avoid doing, some action A that brings about event E. According to the argument's story, Diana creates a zygote *by locally intervening into the history of the universe at  $t_2$* . Diana does not *create* an entire universe or choose the actualization of a specific world from the infinitely many physically possible ones. If she had that power too, then the zygote argument would address the problem of theological fatalism; that is, the compatibility of freedom and responsibility with God's omnipotence and omniscience. But the arguments' aim is to show something about the consequences of *physical* determinism and not of divine power.

What Diana wants to ensure then is that (a) Ernie will perform an action A at  $t_3$  (b) in the circumstances in which all NPCR are satisfied. However, from the conceivability of the fact that there is *a* deterministic universe in which this is physically possible, it does not follow that Diana can control Ernie's future action in the way she does *because* that universe is deterministic. For Diana's responsibility-undermining control is not possible in *every* deterministic universe in which human agents develop from a zygote. And if there are many deterministic universes in which it is *physically impossible* to control Ernie's future actions in the way Diana wants to control them in the example, then the zygote argument cannot prove (support) what it aims to prove (support), that is, that physical determinism is incompatible with responsibility. In fact, as I shall argue later, we can design a sort of zygote-argument about indeterministic universes as well. Thus, whatever explains our intuitions about agents' responsibility in the zygote-creation scenarios, it must be independent of the truth or falsity of physical determinism.

The zygote argument, just as other versions of the manipulation argument,<sup>9</sup> aims to prove that even if all NPCR are satisfied, agents cannot act freely and hence cannot be responsible for their behaviour if they live in a physically deterministic universe. The argument assumes, firstly, that if the universe is deterministic, then Diana *can* create a zygote the physical configuration of which at  $t_2$  ensures that Ernie will perform action A at  $t_3$ . And if she succeeds, then, even if all NPCR are satisfied, Ernie is not responsible for performing A and A's consequences. But then, secondly, it also assumes that there is not any relevant difference between the imagined deterministic worlds in which Diana creates Ernie's zygote and in which his zygote comes to exist in the ordinary way. In either case, whether or not Ernie shall do A at  $t_3$  ultimately depends on his zygote's physical configuration, which is, patently, beyond his own

<sup>9</sup> I assume as most, though not all, philosophers that the zygote argument falls into the broader category of manipulation arguments against compatibilism; however, nothing hinges on this categorization as far as the present reply to the argument is concerned. Another influential version of the manipulation arguments, which I cannot address in the present paper, is Pereboom's 'four-case argument', see Pereboom 1995 and 2001.

control. Consequently, the truth of physical determinism is incompatible with agents' freedom and responsibility.

My claim is that the conclusion of the argument does not follow even if we accept the truth of the premises. The example warrants only the inference that freedom and responsibility are impossible *in those deterministic universes whose physical laws and initial conditions at  $t_1$  are the same as, or at least sufficiently similar to, the universe in which Diana can control Ernie's future actions just by creating his zygote*. However, the possibility of this type of manipulation does not entail that in *every* deterministic universe, in which human agents develop from a zygote, anyone can have Diana's power. For agents' actions cannot be controlled in the way in which the zygote-case describes *only because* the universe is deterministic. Further conditions need to be satisfied as well and it is the occurrence of those other conditions, not physical determinism, which explains why agents may not be responsible in those deterministic universes.

To show why, let us bracket the zygote-scenario for a moment and suppose simply that Diana intends to bring about a state  $\omega$  at location  $d_2$  at  $t_3$  in a deterministic universe by interfering into the evolution of the events at location  $d_1$  at  $t_2$ . She can (has the power to) *guarantee*  $\omega$ 's occurrence in this way only if (a) the world is physically deterministic; (b) she knows the relevant physical laws and the state of the universe at  $t_2$ ; (c) she can bring about the necessary changes at  $t_2$  at  $d_1$ . However, all this is not sufficient. For whether or not  $\omega$  will eventually obtain depends not merely on what happens locally at  $t_2$  at  $d_1$  and the laws, but on the surrounding states of the universe as well. In fact, Diana can guarantee that  $\omega$  will occur only if the universe at  $t_1$ , that is before her intervention, is in some very specific state.

Diana can, of course, have the *further power* to modify either the initial state or the total state of the universe at  $t_2$  in order to guarantee that  $\omega$  is going to obtain a  $t_3$ . However, if we granted her that power, that would render our general case disanalogous to how an agent's zygotes and his behaviour are related in an ordinary deterministic world. We could restore the connection only if we assumed that every deterministic universe *in which zygotes can exist and develop into a human agent must by metaphysical necessity* be such a world in which every external condition necessary for the relevant type of manipulability also obtains. However, as I shall explain in the next section, this assumption is entirely ad hoc and implausible for many reasons.

What follows from this is that if we continue to feel that in the imaginary case Ernie is not responsible for doing A if his zygote was created in a deterministic universe, we do so, not because the universe in which he develops into a person from a zygote is deterministic, but *because of the specific conditions* which must obtain in order for Diana to have the power to manipulate Ernie's future behaviour. In the zygote-example Ernie's responsibility is not undermined because of the truth of physical determinism, but because of the peculiar initial conditions of the deterministic universe; or because of the peculiar circumstances at  $t_2$  which make Diana's intervention into Ernie's future possible.

#### 4 Determinism, Local Manipulability, and the Relevance of External Conditions

Physical determinism is a thesis about how the *complete physical states* of a universe entail each other in the following sense. If physical determinism is true, then a complete description of the physical state of a universe  $u$  at any instant  $t$  together with all relevant



physical laws entail all truths about the physical state of  $u$  at any other instant. More precisely, the total distribution of the values of (fundamental) physical magnitudes in a deterministic universe at a particular instant  $t$  is entailed by the total distribution of the values of those magnitudes at any other instant together with the (relevant, functional, dynamical) physical laws.

It is this entailment-relation on which the possibility of the zygote argument is built. Diana knows the complete physical state of the universe and hence, given physical determinism, she is imagined having the power to control what happens in the universe at  $t_3$  by intervening into its history at  $t_2$ . However, as we have seen above, for the zygote argument to succeed, we must introduce a much stronger assumption about the worlds in which zygotes can be formed or created. We have to assume that if physical determinism is true, then it *must* be possible that some *local intervention* into the state of the universe at  $t_2$  can ensure the occurrence of some specific *local events* at  $t_3$ . More precisely, we must imagine circumstances in which changing the physical constitution/structure of a zygote at time  $t_2$  *can be nomically sufficient* for the performance or avoidance of an intentional action at time  $t_3$ .

Since the possibility of (potentially responsibility-undermining) manipulation does not follow from the truth of determinism, we need to assume that the universe is in such peculiar initial conditions at  $t_1$  which can ground this possibility. Physical determinism in the sense relevant to the issue of compatibilism is global. Hence its truth does not entail that a local intervention into the universe's history at one particular spatial region  $R$  at  $t$  can be sufficient to produce an event at another spatiotemporally distinct – and relatively distant – region  $R^*$  *irrespective* of the general physical state of universe outside regions  $R$  and  $R^*$ .

Moreover, the relevant entailment relation among the instantaneous global states is perfectly symmetric: if, given the universe's physical laws, its physical state at  $t_n$  entails its state at  $t_m$ , then its state at  $t_m$  also entails its state at  $t_n$ . Thus, whatever makes controlling and manipulating the future possible, it cannot be *the mere fact* that a universe is physically deterministic (unless we are ready to admit that by controlling some present state of the universe we can manipulate some of its past states; or that manipulative control is a symmetric relation).

Nevertheless, it seems that we do have real cases when a local intervention into the history of universe appears to be sufficient for the occurrence of a later event.<sup>10</sup> Consider, for instance, a bomb which is designed to explode in a specific future instant at a specific place. It seems that we do not need to assume anything about the initial conditions of the universe in order to imagine that someone can create and plant a timebomb to bring about some specific future events.

However, first, it is important to observe that whether or not someone can create and use efficiently a timebomb does not seem to depend on whether or not she lives in a deterministic universe. Whatever is needed for the efficient operation of a timebomb, a deterministic interpretation of quantum mechanics does not seem to be required.<sup>11</sup> If we

<sup>10</sup> I thank a reviewer of this paper for pressing me on this point.

<sup>11</sup> This does not mean that microscopic indeterminism is always irrelevant to macroscopic processes. But for microscopic indeterminacy to become relevant for macroscopic phenomena, we need to imagine special circumstances in which the effects of microscopic indeterminacy are amplified. A well-known situation of that type is Feynman's case when a bomb is connected with a Geiger-counter and goes off only if the Geiger-counter registers certain readings (see Feynman 1965). But the point is, again, that for such amplification to be possible without human design, the universe must be in a very specific and unlikely initial state.

imagined that Diana-type manipulations are possible because there is an analogy between the possibility of creating and planting timebombs on the one hand and the creation and implantation of a zygote in someone's womb on the other, then it would follow that the possibility of manipulation has nothing to do with the truth or falsity of physical determinism. And then it should be clear enough that what best explains Ernie's non-responsibility is that he has been used by another agent as a means to bring about certain events in the future, and not that the world in which this happens is deterministic.

But second, and more importantly, there is no analogy between the two cases. For what Diana aims to determine is not the 'operation' of the zygote that she creates, but *some specific action* which is a *faraway causal consequence* of the zygote's having a certain physical structure at the time of its creation, which happens in the distant future. Hence, a more adequate analogue between the two cases should assume that just by creating and planting a timebomb at  $t_1$ , one could guarantee that a specific event at a location which is different from the location of the explosion will happen at  $t_3$  as a consequence of the bomb's exploding at  $t_2$ . Diana does not create and implant the zygote with the purpose of the zygote's doing something in the future, but in order to secure, just by creating and implanting the zygote, that *the agent* who develops from it do something. The point is that no matter how much knowledge she has about the laws and the state of the universe at the time of creation, she can have the power to guarantee that a specific future event will happen (or a specific action will be performed by Ernie) only if many other conditions in the universe are favourable.

However, there is no reason whatsoever to believe that zygotes must *always* come to exist in a deterministic universe in such circumstances that make it possible that someone, just by creating zygotes, can ensure how agents developed from them will act later. What the zygote argument must assume is that agents and their behaviour are somehow *the distant causal consequences* of the original creation and implantations of zygotes. They are not the manifestations of *the zygote's intrinsic powers* the exercise of which is relatively isolated from what happens in its environment as is the case with the timebomb; or with any other physical systems in the universe the local state of which at  $t_2$  can guarantee its state at  $t_3$ .

In the zygote-example it is assumed that Diana has the power, just by combining a few atoms at  $t_2$  in region  $R$ , to ensure that Ernie will do A thirty years later at  $t_3$  in region  $R^*$ . But the agent's behaviour and the constitution of his zygote cannot reasonably be considered as parts of an isolated system. Between  $t_2$  to  $t_3$  the zygote, Ernie's embryo and his ever-changing human body are in constant physical interaction in countless ways with their physical environment; so much so that Ernie at  $t_3$  – remember: he is supposed to be a normal human being – will not share a single atom with the zygote Z from which he has developed.

Thus, in order for Diana to be able to control at  $t_2$  what Ernie does thirty years later at  $t_3$ , *the deterministic universe* must be in such a state at  $t_1$  – that is, any time earlier than  $t_2$  – which makes it possible that a *small, localized intervention* that happens in  $R$  – combining a few atoms to create a zygote – be *sufficient* for Ernie's doing A thirty years later in  $R^*$ . Among other things, the state of the universe must be such that none of the future physical interactions between the zygote, the embryo and Ernie's body with their physical environment during the thirty-year interval from  $t_2$  to  $t_3$  *prevent* any possible causal influence of the zygote's atomic structure to what exactly happens in  $R^*$  at  $t_3$ .

To illustrate how strong this assumption is let me use another widely known manipulation-scenario, the so-called Frankfurt-style case. In Harry Frankfurt's original example, a neuroscientist N, who is supposed to be able to interfere into the decision of another agent S, manipulate S's action by implementing the following policy: (a) if S is about to decide so as N wishes, N lets S make 'her own choice'; but (b) if S is about to decide in any other way, N intervenes and makes S to decide as he wants.<sup>12</sup>

Frankfurt introduces this story in order to prove that responsibility does not require the ability to do otherwise. Whether or not he is right about this, or whether or not there is any modification of the original example which can prove this, is a matter of debate. But this is not my present concern.<sup>13</sup> Rather, I would like to call attention to *the extreme fragility* of the *external* conditions which render even this rather direct form of manipulation possible.

In Frankfurt's original example – and in fact in all other modifications of the original example with which I am familiar – S's decision is constrained by the presence of N, but N's influence on S is unconstrained. But if the world is such that S can be prevented from making certain decisions by N's interference, the same must be true about N as well. There can be another manipulator N\* present, who, whenever N is about to interfere into S's decision, somehow makes that intervention impossible. Thus, even if there is a sense in which N has the capacity to interfere into S's decisions – he would be able to manipulate S's choices if N\* were absent – he can never exercise that capacity, because N\*s is in fact present. In this situation, even if N has the power to determine S's choices, S always chooses as she wants. The preventive intervention of N\* 'shelters' S from any interference by N in the sense that N's attempts to interfere will never be successful.

Importantly, both in the original and in the more complex modified scenario, it is entirely irrelevant whether N and N\* are conscious agents or purely physical factors or processes. The point of the more complex Frankfurt-type scenario is that the possibility of successful manipulation does not only depend on the manipulator's power, but also *on the complex causal structure of the world* in which the power is supposed to be exercised on a specific occasion. And there is no reason whatsoever to suppose that in every deterministic universe, in which humans develop from a zygote, the local causal structure that might enable someone to manipulate another agent's behaviour just by creating his zygote thirty years earlier, must always obtain.

Crucially, Frankfurt-style cases are supposed to be possible in the circumstances of physical determinism. Otherwise, the example could not show anything about the compatibility of physical determinism and agents' responsibility, which was its original ambition. Frankfurt's-type manipulations must then be possible *even if* determinism is true. But they are certainly not possible *because* determinism is true. They are possible because of the occurrence of a peculiar causal structure, which may or may not obtain in certain regions of the universe. It is not enough to imagine that the manipulator has the power or

<sup>12</sup> See Frankfurt 1969.

<sup>13</sup> My considered view is that he is not; see Huoranszki 2011 and 2017. Frankfurt-type cases are indeed interesting, but not because they can prove the compatibility of responsibility with agents' inability to do otherwise.

capacity to interfere into S's decision if he decides so. It is also necessary to assume that he can exercise that capacity successfully. But there are many circumstances in which he cannot.

Consequently, it does not follow from the truth of determinism that the constitution of zygotes *can* be, and hence *are*, the 'deterministic causes' of any specific, local action that an agent performs in the future. There are countless background factors that can easily prevent the original structure of the zygote to play a causal role in the agent's action thirty years later.<sup>14</sup>

For there are countless such physical conditions the presence or absence of which can prevent the zygote's structure playing any significant causal role in how an agent makes choices and hence how his actions are produced. It is one thing to admit the obvious that the creation (or the coming to existence) of a zygote is causally *necessary* for the existence of a 'normal' human agent or that it can influence some of his behavioural dispositions. But it is another to assume that the creation and implantation of a zygote can be causally *sufficient* for guaranteeing the performance of a specific future action in the sense that altering the zygote's structure would necessarily result in altering the ways in which an agent will choose and act thirty years later.

An advocate of the zygote-argument should prove that, during the entire interval from the creation of a zygote to the adult agent's action, it is metaphysically, or at least nomologically, impossible that some such condition occurs which is incompatible with the possibility of Diana-type manipulation. But something that is extremely unlikely cannot be proved to be necessary. Even in a deterministic world countless contingent external conditions need to be satisfied which must help rendering Diana-type manipulation possible. There is no reason to assume that those conditions are satisfied in every world in which human zygotes can exist just because those worlds are deterministic.

It might be helpful to restate at this point my principal claim about the Zygote Argument. I do *not* deny that the zygote scenario is conceivable. My point is that its conceivability cannot prove that physical *determinism* entails the *possibility* of manipulating an agent's future actions by designing and creating his zygote; not to mention the even stronger claim – necessary for drawing the argument's desired conclusion – that physical determinism makes it *necessary* that the formation of *every* actual zygote can *causally contribute* to the performance of a future intentional action in a way that can weaken our belief in the agent's responsibility. The possibility of such contribution, or in fact the causal relevance of any specific past event to our present actions, depends on the local causal structure of the universe, and not on the truth of determinism. Thus, what can undermine an agent's responsibility in the zygote-case are only the occurrences of those peculiar and contingent conditions, and not the truth of physical determinism.

<sup>14</sup> It might be objected that agents' genes are relatively isolated systems. However, first, genetic information stored in DNA must be translated in order to result in phenotypic traits; and second, as behaviour genetics have shown, even the genetic structure of an agent's cell can change as a response to interactions with its physical environment. About some important issues concerning the relation between inheritance, genes and behaviour see the excellent survey of Chamey 2012 as well as the following discussion in the relevant issue of BBS.

## 5 Indeterminism and Manipulation

In this section I am going to argue that there need not be any difference between a deterministic and indeterministic universe as far as the possibility of zygote-case scenarios is concerned. If that is right, the truth or falsity of physical determinism does not make any difference as far as our beliefs about responsibility are concerned. What can undermine the belief in agents' responsibility is always the tacit or explicit assumption about the presence of some specific causal structure that renders some responsibility-undermining manipulation possible, and not physical determinism.

To appreciate the significance of this point, consider the following example against the compatibility of responsibility with the truth of physical *indeterminism*. As before, Diana creates Ernie's zygote at  $t_2$  with the purpose to ensure, as much as it is possible, that Ernie will perform action A at  $t_3$ . Given indeterminism, Diana does not have full or perfect control over the physical future; and consequently, over whether Ernie will do A at  $t_3$ . Nonetheless, she may have at least some degree of control over Ernie's future actions.

In an indeterministic world, Diana can have *indeterministic control*. Suppose she has indeterministic control in the following sense: if she creates a zygote with a certain physical configuration  $\varphi$  at  $t_2$ , then either (a) with a relatively high probability  $1-p$ , Ernie will do A at  $t_3$  and all NPCR are satisfied; or (b) with a relatively low probability  $p$ , Ernie will fail to do A at  $t_3$ , but some NPCR are not satisfied. (For instance, Ernie fails to do A because he has become a compulsive A-avoider).

In this scenario, irrespective of whether or not Ernie does A, he cannot be responsible for his action. If he does A, he is not responsible since his action has been successfully manipulated by Diana exactly in the same responsibility undermining way as it is manipulated in the deterministic case. And if he avoids doing A, he is not responsible because he does so as a result of becoming psychologically unable to properly self-control his own behaviour.

Suppose, finally, that there is an indeterministic universe that differs from the one in which Diana creates Ernie's zygote only to the extent that, in this universe, Ernie's zygote comes to exist in the natural way.<sup>15</sup> Thus, the constitution of his zygote makes it *extremely likely* that he will perform a certain type of action A thirty years later in some specific situation, but *impossible* that he will avoid doing A in circumstances in which all NPCR are satisfied.<sup>16</sup>

Would this example convince anyone that physical indeterminism is incompatible with responsibility? Obviously not. The natural response to it is that even if there are *some conceivable indeterministic universes* in which agents' responsibility is undermined because of the rather peculiar physical conditions described, this goes in no way to showing that *physical indeterminism* can undermine the responsibility of

<sup>15</sup> Observe that Ernie's zygote is not created deterministically even in the original, deterministic scenario. Its constitution depends on what Diana wants, not on the prior state of the universe. Hence, whether or not the zygote itself is created deterministically is irrelevant to the argument.

<sup>16</sup> One might object that making something extremely likely is not the same as guaranteeing that it is going to happen. Indeed. But as far as responsibility is concerned, I cannot see any difference between (a) an action that is done as a result of successful manipulation which guarantees its performance and (b) an action which is done as a result of successful manipulation which makes the performance of that action extremely likely. What matters is only the success of the manipulation and the manipulated agent's inability to prevent its success.

agents who develop from a zygote. All that the possibility of such universes shows is that in *some* indeterministic universes with *some* peculiar structure agents' freedom and responsibility can be systematically undermined.

However, exactly the same holds in the case of deterministic universes. It is possible that in *some* deterministic universes with some very specific causal structure someone with Diana's power can control what someone else does intentionally in the future just by manipulating the zygote from which an agent develops; and that, *for this reason*, agents' responsibility would be undermined even if their zygote had developed in the 'natural way'. But this does not show anything about the significance of physical determinism for the issue of responsibility.

What we should prove is that such kind of manipulation is possible in *every* deterministic universe in which humans develop from a zygote *because* the world is deterministic. Otherwise, the zygote-argument cannot prove anything about the relevance of determinism to our beliefs concerning agents' responsibility. I guess everyone would agree that the mere facts that (a) humans develop from a zygote and that (b) persons cannot control the structure of the zygote from which they develop are insufficient for denying the possibility of responsibility. What the zygote argument should prove is that we have reason to deny the possibility of responsibility if we add to these facts that the zygote in question develops into a human agent in a physically deterministic universe.<sup>17</sup>

However, as we have seen, imagining that a zygote develops into a human being in a deterministic universe is neither sufficient nor necessary for explaining why Ernie in the zygote-scenario is not responsible. For the truth of physical determinism itself (together with the two other assumptions) does not entail the possibility of the sort of manipulation that may undermine responsibility. Thus, the zygote argument can show no more about the incompatibility of responsibility and determinism as it does about the incompatibility of responsibility and indeterminism.

We can imagine *some* physically deterministic universes with such initial conditions that make the relevant type of manipulation scenario at least *sometimes* physically possible. It requires more effort to imagine a universe in which the physical conditions necessary for the possibility of this kind of manipulation *always* obtain on every occasion when a human zygote comes to exist. But it is a matter of logic and not of imagination that there is nothing in the *concept* of physical determinism which entails that those conditions *must* obtain in every deterministic world in which zygotes exist. In general, 'original design' of the kind which can intuitively undermine an agent's freedom and responsibility is physically possible only if countless contingent physical background conditions are satisfied in a deterministic universe. But then, it is those entirely contingent background conditions, and not determinism, which best explain why we think that agents in those worlds are not responsible.

As we have seen, the sort of intervention required by the zygote case is physically possible only if nothing between  $t_2$  (the time of the creation of the zygote) and  $t_3$  (the performance of the action) could disrupt the potential physical influence of the original molecular constitution of the zygote to the agent's future specific action. Otherwise, Diana simply cannot control Ernie's future action just by creating the zygote from which he develops. But it is hardly entailed by physical determinism that these further

<sup>17</sup> I thank a reviewer for urging me to clarify this point.

conditions always occur. In fact, the probability that they occur is extremely low. Hence the mere possibility of the zygote-case cannot prove the incompatibility of responsibility with physical determinism for the same reason as the conceivability of some physically indeterministic universe in which agents' responsibility is systematically undermined cannot prove the incompatibility of responsibility with physical indeterminism.

Thus, the zygote argument cannot support incompatibilism. Strangely enough, it can only support the truth of compatibilism, if anything. It might be true that in all those universes in which the peculiar physical conditions that render Diana's control over Ernie's future actions possible are satisfied human agents, who develop from zygotes which come to exist 'in a normal way', are not responsible for their actions. However, if, given the initial state of a deterministic universe, it is physically impossible that these conditions are ever satisfied, then agents *cannot be exposed to the threat* of certain sorts of manipulation. Hence at least one worry about the consequences of physical determinism to agents' freedom and responsibility turns out to be unjustified.

## 6 Rescuing the Zygote Argument?

Some may object that all I have shown is only that zygote-cases are unlikely; but this is irrelevant if they are conceivable without contradiction. For the issue of compatibility of determinism with freedom and responsibility is modal, not statistical.

But if this is meant to be a criticism against my argument, it misses my point completely. For my claim is not statistical, but modal. I say that, since physical determinism does not entail *the possibility* of the sort of control without which the zygote argument cannot even get off the ground, physical determinism cannot undermine agents' responsibility if all NPCR are satisfied. If there are *some* such deterministic universes in which the relevant kind of control is physically impossible – and for all we know our universe is one of them – then it is only the occurrence of certain contingent specific conditions that could undermine agent's responsibility, and not the truth physical determinism.

One may object, further, that all these complications about the conditions in which the zygote-manipulation case is possible are irrelevant. If determinism is true, then what Ernie does at  $t_3$  depends on the earlier states of the physical universe over which he cannot have any conscious control, and this is incompatible with his responsibility. The zygote argument is just an illustration of how 'causal determinism' can undermine agents' freedom and responsibility. Since the purpose of the zygote story is only to 'help our intuitions', no details about how we can move from the example to the conclusion can be relevant to the persuasiveness of the argument.<sup>18</sup>

But first, whatever the zygote argument is supposed to be, it is certainly meant to be *an argument* for something; an argument, which relies on certain premises in order to draw certain conclusions. It is not meant to be a 'mere illustration'. And second, it cannot be an argument about the alleged responsibility undermining effects of *causal* determinism, because the zygote argument does not specify any particular causal process or mechanism the operation of which can ensure that the (artificial or natural)

<sup>18</sup> See for instance Todd 2013.

creation of a zygote will *cause* Ernie's actions thirty years later. In fact, it is an *essential aspect of this argument* that the example does not rely on any specific causal mechanism. And this is the argument's virtue, not its vice.

As Alfred Mele points out (first in Mele 2005), manipulation arguments<sup>19</sup> which rely on examples that involve some specific causal mechanism suffer from a general weakness. Their weakness is that what best explains (our intuition about) the non-responsibility of the manipulated agent is the operation of a manipulative causal mechanism, and *not* determinism as such. For if an agent's actions were the products of an *indeterministic* manipulative mechanism, his responsibility would be similarly undermined. The zygote argument is supposed to be superior to other versions of the manipulation argument precisely because it does not assume the operation of any specific causal mechanism which can explain why the manipulated agent is not held responsible. All we need is just the truth of physical determinism in universes in which ordinary human agents develop from a zygote.

What I have shown is that the 'all we need to add is just the truth of physical determinism' assumption is false. For the zygote-scenario is possible only if many contingent physical conditions are satisfied in a universe in which human agents develop from a zygote. Whether or not these conditions are satisfied is entirely independent of the truth of physical determinism, which is a view about laws and the nomic structure of the universe; and which is not to be confused with assumptions about local causal structures determined by the universe's initial conditions and its history. The physical configuration of a zygote can of course contribute to the development of certain *generic* phenotypical (or perhaps even psychological/behavioural) *traits* of a human being, no matter whether physical determinism is true or false. But the molecular structure of a zygote can ensure any particular future action of an agent only in very special physical circumstances. Hence the kind of intervention described by the zygote example is possible only in some deterministic universes with peculiar initial conditions. But then it is those peculiar initial conditions, and not the mere truth of physical determinism, which best explain why human agents are not responsible *there*.

Consequently, the zygote argument is susceptible to an objection similar to the one Mele raises against alternative formulations of the manipulation argument. What can best explain (our intuition) that Ernie in the example is not responsible is not physical determinism, but the peculiar physical circumstances in which an agent's zygote was created and which need not obtain just because physical determinism is true; and further, which can obtain even if physical indeterminism is true. In most alternative formulations of the manipulation argument, the peculiar circumstances are local. In the case of the Zygote Argument, the peculiar circumstances are global. But in both cases, what can, if anything, undermine an agent's responsibility is the occurrence of the peculiar circumstances which lead to his action, and not the truth of physical determinism.<sup>20</sup>

<sup>19</sup> Like Pereboom 1995 and 2001.

<sup>20</sup> I am grateful to a reviewer for comments on an earlier version of this paper.



**Funding** Open Access funding provided by Central European University.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Arrow, K. J. (1963). *Social choice and individual values* (2nd ed.). New York: John Wiley.
- Chamey, E. (2012). Behavior genetics and postgenomics. *Behavioral and Brain Sciences*, 35, 33–410.
- Feynman, R. (1965). *The character of physical Laws*. London: Modern Library.
- Fischer, J. M. (2011). The zygote argument remixed. *Analysis*, 71, 267–272.
- Frankfurt, H. (1969). Alternate possibilities and moral responsibility. *Journal of Philosophy*, 66, 829–839.
- Ginet, C. (1990). *On action*. Cambridge: Cambridge University Press.
- Huoranszki, F. (2011). *Freedom of the will. A conditional analysis*. New York: Routledge.
- Huoranszki, F. (2017). Alternative possibilities and causal overdetermination. *Disputatio*, 9, 193–217.
- Kane, R. (1996). *The significance of free will*. Oxford: Oxford University Press.
- Kearns, S. (2012). Aborting the zygote argument. *Philosophical Studies*, 160, 379–389.
- Mele, A. (1995). *Autonomous Agents*. New York: Oxford University Press.
- Mele, A. (2005). A critique of Pereboom's 'four-case argument' for incompatibilism. *Analysis*, 65(1), 75–80.
- Mele, A. (2008). Manipulation, compatibilism, and moral responsibility. *The Journal of Ethics*, 12, 263–286.
- Pereboom, D. (1995). Determinism *al Dente*. *Nous*, 29, 21–45.
- Pereboom, D. (2001). *Living without free will*. Cambridge UK: Cambridge University Press.
- Schlosser, M. E. (2015). Manipulation and the zygote argument: Another reply. *The Journal of Ethics*, 19, 73–84.
- Todd, P. (2013). Defending (a modified version of) the zygote argument. *Philosophical Studies*, 164, 189–203.
- Van Inwagen, P. (1983). *An essay on free will*. Oxford: Clarendon Press.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.