Synthese (2022) 200:254
https://doi.org/10.1007/s11229-022-03523-2

ORIGINAL RESEARCH

Anscombe and Intentional Agency Incompatibilism (for human *and* animal agents)

Erasmus Mayr[1]

Received: 23 March 2021 / Accepted: 23 November 2021

Abstract
In "Causality and Determination", Anscombe stressed that, in her view, physical determinism and free action were incompatible. As the relevant passage suggests, her espousal of incompatibilism was not merely due to specific features of human 'ethical' freedom, but (also) due to general features of agency, intentionality, and voluntariness. For Anscombe went on to tentatively suggest that lack of physical determination was required for the intentional conduct of animals we would not classify as ‚free', too. In this paper, I examine three different lines of argument to establish Anscombe's latter suggestion, which are based on general considerations about the causal efficacy of psychological-agential phenomena, the nature of agency, and the specific features of intentional agency. I start with Anscombe's own claim from "The Causation of Action" that microphysical determinism would make psychological and personal phenomena epiphenomenal, before I turn to the view of 'Agency Incompatibilism', that genuine agency requires the absence of antecedent necessitation, and, lastly, to concerns about some crucial features of intentional agency which we find in both human and animal agents.

Keywords Elizabeth Anscombe · Mechanism and microphysical determinism · Agency Incompatibilism · Helen Steward · Active Powers and antecedent determination · Practical deliberation and determinism

In "Causality and Determination", Elizabeth Anscombe not only forcefully argued that causation and determination do not always go together. She also stressed that physical indeterminism was, in her view, a necessary precondition for free will and rejected attempts to reconcile such determinism with the possibility of 'ethi-

✉ Erasmus Mayr
erasmusmayr@yahoo.com

[1] Institute for Philosophy, Friedrich-Alexander-Universität Erlangen-Nürnberg, Bismarckstr. 1, D-91054 Erlangen, Germany

Published online: 03 June 2022

cal' freedom. Interestingly, the relevant passage suggests that this rejection was due not merely to specific features of 'ethical' freedom, or considerations about moral responsibility, but (also) due to general features of agency, intentionality, and voluntariness, which we find not only in humans, but also in other animals. For Anscombe went on to tentatively suggest that lack of physical determination is required for the intentional conduct of animals we would not classify as ‚free', too.

"The truth of physical indeterminism is […] indispensable if we are to make anything of the claim to freedom. But certainly it is insufficient. The physically undetermined is not thereby 'free'. […] Nevertheless, there is nothing unacceptable about the idea that that 'physical haphazard' should be the only physical correlate of human freedom of action; and perhaps also of the voluntariness and intentionalness in the conduct of other animals which we do not call 'free'. " (1971, 146).

Anscombe's tentative farther claim has close affinities to a position that has become prominent (again) in recent years, especially through the work of Helen Steward, under the label of 'Agency Incompatibilism': I.e. the view that agency as such – be it human or animal agency – would be impossible in a deterministic world. While Anscombe, in the quoted passage, speaks only of 'voluntary' and 'intentional' conduct, rather than of 'agency' *tout court*, which makes her claim at least *prima facie* narrower than Steward's, it is at least clear that the latter would imply Anscombe's claim of 'Intentional Agency Incompatibilism', too.

But should we accept an 'Intentional Agency Incompatibilism' of the form Anscombe suggests? Why should we think that human or animal behaviour, insofar as it can count as an instance of intentional[1] agency, requires physical indeterminism? In this paper, I want to critically assess three lines of argument that purport to establish that physical determinism, at least in certain forms, would rule out the possibility of human or animal intentional agency. In Anscombe's own writings, we can find two different lines of thought supporting this conclusion. While she did not offer an argument for her suggestion in "Causality and Determination", she presented what is presumably her central argument for the incompatibility claim in "The Causation of Action", where she tried to show that microphysical determinism implies mechanism. In other writings she expressed some sympathy for the idea that determinism would undercut the point of deliberation and choice, which suggests that, even independently from the mechanism argument, she considered physical determinism to be at odds with some of the specific characteristics of *intentional* agency. In addition to these arguments drawn (to a greater or lesser extent) from Anscombe's own work, it is natural, in assessing her claim, to also look at the main arguments in favour of 'Agency Incompatibilism', given both the current prominence of that position and its affinity to Anscombe's claim.

The order in which I will discuss these lines of argument will be systematic, rather than historical, namely one of successively narrowing scope: I will begin with Anscombe's argument that microphysical determinism would imply mechanism and epiphenomenalism about psychological phenomena *in general*, which result would (also) make intentional agency illusory (Sect. 1). I will then look at arguments that focus on *agency*, rather than the causal relevance of psychological phenomena in

[1] I will only focus on intentionalness in the following, setting aside voluntariness.

general, and that try to establish that agency as such is impossible under determinism (Sect. 2). Lastly, narrowing the focus even more, I will look at arguments focussing on the specific characteristics of *intentional* agency (Sect. 3). With regard to all three lines of argument, my conclusions will be sceptical, however: Neither the causal relevance of psychological phenomena, nor agency, nor the intentional character of actions, require physical (or microphysical) indeterminism, at least when we look beyond the realm of human agents.

Before I set off, an advance warning is in order: While Anscombe, in the passage quoted above, speaks specifically of "the conduct of other animals which we do not call 'free'", her claim, as I understand it, is not merely one about animal intentional agency. If she is right, key features of intentional agency as such, as we find it in both humans *and* other animals, are incompatible with physical determinism, or certain forms of it. In consequence, the differences between human and non-human animal agency will not be relevant to all of the arguments that are discussed in the following, and will only become crucial in the last section.

1 Microphysical determinism and mechanism

In "The Causation of Action", Anscombe argues that (micro)physical determinism would not only make human actions predetermined, but would also imply 'mechanism'. Since mechanism might make intentional agency illusory, this seems one promising route towards establishing the incompatibility of (micro)physical determinism and intentional agency.

1.1 The threat of mechanism

The background to the thesis of mechanism is the idea that we can distinguish between lower and higher levels of (descriptions of) reality, such that the particles and phenomena at the lower levels 'make up' or 'compose' those at the higher levels. According to the thesis of mechanism, there is some level of physical description, referring to physical particles and phenomena which compose all the higher-level entities and phenomena, such that only those physical particles and phenomena (and yet lower level ones) can be considered as causally efficacious and as having any real causal forces, while the higher-level phenomena cannot. For pragmatic reasons, it might still be "a highly convenient, nay indispensable, *facon de parler*" (1983, 103) to speak of the higher-level phenomena (such as mental phenomena or animal actions) as 'causing' other phenomena. But it is no more than that: a "*facon de parler*", and the higher-level descriptions will be, as Anscombe puts it, just "supervenient descriptions"[2] (loc.cit.). All the real causal work is done at the lower level; the higher-level entities and phenomena are no more than 'shadows' of what goes on there and are causally epiphenomenal. Anscombe claims that such mechanism would be the necessary consequence of determinism at the fundamental level – or rather,

[2] Anscombe uses ‚supervenient' differently from how it is normally used in contemporary debates about downward causation, but I will follow her usage here.

given that contemporary quantum physics tends to view the 'most' fundamental level as indeterministic, of determinism at one relatively fundamental level with respect to all comparatively higher levels of descriptions:

"given a strict and total determinism relative to the particles of which all things are composed, I think two things follow: one, that these descriptions of action and their causation in human (and animal) histories *are* supervenient descriptions, and two, that actions *are* determined" (1983, 104).

Though talk of 'particles' may suggest determinism at a fairly 'deep' microphysical level here, it is irrelevant where, within the 'hierarchy' of levels of descriptions, the deterministic level is placed. Once there is determinism at *some* level, all descriptions at higher levels which are 'based' on that deterministic level must be supervenient (loc.cit.). Thus, once there is determinism at some lower level on which the higher levels of mental and agential phenomena are 'based' (e.g. the level of neuronal networks), this would make mental and agential phenomena causally epiphenomenal.

Even if we assumed that Anscombe was right that (micro)physical determinism implies mechanism, it is not immediately clear that such determinism would thereby threaten our understanding of actions as intentional or undermine our explanations of intentional actions in terms of the intentions with which the agents acted. For, as Anscombe insists, the latter explanations are not themselves causal explanations (at least in the 'modern' sense) (1983, 106 f.). So why should the fact that higher-level events turn out to have no causal efficacy affect the standing of explanations by intention? It is true that Anscombe believes that there are other explanations of actions which *are* causal. This class includes explanations in terms of 'mental causes' (e.g. when seeing a frightening face at a window makes a child jump back (1957, 16)). Even when a settled, standing intention I have to (generally) do X makes me directly do X, without further deliberation about the particular case, the existence of my intention, so Anscombe, causes my action (1983, 95). If mechanism were true, neither of these purportedly causal explanations could pick out a factor which was really causally relevant. But why would explanations by intentions be necessarily undermined by this, given that they rely on a different explanatory pattern?

Still, we cannot suppose that mechanism would leave explanations by intentions untouched. When we act intentionally, Anscombe argues in *Intention*, our "practical knowledge is 'the cause of what it understands'", adding that this "means more than that this knowledge is observed to be a necessary condition of the production of various results" (1957, 87). While the reading of the passage is disputed, I take her to (also) make the claim that without the agent's having the practical knowledge in question she wouldn't perform the movements she does perform in the course of her intentional action.[3] It is very hard to see how practical knowledge could be genuinely necessary for this performance if mechanism were true and all the causal work was done at the (micro)physical level.

[3] I do not take her to merely state that without the practical knowledge those movements would not be (part of) an intentional action. In this, I concur with Schwenkler (2019, 173 ff.), who argues that Anscombe does not qualify practical knowledge *merely* as a formal cause, but also ascribes to it the role of an efficient cause.

In the passage I have just quoted from Anscombe is only talking about *human* intentional action: So what about the intentional actions of non-human animals? Whether such animals, too, in acting intentionally, have practical knowledge of what they are doing is a disputed question.[4] Fortunately, we don't have to decide this question here. For what Anscombe says about animal intentional behaviour makes it clear that she thinks that such behaviour is at least connected to the animal's having thoughts, perceptions and knowledge about the circumstances of the action (see 1957, 86). It is most natural to read her as claiming that our understanding of the animal's behaviour as intentional is tied to the possibility of explaining the animal's movements by referring to its knowledge and perception of those circumstances. We see the cat as (intentionally) crouching in order to catch the mouse because we can explain her crouching by saying things like 'The cat just saw the mouse moving over there'. Mechanism would undermine the possibility of such explanations, because it would imply that the animal's perceptions are not really relevant to what went on.

1.2 The thesis of micro-determinism

So, if Anscombe is right that (micro)physical determinism would imply mechanism, such determinism would indeed undermine the possibility of intentional agency in both humans and other animals. But would (micro)physical determinism really imply mechanism? To answer that question we first need a more precise formulation of the thesis of micro-determinism, or of "strict and total determinism relative to the particles of which all things are composed" (1983, 104). For the sake of simplicity, I will not worry about indeterminism at the (putatively) most fundamental level of physics, but will just focus on some fairly 'low' physical level which is meant to be governed by deterministic laws; when I talk of 'micro-determinism', this is meant to be understood as determinism for that level.

The thesis of determinism is often formulated as the general claim that for every event in the universe there are prior circumstances which, together with the laws of nature, necessitate its occurrence (including all its features),[5] or, as the 'global claim' that the initial global state of the world together with the laws of nature completely 'fixed' how the world would evolve at each future step.[6] A parallel formulation of the claim of microdeterminism would be, e.g., that for every microphysical event there are antecedent microphysical circumstances, which, together with the laws of nature, necessitated its occurrence and all its features. But since our specific question here is whether the causal efficacy of higher-level phenomena is compatible with microphysical determinism, such general formulations of micro-determinism are not the most helpful ones for our purposes. The reason for this is the same as one often noted in the debate about determinism and free will: What threatens free will, from the incompatibilists' perspective, is specifically determinism "in the wrong places"[7], i.e. within the processes suitably connected to human choice and agency. It is for this rea-

[4] Gustafsson (2016, 234) denies that they have.

[5] See, e.g., Steward (2012, 9).

[6] Cf., e.g. Holton (2013, 87).

[7] Steward (2012, 9).

son that the falsity of universal determinism will not *per se* guarantee for incompatibilists that free will is possible, because universal determinism might be falsified by the existence of some undetermined events which are completely unrelated to human choice and agency. *Pari passu*, if Anscombe's point about mechanism is correct, the threat of mechanism will not necessarily be removed by the falsity of determinism at the relevant micro-level, since such micro-determinism, too, could be falsified by the existence of indeterministic micro-events completely unrelated to higher-level phenomena. The threat of mechanism would remain as long as all microphysical states and events relevantly connected to higher-level phenomena were necessitated by antecedent microphysical events.

I will therefore use the following slightly restricted version of the claim of microphysical determinism, which is already tailored to the issue of mechanism:

Micro-determinism^{A}: All microphysical states and events involved in, or (contributing to) constituting, higher-level phenomena are such that for these states and events there are antecedent microphysical states and events which (together with the laws of nature) necessitated their obtaining viz. occurrence.

Two things should be noted about this formulation. First, I will not put weight on the distinction between states and events, because for our particular question not much hinges on it. Second, the claim of micro-determinism, as I understand it, only concerns the determination of *microphysical* phenomena. It does not, by itself, state that higher-level phenomena are determined, too; the latter claim is only meant to be an implication of micro-determinism *plus* the claim that the microphysical phenomena 'compose' the higher-level ones.

1.3 Micro-determinism without mechanism

At first glance, there is much to be said for thinking that Micro-determinism^{A} would make higher-level phenomena epiphenomenal. If all microphysical states relevantly involved in the higher-level phenomena are already fully determined to obtain by other microphysical states, what causal work is there 'left over' for the higher-level phenomena to do? None, it seems, with regard to subsequent microphysical states. And equally little with regard to subsequent higher-level states, since the latters' occurrence will already be made necessary by the obtaining of the microphysical events which 'compose' them.[8] *Some* determination gaps in the microphysical causal processes seem therefore needed if higher-level phenomena are to do any causal work.[9]

But this appearance is deceptive: Micro-determinism^{A} does not necessarily deprive higher-level phenomena of their causal relevance. To illustrate this, I want to use a scenario suggested by David Humphreys. While developed against Jaegwon Kim's argument for epiphenomalism about higher-level property instances, this scenario also shows that Micro-determinism^{A} allows for the existence of higher-level property instances with genuine causal relevance.

[8] Compare Kim's arguments against downward causation and for epiphenomenalism about irreducible mental properties, e.g. (1993).

[9] For a similar argument see Steward (2012, 244).

As Humphreys argues, one way in which novel higher-level property instances can emerge is through 'fusion' of lower-level property instances. In a fusion-operation, the lower-level property instances are combined in such a way that, after this operation, they "no longer have an independent existence within the fusion". They "no longer exist as separate entities and they do not have all of their i-level causal powers available for use at the (i + l)st level. Some of them, so to speak, have been 'used up' in forming the fused property instance" (1997, 10).

Humphreys uses quantum entanglement states as a possible illustration of this phenomenon (1997, 15 f.). But whatever you may think about this particular case, what Humphreys describes seems clearly a possible scenario. Furthermore, in this scenario it is plausible to ascribe genuine causal efficacy to the new higher-level property instance which has come about by the fusion (call it the 'fusion-state'). The fusion-state is a much better candidate than its components for being the (or a) cause with regard to further changes which come about in consequence of the fusion. For at the time the fusion-state exists, its components do not exist as separate entities which could do the causal work in its place and would 'compete' with it for causal influence. The alternative to taking the fusion-state to be causally efficacious is to assume that its components caused the further effects at a temporal distance, being causally efficacious with regard to the further effects already *before* they fused, thus 'by-passing' the fusion-state. But why should we say that – and not rather say that the components cause the effects *via* the fusion-state which they enter into? A slight addition to the case will make the former alternative even more implausible. Imagine that the same higher-level property instance, followed by the same consequences, can come into existence, on different occasions, by the fusion of somewhat different kinds of micro-property instances.[10] Then the presence of the fusion-state is clearly more relevant than the earlier presence of the specific microproperty instances for explaining these consequences.[11] So, there is no good reason to regard the fusion-state as generally epiphenomenal.

Importantly, Humphreys' fusion cases are compatible both with Microdeterminism^{A} and the idea that the higher level phenomena are 'made up' from microphysical ones. As regards Microdeterminism^{A}, in fusion cases all microphysical effects caused by the fusion-state may well be causally necessitated by antecedent microphysical events.[12] For the occurrence of the fusion may be determined by antecedent microphysical states, and the fusion-state may, in turn, necessitate its own causal consequences (microphysical consequences as well as higher-level ones).[13] In that case, all

[10] This would be similar to the ‚multiple realizability' of higher-level properties, see Humphreys (1997, 12).

[11] For the presence of the higher-level property instance is ‚more proportional' to the occurrence of the consequences than the presence of the lower-level property instances, see Yablo (1992, 277).

[12] As Humphreys notes: "We may maintain that all i-level events are determined by i-level antecedents, but often this will be by way of j-level intermediaries." (1997, 14).

[13] For this reason, the existence of fusion cases is also compatible with the truth of microphysical determinism in the global formulations presented earlier, i.e. as the claim that for all microphysical events there are antecedent microphysical circumstances that, together with the laws of nature, necessitate their occurrence, or that the initial microphysical state of the world and these laws already 'fixed' how the world would evolve later. Even if these claims are true, they do not imply that the determination chains always

occurrences at the microphysical level will be determined by earlier microphysical events (though sometimes only indirectly via determination of higher-level events). As regards the 'making up' claim, the higher-level phenomenon arises from a fusion of the lower-level ones, and may, for its existence, depend on the fact that the lower-level phenomena continue to exist as parts of it (though not as parts with an independent existence).

So, if Microdeterminism^{A} is the correct rendering of the thesis of micro-physical determinism (insofar as the latter claim is relevant for the issue of mechanism), the latter thesis does not imply mechanism. Thus understood, it only claims that all microphysical events (or states) are necessitated by earlier such events (or states). It makes no claim about *how* they are necessitated and does not exclude that the necessitation runs via necessitation of higher-level states which lie 'in between' on the causal pathway.

What Microdeterminism^{A} may still imply, of course, is – as Anscombe thought – that all higher-level phenomena are necessitated to occur by prior circumstances. As stated, even in the fusion cases the fusion-state may be so necessitated. But *this* implication would not establish mechanism, either, since necessitation by prior circumstances does not, for Anscombe, by itself deprive an event of its causal efficacy. The latter point is clear from her discussion in "The Causation of Action". On the one hand, she clearly does not think that the individual microphysical phenomena would lose their causal efficacy if microphysical determinism were true, even though the occurrence of each microphysical event would then be necessitated by prior events (see, e.g., 1983, 103). On the other hand, Anscombe explicitly considers and regards as consistent a position which combines a rejection of mechanism with an acceptance of higher-level determinism (1983, 105 f.). Such a position would take higher-level phenomena to be both causally efficacious and predetermined. So, even if all higher-level phenomena were predetermined by antecedent circumstances, this would not, for Anscombe, imply mechanism.

To rescue Anscombe's argument that microphysical determism leads to mechanism, we would have to understand micro-determinism or the composition claim in a more demanding way. Let us consider three possible moves of this kind.

First, my use of Humphreys' scenario exploited the fact that Microdeterminism^{A} only requires that at *some* prior time there were microphysical factors which made the subsequent effect necessary, but not that there were such factors at *all* antecedent times. Perhaps, though, micro-physical determinism had better be understood as the claim that for *all* antecedent times there are (concurrently present) necessitating microphysical factors for the effects which occur later. The latter condition is not met in the fusion cases. For, as long as the micro-physical factors are fused into the higher-level phenomenon, there are no such concurrently existing or occurring microphysical factors for the subsequent effects.

But this alternative reading of the claim of micro-physical determinism would be fairly implausible, since it saddles the determinist with a problematic view which she is not committed to in virtue of her claim that 'the past determines the future'.

run exclusively via other *microphysical* events, and therefore do not give us reason to regard the fusion-state as epiphenomenal.

Namely with the view of a necessarily step-wise evolution of the universe, such that whenever features of one state of the universe necessitate features of a later state they must do so via necessitating features at each intermediary step. It excludes the possibility of necessitation at a (temporal) distance or of necessitation by temporally extended processes, whose causal efficacy cannot be reduced to the efficacy of instantaneous states obtaining during this process. This seems an undue weakening of the determinist claim.

Furthermore, even if microphysical determinism were understood in this stricter way, it is far from clear that this would entail mechanism. We could still imagine that the microphysical states (without going out of existence) contribute to the coming into existence of an additional, concurrent, higher-level state. If the consequences of this new state are both dependent on its existence and are always the same, regardless of which specific micro-physical states form the basis for its coming into existence, we would have good reason to regard the new state as causally relevant for these consequences. For that state's obtaining, rather than the obtaining of the specific microphysical states, would 'make the difference' to whether the consequences obtain.[14]

Second, one might argue that the existence of higher-level phenomena which can cause microphysical effects is *ipso facto* incompatible with microphysical determinism. Would not any possible changing of the microparticles' behaviour by higher-level phenomena conflict with the assumption that once the distribution of the microproperties is fixed for time t, the distributions for all later times t' are fixed, too? But, as we have seen, there is no necessary tension between the causal efficacy of higher-level property instances and the 'fixing' idea, because which higher-level properties are instantiated might be fully necessitated by antecedent microproperty instances. The tension only arises when we add to the original determinism claim the further claim that *only* earlier microphysical phenomena can determine later ones, or can contribute to determining later ones, *and nothing else can*. But the latter claim – a claim of 'causal determination closure' of the microphysical level against downward causal influences[15]– is a different claim from microphysical determinism. Determinism is a positive thesis about determination; it is neither a claim about the causal pathway by which this determination is ensured nor a negative claim about which items *cannot* determine the effects. Also Anscombe's own understanding of determinism comprises only the former positive thesis.[16]

Third, one might try to defend the connection between (micro)physical determinism and mechanism by giving a stronger interpretation to the phrase "the particles of which all things are composed". I have understood the composition claim to be not only about higher-level objects, but about higher-level phenomena in general.

[14] Again, the presence of the higher-level property instance would be ‚more proportional' to the occurrence of the consequences, Yablo (1992, 277).

[15] Jaegwon Kim has prominently used a similar principle to argue against the possibility of downward causation, e.g. Kim (1993). Compare Humphreys' distinction between 'i-determinism' and 'i-closure', (1997, 6).

[16] „For a result to be determined is for no other result to have been antecedently possible." (1983, 103). Nor can we read the negative claim into Anscombe's formulation of "strict and complete" (1983, 103) determinism. 'Complete' here is plausibly read as '*all* features of the latter microphysical events are necessitated'.

Should we read it as even stating that all higher-level phenomena are 'nothing over and above' the lower-level phenomena from which they 'arise'?

The composition claim, thus understood, and the thesis of micro-physical determinism would indeed entail mechanism – but only at the price of making the issue of determinism completely irrelevant for the question of mechanism. For on that reading the composition claim would entail mechanism all on its own. This cannot be what Anscombe had in mind: She clearly considered the latter issue as important for the question of mechanism. (Besides, the composition claim thus understood could hardly claim much plausibility. It is widely held that the special sciences are concerned with higher-level phenomena that are not simple 'aggregates' of microphysical phenomena.)

2 Agency incompatibilism

Thus, Anscombe's argument that microphysical determinism would lead to epiphenomalism about higher-level phenomena should be resisted. But even if intentional agency is not threatened by microphysical determinism in that way, microphysical (and other forms of physical) determinism may still imply that all our actions are predetermined. That may be bad enough – since being a true agent may require that one's action is not antecedently (fully) determined. This is the claim of Agency Incompatibilism, which, in recent years, has most prominently been defended by Helen Steward.

There are different ways to motivate this kind of position, and in this section I want to look at what I consider to be the two main ones. The first way has to do with considerations about the possibility of causal activity and rests on the idea that, in acting, an agent must exercise her active or causal powers. These powers, it has sometimes been argued, could not be such that their exercise was predetermined by antecedent factors. Alternatively, one might hold that even apart from causal considerations, our concept of agency is such that only free agents can be true agents. This second strategy is more prominent in Steward, who holds that agents are essentially 'settlers' of questions which are antecedently open (2012, 39). But there are also at least hints of the first strategy in Steward, and this 'causal' strategy has been an historically influential one. I will therefore consider it first (2.1.), before turning to the second one (2.2.).

2.1 Causal activity and predetermination

When agents act, their actions are not things that just befall them as passive sufferers, but they are active with regard to what they are doing. The distinction between activity and passivity is one we are pretheoretically familiar with: We ordinarily think that there is all the difference in the world between Jim's falling to the ground because he is pushed and loses his balance (passivity) and Jim's throwing himself on the ground as part of a film stunt (activity and agency). It is also a distinction without which we would not have our concept of agency. But how can we philosophically best make sense of it? On one traditional view, we can draw the distinction in terms of whether

the change at issue involves the agent's active or passive powers, where the former are powers to produce, the latter powers to undergo change.[17] This way of drawing the distinction has been especially attractive to those agent-causalists who hold that agents, in exerting their active powers, causally produce, *qua* substances, certain effects. Steward herself is clearly sympathetic to these ideas: She takes her position to be a version of agent-causalism (see 2012, 199 f.) and also believes that agents, in 'doing causal work' exert their causal powers (2012, 209). (Also in Anscombe, we find much that is congenial to the agent-causalist position. For instance, many of the specific causal predicates she lists in (1971, 137) can take on as subjects persons, agents, or other substances, which suggests that Anscombe is quite willing to accept genuine substance-causes alongside other causes.[18])

Understanding agential activity in terms of the agent's active powers will lead one to embrace Agency Incompatibilism if one additionally holds that active powers are such that their exercise cannot be antecedently determined. This further view, too, has been held by several agent-causalists, probably most famously by Thomas Reid, who thought that active powers must be such that their bearer can not only exert, but also refrain from exerting them. For, as he claimed, "[p]ower to produce any effect implies power not to produce it" (1788, 35). My following discussion of the 'causal' route to Agency Incompatibilism will concentrate on this last crucial step:[19] Why should we believe that A can only have the power to produce X, and can only be considered as the real cause of an effect, if A could refrain from producing X and her causing the effect was not antecedently determined?

We should start by noting that *prima facie* evidence speaks against such a requirement. In our ordinary usage of causal terms, we see no contradiction in assuming that A produced an effect, but that something else made it necessary that A produced this effect. 'Who made that noise? It woke me up.' – 'Sorry, I did, I gave a shout.' – 'Why did you?' – 'I just had to shout, I was too frightened by the dark.' This dialogue seems perfectly in order, and the second answer does not undermine the appropriateness of the first one (though it may rule out the person's moral responsibility for shouting). The person who shouted did act and produced a result, even though something made it necessary for her to do what she did. In our ordinary explanatory practices, we distinguish between the question of whether someone is *causally* responsible for a

[17] E.g. Locke (1975), 234: "*Power* thus considered is twofold, viz. as able to make, or able to receive any change: The one may be called *Active*, and the other *Passive Power*.".

[18] Her remarks in (1983), 91, on the variety of possible ‚causes' are equally liberal.

[19] I will be interested only in the general claim that active powers are at odds with predetermination, though, not in Reid's specific version of this claim. There are several passages which suggest that Reid took the connection between '(active) power', 'agency' and 'cause', on the one hand, and the absence of necessitation, on the other hand, to be a conceptual one (e.g. 1788, 288 f.) and that he took – contrary to Locke – 'active power' to include, by its meaning, the power not to produce the effect. (See also Rowe 1991, 49 f., and 52 f. on Reid's 'strict' sense of cause.) Given the feature of everyday causal language I discuss presently, I take it to be fairly implausible that Reid's above claim follows on merely conceptual grounds, as long as we stick to the ordinary meaning of 'power to produce', as I will try to do in the following. (Things are different when we follow what is, in Reid's view, the 'strict' sense of the terms 'power' and 'agency' (see also Van Cleve 2015, 378 f.), but I will not do this here.) So, differently from Reid, I will consider the claim that activity is incompatible with predetermination as a substantive philosophical thesis rather than as a conceptual claim.

result and the further question of whether she acted freely. Why should we think our ordinary practice goes wrong in drawing this distinction? In the following, I want to discuss what I take to be the two strongest grounds for suspecting that it does.

2.1.1 The threats of redundancy and mere passivity

The first is the worry that antecedent determination might make the role of the putative cause redundant. What would there be 'left to do' for it if antecedent factors already made it necessary that the effect came about? It is useful to split this general worry into two more specific ones: A 'mere bystander' concern, that prior determination would not leave *any* role for the putative cause, and a 'mere passive sufferer' concern, that prior determination might leave a role, but not an active or causally productive one.

With regard to the 'mere bystander' concern, it is crucial to note that the fact that antecedent circumstances make it necessary that an effect occurs doesn't *ipso facto* make the role and contribution of a putative cause or agent A redundant. For the circumstances may make the effect necessary only by making it necessary that A makes this contribution or brings about this effect. In that case, far from being a mere 'bystander', A will have an essential role to play: Without A, it might even be impossible for the effect to occur at all. Imagine, for illustration, that a group of bank-robbers utter threats which (together with the clerk's perception of these threats and the absence of anything which could destroy the safe, end his life or rob him of his physical ability to move in the next five minutes etc.) make it necessary for the clerk to open the safe. Assume the threats are so dire that the clerk literally cannot resist them and is psychologically compelled to comply with the robbers' commands. In this case, the robbers' threats (together with the other circumstances mentioned) make it antecedently necessary that the clerk opens the safe and that the safe will be open in a few minutes. But this does not make the clerk's role redundant: Quite the reverse! If he were not there, and did not follow the robbers' commands, it may well be impossible for the safe to be opened, because he may be the only one who can open it.

Thinking that prior necessitation makes the agent's role *perforce* redundant, rests on a failure to distinguish between two different kinds of scenarios: (1) A scenario where antecedent circumstances make it necessary that this effect will occur *independently* of the agent and of what she is doing, and (2) a scenario where they make it necessary that the agent will produce this effect and it is only in this way that they necessitate the effect itself. Scenario (2) is the one of the clerk's being compelled to comply with the bank robbers' orders; scenario (1) is the very different scenario of the robbers' themselves knowing the safe code and having already set in motion the opening process which will necessarily lead to the safe's being opened in a few minutes. While the agent's own contribution is indeed redundant in scenario (1) – here the clerk can neither contribute to nor prevent the safe's opening –, and we cannot consider her a cause of, or active with regard to, the effect, things are very different

in scenario (2), where the agent's contribution may even be strictly necessary for bringing about the effect.[20]

Let us turn to the 'mere passive sufferer' concern: Even when A's contribution is crucial, might necessitation by earlier factors not reduce her to a 'passive victim', like a mere conduit or an 'arena' where things play out? Imagine that it is antecedently determined that you are infected by a virus at a certain time. Obviously, you play a role in the latter process: Without your being alive and having the physiological constitution you have, the infection would not be possible. But this doesn't give you an active part in becoming infected: You did not cause your infection,[21] you passively underwent it. Might determination of your behaviour by antecedent factors not generally make you passive in this way?

In order to see whether it would do so we have to examine more closely what precisely in the infection case makes you passive rather than active with regard to the infection. That some factors make it antecedently necessary that you get infected is only one possible explanation here. There is also the alternative – and, I submit, better – explanation that it is the fact that factors *extrinsic* to you played the crucial part in making your infection necessary. What makes the latter explanation preferable is that our qualification of an entity as 'producing change' rather than 'undergoing change' is always connected to explanatory interests. We want to know where to look for identifying the crucial factor which explains the change. If its place is among the object's own intrinsic properties and powers, the object is, *ceteris paribus*, considered as active; if it is among factors extrinsic to the object, the object is passive. As Harré and Madden aptly put it, how active an object is depends on "the degree to which we assign responsibility for particular behavioural manifestations between intrinsic conditions and extrinsic circumstances" (1975, 89).

Drawing the distinction between activity and passivity in this way presupposes a distinction between intrinsic properties and extrinsic factors. I cannot, in this paper, provide any detailed account of this distinction. But it should be clear that the distinction cannot simply be drawn in spatial terms, i.e. cannot simply be drawn according to whether the features in question are located, or based on parts, "within the spatial envelope of the thing" (Harré/Madden 1975, 87) or outside it. When you have been injected a poison which is destroying blood-cells in your body, it is not *you* who is actively causing the destruction of these blood-cells, even though the virus is working 'inside' you. According to a more plausible view (and one more in line with Harré's and Madden's own proposal, loc.cit.), the intrinsicality of a feature of an object depends on whether this feature is suitably connected to what makes the object the kind of object it is, or to its nature. E.g. when we ask whether you are active as a person, the answer will depend on whether the properties and powers that are chiefly responsible for your behaviour are suitably related to your rational nature. But I cannot argue for this view here: The crucial thing is that there is an intuitive distinction to be drawn here, wherever the line ultimately lies.

Let us turn to the question of what it means that the intrinsic properties and powers of an object, rather than circumstances extrinsic to it, are crucial to the produc-

[20] For this argument see also Mayr (2011, 202).

[21] Unless you self-injected the virus, but that would be a very different matter.

tion of an effect. If Harré and Madden are right, this is a matter of degree. The most promising way to spell out their basic idea seems to me the following: The object's intrinsic properties and powers are crucial if (and to the extent that) the production of the effect is relatively independent from extrinsic circumstances. Such relative independence holds if (i) under the very same extrinsic circumstances as the ones that actually obtained at the time, the effect need not have occurred, and (ii) even if the extrinsic circumstances had been different, the effect might still have occurred. If both conditions are met, then what we have to primarily mention in explaining the occurrence of the effect is the object with its properties and powers, while extrinsic circumstances are clearly less important (and become less and less so, the greater the scope of variations for which (ii) is true).[22] For it was the object with its properties and powers which 'made the difference', or at least, made a significantly greater difference than the extrinsic circumstances did. The precise extrinsic circumstances that obtained did not have to obtain for the effect to occur, if (ii) is met, and did not ensure that the effect occurred, if (i) is true.[23]

If this is right, it is not determination by antecedent circumstances *per se* which undermines your active role, but the fact that factors extrinsic rather than intrinsic to you were crucial in determining what happened. In particular, determination by your own nature or by features that make you the kind of entity you are, does not undermine your active role.[24] This fits well with our intuitive assessment of particular cases. Imagine that an agent's character is such that she must act in a certain way on a particular occasion. Just think of the often-cited case of Luther's 'Here I stand, I can do no other' and assume that, being the person he was, Luther really could not on that occasion act otherwise. Maybe, under these circumstances, Luther was not morally responsible for what he was doing. But it seems highly implausible to claim that he was not *causally* responsible, or not a true agent in doing what he was doing. After all, what his character made necessary was that Luther *acted* in the way he did! Luther's behaviour (we assume) was a necessary expression of the person he was, and it could only be so[25] because he was acting. Had Luther's behaviour been non-actional (e.g. the result of a spasmic seizure whose causal influence on his body's motions bypassed Luther's agency), we could no longer say that it was Luther's character which made this behaviour necessary.

Given that the difference between activity and passivity can plausibly be drawn in terms of whether to locate the crucial factor for the occurrence of an effect among features intrinsic or extrinsic to you rather than in terms of whether your behaviour is

[22] For a development and defense of this proposal see Mayr (2011, 205 f.).

[23] Conditions (i) and (ii) can be compared to Yablo's idea of ‚proportionality' of cause to effect, which Yablo takes to spell out the core notion that "the cause was the thing that ‚made the difference' between the effect's occurring and its not" (1992, 274).

[24] It is crucial to note that drawing the distinction between activity and passivity in terms of the relevance of your properties does not mean that these properties *rather than you* are causally active. Your properties make you active and their possession enables you to cause effects; they are not ‚competing' with you for causal influence.

[25] Under the circumstances of this case, that is. In other cases, we can imagine that, e.g., a blush is a necessary expression of who the person is, too.

antecedently determined or not, we should therefore reject the worry that antecedent determination would *perforce* exclude an object's causal or active role.

2.1.2 Causation and the initiation of change

Let us to turn a second reason for holding that a truly active and causal role is incompatible with antecedent determination. This reason is based on an acceptance of Locke's point that a truly active power must be a power to begin an action 'from scratch'.

Locke makes this point in the course of his argument that the observation of physical bodies does not really give us the idea of active powers. When we see one billiard-ball hit another ball and set that in motion, we do not really see, so Locke, the first ball producing the motion of the second, but only its 'transferring' the motion it had itself received earlier from another source. For we do not see a true "beginning of motion" and thus "not the Production of the Action, but the Continuation of the Passion" (1975, 235). (It is only when we reflect on our minds and wills, so Locke, that we get an idea of true 'Production').

This line of thought can easily lead one to think that true activity is incompatible with predetermination since the latter would exclude the possibility of starting an action 'from scratch'. But an argument based in this way on Locke's claim about the origin of our idea of active power would be unconvincing. This is so even if we were to grant Locke's point that, in the billiard-ball case, we would not regard the first ball as 'active' when it sets the second one in motion.[26] For any plausibility this point may have rests on the fact that the kind of change that the first ball induces in the second one is the very same kind of change as the one that was originally produced in it: The first ball seems just a transfer conduit for the change, because it "only communicates the motion it had received from another, and loses in it self so much, as the other received" (Locke, 1975, 235). Had the changes at issue been fundamentally different, we would not be tempted to describe the case as one of mere 'Continuation of Passion', but would ascribe to the first object a transformational activity. Imagine you make an object magnetic by applying a sufficiently strong electric current to it. Once the object has become magnetic, it will, by necessity (we will assume), attract certain metals. When we observe the overall process 'applying the electric current – the object's becoming magnetic – its attracting a metal', we can easily imagine that the steps are deterministically connected. But would we therefore say that attracting the metal was only a 'Continuation of Passion' on the object's part? Hardly so.

Anscombe's own arguments in "Causality and Determination" provide additional support for rejecting the Lockean argument. As she insists, contra Locke, we *can* perceive the causal activity of physical bodies, when we see, e.g., fire burning paper (1971, 137). And if the argument from the last paragraph is correct, there is no reason to think that what we see there is, in general, not 'real' causal activity, but only a "Continuation of Passion".

[26] Even this is doubtful. See Hyman (2015, 27 f.), for a criticism of Locke's example. For the following critique of Locke's argument see also Mayr (2011, 201 f.).

2.2 Agency and settling

We can therefore conclude that there is no good reason to think that true causal activity is incompatible with prior determination. But this does not yet complete the case against Agency Incompatibilism, because, even apart from concerns about causality, our concept of agency might be such that only free agents are truly agents. This is Helen Steward's central argument in her *Metaphysics for Freedom*, which rests on the idea that agents are essentially 'settlers', i.e. beings that, in acting, resolve a question which was, up to that time, 'open', i.e. could have been resolved one way or another (2012, 39). The possibility of such settling, Steward argues, is incompatible with determinism: For "surely it is a condition of being truly able to settle something that it has not already been settled in advance of one's potential intervention." (loc. cit.)

Steward adduces several reasons for holding that an incompatibilist notion of 'settling' is part of our conception of agency. Some of her arguments are directed specifically at reductionist compatibilist theories of agency which try to analyse the agent's own role in acting in terms of the causal contributions of her mental states. These arguments need not concern us here: While it is highly plausible to hold that this reductionist project cannot adequately capture the agent's *own* role, this, it seems, has nothing essentially to do with the issue of determinism. There are, after all, non-reductionist accounts of (free) agency which are compatibilist.[27] But, in addition, Steward also relies on the idea that an agent can only be considered as the source or origin of her action, if the latter is not pre-determined: "The agent can be the arché, precisely because she is a settler: because the chain of conditions from which her action results cannot be traced back along lines of inevitability beyond her. She really is an initiator, in a sense, of what happens; (…) in the sense that it has, at any rate, no prior necessitating conditions." (2012, 246).

Why should we think that sourcehood requires the absence of 'prior necessitating conditions'? Steward's idea here seems to rests on the picture that a true source or agent must 'set things in motion', or 'get things going', which is most naturally conceived of as initiating a new causal chain. Steward specifically claims that, as part of our folk psychological notion of agency, "[a]gents' actions are conceived of by us […] as newly initiated injections into the course of history" (2012, 78). She also calls substance-causes "movers" (2012, 212), which similarly suggests the picture of their setting things in motion which would otherwise remain at rest, and thus of 'starting a causal chain anew'.

But this is not the only way we can think of things as 'sources' of what occurs – and, interestingly, some of Steward's own remarks describe an attractive alternative. Not all our actions, Steward points out, are consciously initiated. What makes my sub-intentional actions nonetheless *my* actions, is their integration in a complex system where the personal and conscious level *can* 'supervise' and control what is going on. It is "in virtue of my possession of an ongoing *capacity* to prevent altogether, stop in its tracks, reverse, alter, change the direction and speed of, or otherwise affect the motion in question" (2012, 52) that the motion is a result of my agency. Steward

[27] E.g. Markosian's compatibilist agent-causal view of free action, (1999).

states that this capacity is sufficient for "my activity" to "constitute[s] a settling by me of what in fact occurs" (loc. cit.).

For such cases of 'settling', the simile of 'starting off a causal chain anew' does not seem apt, though. The right comparison is to something that is going on anyway and which I supervise and can intervene in. The crucial contrast here is not between no change occurring vs. my putting things in motion, but rather between the 'normal' course of events in a certain context (i.e. the course that would occur if I didn't intervene) vs. my changing this course. Drawing this latter contrast requires a distinction between what counts as a factor which is already part of the 'normal' course of events in that context, on the one hand, and what counts as an external intervention into that course, on the other hand. But we need not necessarily presuppose that the latter intervention (i.e. my intervention) is not, itself, determined by any other features whatsoever. More specifically, we need not presuppose so, as long as the intervention is not determined by factors which are part of the 'normal' course of events *in that context*. We only have to assume that the intervention can make a difference to that course of events – and this can be true even if it is determined.

Given this alternative way in which an agent can be a 'source' of what occurs, 'sourcehood', as such, can hardly be seen as incompatible with determinism. It may be incompatible with determination by a certain set of features – i.e. namely those features which are part of the 'normal' course of events in the context the agent intervenes in; for if the agent's intervention was determined by such features, we could not properly distinguish her intervention from this normal course. But incompatibility with this specific kind of determination is a recognizably different thing from incompatibility with (physical) determinism as such.

3 Intentional agency and indeterminism

Since neither the falsity of mechanism nor the possibility of agency as such require physical indeterminism, let us turn to the question whether the characteristic features of *intentional* agency require it. In line with Anscombe's remark in (1971, 146), we have to look for an explanation of this incompatibility which not only applies to human agents, but also to non-human higher animals which we would not consider as 'free'. From Anscombe's remarks in other places (e.g. 1957, 86) we can see that one feature that crucially distinguishes humans from other animals, in her view, is the latters' lack of linguistic capacities. Connected to this is the fact, stressed by Anscombe, that the wants and intentions of humans, differently from those of other animals, can concern 'generalities' rather than merely particular objects, times, and circumstances. "The human wants things like health and happiness […] and virtue and prosperity, he does not simply want, e.g., that such-and-such-a-thing should be in such-and-such-a-place at such-and-such a time" (1962, 98). But these differences do not mean that animals cannot act intentionally.

"[W]e certainly ascribe intention to animals. The reason is precisely that we describe what they do in a manner perfectly characteristic of the use of intention concepts: we describe what *further* they are doing *in* doing something … the cat is stalking a bird *in* crouching and slinking along with its eye fixed on the bird and its

whiskers twitching. … The enlarged description of what the cat is doing is not all that characterises it as an intention …, but to this is added the cat's perception of the bird, and what it does if it catches it. The two features, knowledge and enlarged description, are quite characteristic of description of intention in acting." (1957, 86).

These features make it appropriate to answer the question 'why is the cat crouching?' by saying 'because she is stalking the bird', though the cat could never give that answer herself. In giving that explanation of the cat's behaviour, we are not merely embedding her crouching within a more-encompassing process. Our explanation also brings in a normative element (if only a fairly 'flat-footed' one), by introducing a standard for 'success' for what the cat is doing which she can meet or fail to meet. If she is truly crouching because she is stalking the bird, things can *go wrong* (the bird may become aware of her presence and escape), and if the cat finally catches the bird she has succeeded in what she was up to.

Which features of intentional behaviour might make it incompatible with (micro-) physical determinism? I will consider two candidates, namely the connections between intentional agency and, on the one hand, the adaptation of behaviour to changing circumstances (3.1.) and, on the other hand, practical deliberation (3.2.).

3.1 Determinism and adaptation

A piece of behaviour must be embedded within a fairly complex structure or 'surrounding', characterized by (potential) adaptation to changing circumstances, to count as intentional. If the cat is crouching because she is stalking the bird, a great many things must be true of her that go beyond what she is presently doing, and even beyond what she will do. It must be true, e.g., that the cat will change course if the bird does, or, even if the bird remains still, that she would change course if the bird started to move. One might worry that this complex adaptive structure would *per se* be at odds with determinism. But why should it be? The structure can be spelled out in terms of material and counterfactual conditionals about what the animal does, or would do, in reaction to changing circumstances, and these conditionals can be true even if determinism holds (just as they can be true under indeterminism). Of course, if determinism is true, then the animal's capacity to adapt will be limited in some respects; but this is something we should realistically expect anyway.

3.2 Determinism and practical deliberation

Second, in the case of us humans at least, intentional agency is connected to the possibility of practical deliberation, and the latter might only make sense if determinism is false. In a book review, Anscombe expresses some sympathy for the idea, ascribed to Aristotle, that determinism would imply that "deliberation and choice are futile" (1980).[28] This idea is sometimes motivated by the thought that practical deliberation and choice presuppose that different options are open to the agent and are pointless when one outcome is already necessary. Anscombe in (1956, 6) identifies the latter thought in Aristotle's discussion of the sea-battle in "De Interpretatione", where it

[28] I am indebted to Christian Kietzmann for pointing out this passage and the following one to me.

"is first of all the nature of deliberation that makes him think that the fact of human action proves the dialectic must be wrong" (i.e. that what happens in the future *must* happen).[29] The idea that deliberation requires the openness of different options is one which has seemed compelling to many other philosophers, too. E.g. Peter Geach has argued that "'appeals to reasons'" only make sense if we presuppose free will and assume that how persons assess arguments "is not already determined by the bent of their minds" (2000, 78).

There are very different reasons for believing that practical deliberation would be undermined by physical determinism. We can, schematically, distinguish between those reasons which have to do with the specific object or typical end-point of such deliberation and those reasons which have to do with the kind of process deliberation is meant to be, namely one which is responsive to reasons. First, one can argue that, since practical deliberation is about what to do and is meant to lead to a decision between options, it only makes sense when we have different alternatives open to us, and determinism would rule out the existence of relevant alternatives. However, this line of argument does not seem compelling: As many philosophers have convincingly argued, practical deliberation about what to do does not plausibly presuppose that there are alternative possibilities in a sense which would require indeterminism.[30] What practical deliberation requires in order to make sense is only that our course of action is not antecedently settled independently from the deliberation process. But this is compatible with that process, and our resultant action, being determined by antecedent conditions, as long as this determination goes via the steps in the deliberative process.

Let us turn to the second kind of reason for thinking that physical determinism would undermine the possibility of practical deliberation, i.e. that deliberation must be responsive to reasons, and to considerations about the goodness of certain courses of action or the truth of propositions. Such responsiveness may seem impossible if our acceptance of the steps in our deliberation is already determined *by something else*: I.e. not by our perception of reasons and our responsiveness to them, but by factors which are themselves non-rational and insensitive to reasons.[31] And would not microphysical determinism as well as other forms of physical determinism imply that all the steps are determined by *something else*? For microphysical features cannot themselves be sensitive to reasons. So, once our behaviour is determined by such features, how can it also be a response to reasons?

This worry seems to apply when no actual deliberation occurs, too. Even when we don't actually deliberate, our intentional agency is responsive to reasons, to some degree. When we act intentionally, we usually adapt, e.g., our course of action, when we see another one as better. But how can such adaptations count as responses to reasons if they are physically determined?

[29] However, Anscombe also warns against interpreting this to mean that "'the nature of deliberation presupposes freedom of the will as a condition.' That is not an Aristotelian idea." (loc. cit.).

[30] See Bok (1998, 104 ff.) for an excellent elaboration of this point.

[31] Geach (2000, 80) also formulates the problem as one of determination by *non-rational* causes (though, in other places, he raises the worry for determination in general).

To answer this question, it is useful to first distinguish two different functions practical deliberation can fulfil, and two corresponding kinds of reasons-responsiveness. After having done so, we can ask which of these functions would be at odds with determinism.

First, there is reasoning from settled ends to ways of implementing these ends. This can take the form of means-end reasoning or of reasoning by specification (e.g. reasoning from one's end to become a good sportsman to what it means to be a good sportsman in a given situation). Second, there is reasoning and reflection *about ends*: We humans can ask ourselves whether the ends we have, even very general ones, are indeed good ones or the right ones to have (see Anscombe 1962, 98). So, we humans can be responsive to purely instrumental reasons and to 'specification' reasons, and to reasons for and against having an end in the first place.

How about non-human animals – can we find these two functions of deliberation and forms of reasons-responsiveness in at least some higher animals, too? It is already a matter of dispute whether animals without language can deliberate practically at all.[32] But, as researchers (as well as philosophers) have argued, there are reasons to ascribe to several animals at least forms of intelligent problem-solving processes which are very close to our forms of instrumental-causal reasoning.[33] Things are trickier with regard to reasoning by specification, and if Anscombe is right that animals' wants and intentions cannot concern 'generalities', animals cannot take specification steps from general ends to more particular ends and circumstances. More importantly, however, there seems to be a broad consensus that non-human animals have no capacity to engage in reflection about their ends: They cannot ask whether the ends they have are really good ones to have. Anscombe, for one, took this to be a distinguishing feature of humans (1962, 98) – and scientists who are otherwise sympathetic to the idea that animals can act for reasons and deliberate tend to agree with her.[34] This might be seen merely as a consequence of the fact that research into animal intelligence focuses on the question of intelligent problem-solving with regard to set ends.[35] But there is a more principled reason against attributing to animals lacking language a capacity of reflecting on their ends. This capacity would presuppose a distinction between having something as one's end and considering it as good. For animals without language this distinction seems impossible to draw, since the fact that the animal pursues X as its end is our basis for ascribing to it the view that X is good. (For us language-using humans, things are different: We can pursue an end, though we expressly judge it to be bad.) Thus, there are good reasons to think that animals are at most capable of something like reasoning from set ends

[32] For a summary of the debate see, e.g., Glock (2009, 247 f.).

[33] See, e.g., Tomasello & Call (1997, 10 ff.). These processes could even involve forms of inference-drawing, see, e.g., Call (2006, 221 f.) and Glock (2009, 249 f.).

[34] See, e.g., De Waal (2006, 174) and Tomasello (2016, 110 and 148).

[35] For illustration take, e.g., the very recent overview of research into decision-making in (some higher) animals in Hunt et al., (2021).

to means, but not of reflecting about their ends, or of choosing a (new) end because they consider it to be good.[36]

Let us now turn to the question of which forms of deliberations, and of reasons-responsiveness, would be really threatened by physical determinism. With regard to the first form – reasoning from set ends to how to implement or realize them –, it seems very implausible that such reasoning, and intelligent processes very akin to it, would be ruled out by (micro)physical determinism. For we can easily imagine systems which implement something very similar to such reasoning and are responsive to what is conducive to reaching a list of pre-set aims, but which work in ways we take to be completely physically determined. Computers programmed to calculate how to optimize certain pre-specified features of a process (e.g. the reduction of energy costs) seem to work in just this way. They are sensitive to (some) instrumental reasons (e.g. to the fact that insulation measures reduce heating costs), because they have been constructed to work this way. But this responsiveness does not require that their functioning be indeterministic.

When we turn to the second form – reflection about ends – things may be different. Our capacity to choose new ends because we judge them to be good seems particularly hard to reconcile with determination by (micro)physical factors since it appears to make us capable of transcending any fixed program which could settle in advance how we will act. Whether this capacity is really at odds with physical determinism, is a question I cannot decide here.[37] But neither do I need to do so in order to answer our overall question of whether physical determinism would rule out intentional agency as such. For, as we have already seen, non-human animals - even those we consider capable of intentional agency – lack the capacity of reflecting on their ends. So, even if this capacity were incompatible with determinism, animal intentional agency and, by extension, intentional agency as such, would not thereby be incompatible with determinism, too.

4 Conclusion

I have reviewed three different ways to support Anscombe's suggestion that the absence of (micro)physical determination is required for intentional agency. For all three, my results have been negative. But this, I believe, fits quite well with another part of Anscombe's view about the relation between explanations by intention and physical explanations. For directly after making the tentative claim of Intentional Agency Incompatibilism, she goes on to say:

"The freedom, intentionalness and voluntariness are not to be analysed as the same thing as, or produced by, the physical haphazard. Different sorts of pattern altogether

[36] The fact that animals may have the first kind of capacity but not the second does not mean that, in us humans, too, these capacities work separately. Rather, it is plausible to hold that the way our (i.e. human) instrumental reasoning proceeds is fundamentally influenced and shaped by our capacity to reflect on our ends.

[37] For a way to resist this conclusion see Holton (2013).

are being spoken of when we mention them, from those involved in describing elementary processes of physical causality." (1971, 146).

The patterns may well be *so* different that even if the one we use for physical explanations turns out to be a deterministic one, this need not undermine the pattern we rely on in intentional explanation.

Acknowledgements For very helpful comments on earlier versions of this paper, I am indebted to Stefan Brandt, Nora Heinzelmann, Christian Kietzmann and Konstantin Weber, as well as to two anonymous referees.

Funding Open Access funding enabled and organized by Projekt DEAL.

References

Anscombe, E. (1956). Aristotle and the Sea Battle. *Mind*, 65/257, 1–15

Anscombe (1957). *Intention*. 2nd ed. 1963, paperback edition Harvard University Press 2000

Anscombe (1962). Authority in Morals. Repr. in M. Geach & L. Gormally (Eds.), *Faith in a Hard Ground. Essays on Religion, Philosophy and Ethics by G.E.M. Anscombe*, (pp. 92–100). Imprint Academic 2008

Anscombe (1971). Causality and Determination. Repr. in Anscombe, *Metaphysics and the Philosophy of Mind (The Collected Philosophical Papers of G.E.M. Anscombe Vol. 2)*, (pp. 133–147). Blackwell 1981

Anscombe (1980). Matters of Consequence. *The Times Literary Supplement*, 20 June 1980, p. 701. The Times Literary Supplement Historical Archive, http://tinyurl.galegroup.com/tinyurl/8pCMB3. Accessed 18 Jan. 2019

Anscombe (1983). The Causation of Action. Repr. In M. Geach & L. Gormally (Eds.), *Human Life, Action and Ethics. Essays by G.E.M. Anscombe* (pp. 89–108). Imprint Academic 2005

Bok, H. (1998). *Freedom and Responsibility*. Princeton University Press

Call, J. (2006). Descartes' two errors: Reason and reflection in the great apes. In S. Hurley & M. Nudds (Eds.), *Rational Animals* (pp. 219–234). Oxford University Press 2006

De Waal, F. (2006). *Primates and Philosophers. How Morality Evolved*. Princeton University Press

Geach, P. (2000). Intention, freedom and predictability. In R. Teichmann (Ed.), *Logic, Cause & Actions. Essays in honour of Elizabeth Anscombe* (pp. 73–81). Cambridge University Press

Glock, H. J. (2009). Can Animals Act For Reasons. *Inquiry*, 52(3), 232–254

Gustafsson, M. (2016). Anscombe's Bird, Wittgenstein's Cat: Intention, Expression, and Convention. *Philosophical Topics*, 44(1), 207–237. Essays on the Philosophy of Frederick Stoutland

Harré, R., & Madden, E. (1975). *Causal Powers: A Theory of Natural Necessity*. Blackwell

Holton, R. (2013). From Determinism to Resignation; and How to Stop It. In A. Clark, J. Kiverstein, & T. Vierkant (Eds.), *Decomposing the Will* (pp. 87–100). Oxford University Press

Humphreys, P. (1997). How Properties Emerge. *Philosophy of Science*, 64, 1–17

Hunt, L. T., et al. (2021). Formalizing planning and information search in naturalistic decision-making. *Nature Neuroscience*, 24, 1051–1064

Hyman, J. (2015). *Action, Knowledge, and Will*. Oxford University Press

Kim, J. (1993). The Nonreductivist's Troubles with Mental Causation. In J. Heil & A. Mele (Eds.), *Mental Causation* (pp. 189–210). Clarendon Press
Locke, J. (1975). *An Essay concerning Human Understanding*. Ed. P. Nidditch. Oxford University Press
Markosian, N. (1999). A compatibilist version of the theory of agent causation. *Pacific Philosophical Quarterly*, 80, 257–277
Mayr, E. (2011). *Understanding Human Agency*. Oxford University Press
Reid, T. (1788). *Essays on the active powers of man*. Repr. Lincoln-Rembrandt Publishing 1990
Rowe, W. (1991). *Thomas Reid on Freedom and Morality*. Cornell University Press
Schwenkler, J. (2019). *Anscombe's Intention. A Guide*. Oxford University Press
Steward, H. (2012). *A Metaphysics for Freedom*. Oxford University Press
Tomasello, M., & Call, J. (1997). *Primate Cognition*. Oxford University Press
Tomasello (2016). *A Natural History of Human Morality*. Harvard University Press
Van Cleve, J. (2015). *Problems from Reid*. Oxford University Press
Yablo, S. (1992). Mental Causation. *The Philosophical Review*, 101(2), 245–280