

# The Ontology of Organismic Agency: a Kantian Approach

## Authors

Hugh Desmond and Philippe Huneman

## Abstract

A key development in biology has been the increasing tendency to view organisms not just as programmed by genes and shaped by natural selection, but also as agents, which react in goal-driven ways to changes in their environment. This reopens old questions about how justified it is to ascribe agency to entities like bacteria or plants that lack rationality and even a nervous system. Is organismic agency genuinely ‘real’ or just a fiction? In this paper we identify the Kantian approach to this issue, and contrast it with the main positions available today, namely instrumentalism and realism. Attributing agency to organisms is not a merely a heuristic or predictive tool, nor is it to be taken literally in the way we attribute physical properties such as mass or acceleration. Rather, it is inevitable consequence of our own rational capacity: as long as we are rational agents ourselves, we cannot avoid seeing agency in organisms.

## Introduction

Stags lock antlers to gain access to mates. Arctic poppies rotate and track the sun in order to maximize solar exposure. Bacteria swim up a sucrose gradient in order to get better access to the source of sucrose. These are some examples of seemingly goal-directed behaviour that is ubiquitous among organisms. Yet, in much of 20<sup>th</sup> century mainstream evolutionary theory, the *artefact metaphor* for organisms was the dominant one: the organism was understood to be designed by natural selection in much the same way Paley’s watch was designed by an intentional creator. All apparently purposive behaviour was analysed to be a designed function of the organism, and purposive in name only (‘teleonomy’, cf. Pittendrigh

1958; Ernst Mayr 1961). In reality, all purposive behaviour is the product of the non-teleological process of design by natural selection (cf. Lewens 2005).

Today the *agential metaphor* is (once again) becoming increasingly influential. Organisms are agents that interact with their environments with goals and purposes; thus, their behaviour must be analysed with reference to the goals of organisms as wholes rather than mere collections of parts. This focus on organisms (rather than genes) is linked with several developments in evolutionary biology, most notably the Extended Synthesis, which among other things, attempts to give development and ecology a more central explanatory and causal role in evolutionary theory than was previously the case.

It is wrong to think that the agential metaphor is a candidate to ‘supersede’ or ‘replace’ the artefact metaphor. For instance, organisms engage in self-destructive behaviour that can only be explained as for the good of the group they are part of, or as for the good of their genes. This can only be understood through evolution by natural selection, or by maximizing inclusive fitness (cf. Huneman forthcoming). More fundamentally, the agential approach to organisms takes organismic purposes as given; to account for why organisms have such-and-such purposes and not others, we need a design explanation. Thus, stags tend to lock antlers because those stags who tended to not engage in sexual competition did not get to transmit their genes to the next generation. The same with rotating arctic poppies and swimming bacteria. The agential approach does not account for the genesis of organismic purposes – shelter, nutrition, social dominance, all forms of self-maintenance, reproduction.

But yet scientific explanations are replete with implicit references to goal-driven and goal-maximizing behaviour of organisms. Organismic purposes and organismic agency seem to be, in some yet to be specified sense, explanatorily indispensable. The question we tackle in this paper concerns what this explanatory indispensability means for the ontology of the agential metaphor. Is organismic agency merely a metaphor, a fiction of the human mind – perhaps an anthropomorphic projection – or is robustly real in the same way that the wings of a bird are, or the claws of a bear? This question is not decided by showing how organismic purposes are products of evolution by natural selection (they are, except perhaps for some that may be the product of purely physical processes: see later). We take it for granted that organismic purposes have arisen over time through some causal process; what matters here is what they are.

Often this issue has been framed through the resources in the philosophy of mind literature: how is it that we ascribe *intentions*, *reasons*, and *cognitive states* to organisms? (Dennett 1971, 1995; Okasha 2018). By contrast, we will draw on resources in the literature

on scientific explanation and the metaphysics of science. In particular, we will introduce the notion of *agential explanations* ('the bacterium swam up the sucrose gradient because it is an agent'), and contrast the *ontic* and *non-ontic views of explanation*. According to the ontic view of explanation, an explanans explains in virtue of referring to entities or processes that are part of the accepted ontology. In most views of causal explanation, the explanans refers to part of the causal nexus (Woodward 2003; Strevens 2008; Craver 2014). However, the ontic view also allows for non-causal explanations: the explanans may also explain in virtue of referring to mathematical entities or properties (Lange 2013). In this latter case the ontology is not purely physicalist, but is supplemented by mathematical entities and/or properties.

With regards to agential explanations, this distinction between ontic and non-ontic views can help systematize some approaches to organismic agency. One prominent ontic view (Walsh 2015, 2012) has entailed expanding the ontology to include purposes; agential explanations then simply explain in virtue of referring to these purposes. However, non-ontic views have been far more common. One common version of the non-ontic view has been to account for the explanatory role of agency as a heuristic or methodological tool for scientific investigation. An agential explanation is a first step towards explaining organismic behaviour; a full explanation cashes the behaviour in causal mechanisms. Thus apparent agential behaviour can be explained as the result of physiological mechanisms programmed or designed by natural selection (e.g. Ruse 2000). Another non-ontic view holds that agency is not a mere heuristic, to be discarded when a causal explanation is available, but is a powerful predictive tool (Dennett 1995, 1971; Okasha 2018). What both non-ontic views have in common is that they do not explain in virtue of referring to parts of the ontology.

Our paper will explore a Kantian approach, which we believe offers an interesting alternative to the ontic/non-ontic dichotomy. The choice of Kant is not artificial, as Kant – uniquely among historical authors – is quoted as an important precedent to agential approaches to biological organisms (Weber and Varela 2002; Kauffman 1993, 2000; Kitano 2001; Moreno and Mossio 2015).

Drawing on a historical author for a contemporary debate risks either anachronism – twisting the author's words to fit contemporary context – or, if too much weight is given to a nuanced and scholarly interpretation of the author, an ultimate failure to show how the author's thoughts are still relevant today. In this paper we will err on the side of scholarly accuracy. Kant's own writings on this subject are subtle and complex, and sometimes the Kantian approach is represented as a projectivist approach (e.g. Lewens 2007), or an instrumental approach for a computationally limited human mind (e.g., Weber and Varela 2002). We believe

that a closer reading of Kant's work suggests that the very ontic/non-ontic distinction is undermined.

By contrast, we will argue that Kantian approach entails viewing agential purposes (or purposiveness) as a "demand of reason". Purposiveness is necessary given our rational nature; it is not a consequence of our limited computational capacity, or of our contingent evolved nature that causes us to see agency even in places where there is none (cf. the so-called 'agency detection' cognitive modules, Atran 2002; Barrett 2000).

This paper will not broach the important question whether the Kantian approach can be 'naturalized' on par with the other approaches to agential explanation; this will be a topic for a future paper. The focus here is identifying the Kantian approach to agential explanations and contrasting it negatively with the available approaches. Making the positive case is outside the scope of the paper, and would likely require a departure from a close reading of Kant's work.

The paper is structured as follows: in the first sections we show how organismic agency is to be defined and problematized, and in doing so contrast it with the problem of design and biological functions, which has received most attention (by far) in contemporary philosophy of biology. In the second we lay out, with as much interpretative accuracy as is possible in a philosophy of science paper, what Kant's views on organismic agency were, and we translate this to contemporary context. In the third section we draw out the consequences of the Kantian approach for how we are to think about the ontology of purposes when we refer to organismic agency.

## **1. Artefacts and Agents**

Much of mainstream 20<sup>th</sup> century evolutionary biology operated within the framework of what is called 'the Modern Synthesis', a term coined by Julian Huxley (Huxley [1942] 1974). The Modern Synthesis was forged in the 1930s and 1940s by Ronald Fisher, Sewall Wright, Theodosius Dobzhansky, John Haldane, among others, and is often described as synthesizing Mendelian genetics with Darwin's theory of natural selection. It was very much focused on how allele (different versions of the same gene) frequencies change over time in response to evolutionary forces, such as natural selection, mutation, drift, or migration. It is the basis for Jacques Monod's influential portrayal of evolution as '*hasard et nécessité*', chance and necessity: evolutionary history can be explained by chance genetic mutations followed by selection of the most fit.

Organisms were given short shrift, and in essentially analysed as epiphenomena arising from changes in underlying allele frequencies. This view of organisms received an influential popularization in the work of Richard Dawkins (Dawkins 1976, 1982), who introduced a dichotomy between replicators (alleles) and interactors (organisms). Organisms are mere tools in a never-ending arms race between genes; genes are the genuine actors through which evolutionary history can be understood. Even complex behaviours such as beaver dams are expressions ('extended phenotypes') of the underlying genotype. So while it may *seem* that organism undertakes behaviour for its own benefit (e.g., behaviour to secure food or to fend off predators), it is actually for the benefit of the gene, which gets to replicate when the organism does well.

Thus, to anticipate, the theoretical resources of the Modern Synthesis were used to derive a decidedly non-realist interpretation of apparently purposeful behaviour. Organisms are simply collections of 'traits', where each trait is a product of some underlying combination of alleles. Furthermore, each trait is optimized to maximize the transmission of alleles to the next generation. This yields a view of organisms as artefacts, where the parts are traits, and the functionality of each part is designed by natural selection (see Lewens 2005 for an in-depth discussion of the artefact metaphor).

This interpretation was strengthened by the usage of some analogies borrowed from computer science. Organismic behaviour was often described as *programmed*:

The purposive action of an individual, *insofar as it is based on the properties of its genetic code*, therefore is no more nor less purposive than the actions of a computer that has programmed to respond appropriately to various inputs. (Mayr 1961, 1504, our emphasis)

So even if the behaviour of an organism may seem goal-directed, that is only because its genetic code has been 'programmed' by natural selection to direct the organism to react in certain ways to certain inputs, and other ways to other inputs. Organisms are no more goal-directed than computers are.

However, we emphasized Mayr's qualification 'insofar as it is based on the properties of its genetic code'. This is crucial to realize that, even for biologists operating squarely within the Modern Synthesis, they did not wish to push the artefact or computer analogy too far. Mayr is not claiming that individual organisms behave exactly like pre-programmed computers; only that some aspects of their behaviour are determined by environmental inputs in the way a computer program responds to user inputs.

To see how the artefact metaphor was never taken literally, we can go back to the very foundational basis of the Modern Synthesis: the analysis of phenotypic variance as proposed by Fisher (Fisher 1919). This states that, in general, only a part of the variation of phenotypes in a population is explained by a corresponding variation in genotype. The rest is variation in environment (impacting how the organism develops), or variation in how genotype and environment correlate (cf. e.g. Hamilton 2009).

$$V_P = V_G + V_E + V_{G \times E}$$

Thus, no biologist holds that organismic behaviour is in all its respects determined by a genetic program – even Dawkins devotes a chapter in *The Extended Phenotype* to debunking the ‘myth’ of genetic determinism. The programming analogy cannot be taken too far.

Similarly, the artefact metaphor cannot be taken too far either. A genotype may be designed for a particular type of environment, i.e., there may be a particular ‘normal’ environment in which the bulk of the selection for that genotype occurred. In that normal environment, the genotype develops into an adaptive phenotype. However, in reality environments are highly heterogeneous, so in a population of identical genotypes, only a fraction will develop in the normal environment. Other environmental inputs – inputs that differ from the ‘normal’ environment – cause the organism to diverge from its ‘designed’ phenotype.

So even if we acknowledge that the artefact metaphor was never taken literally, developmental biologists have, by and large, rejected that the artefact metaphor is even a good metaphor of organisms. Developmental biology focuses in detail on the precise processes which govern organism’s development – for instance, embryogenesis. These processes are themselves products of natural selection, but yet are sometimes decoupled from genetic evolution in important ways, where a phenotype may be maintained in a lineage even though the underlying genotype changes (Müller and Newman 2003). These processes can constrain phenotypic evolutionary novelty that is not comprehensible solely in terms of genotype variation.

This developmental perspective on organisms and on organismic evolution has been gaining influence, to the extent that many are now calling for an ‘Extended Synthesis’ that revises the Modern Synthesis in important ways. The Extended Synthesis is very multi-faceted, and while there is currently no canonical way of summarizing it, a highly simplified summary that emphasizes the importance of developmental processes for evolution is that the Extended

Synthesis views evolution as ‘the evolution of epigenetic processes’, as opposed to the evolution of allele frequencies for the Modern Synthesis (Müller and Newman 2003).

Most relevant for the purposes of this paper is how the phenomenon of phenotypic plasticity is treated in the Extended Synthesis. Plasticity in a trait refers to environmentally-induced phenotypic variation, and is the term  $V_E$  in the equation above. It is an incredibly basic phenomenon: it simply refers to how different environments cause genotypes to develop into different phenotypes. This can include seemingly agential behaviour: on response to chemical cues emitted by sea slugs, bryozoans will develop spines to defend themselves (Godfrey-Smith 1996). However, it can also refer trivial phenomena such as the stunting of growth of a plant in response to poor nutrition. Strictly speaking, there are few if any organisms that lack some form of phenotypic plasticity in some of their traits.

The phenomenon of plasticity was not considered to be of any special significance until the work of Bradshaw (Bradshaw 1965) – before him, environmental perturbation was often viewed as a source of noise. What Bradshaw showed was that plasticity in a trait can be adaptive, mainly in heterogeneous environments. He distinguished between four ways in which plasticity can be adaptive (Bradshaw 1965, 21): (1) when the environment changes on a time-scale that is equal or shorter than generation time, (2) when the environment varies over very short spatial scales, (3) when the magnitude of environmental variation is very large, (4) when it is beneficial to maintain a stable phenotype in a population while maintaining genetic diversity.

The parallels between phenotypic plasticity and cognition have not gone unnoticed. Theorists and philosophers concerned with how cognition evolved often take the evolution of adaptive phenotypic plasticity to be a model (van Duijn, Keijzer, and Franken 2006; Lyon 2017; Calvo Garzón and Keijzer 2011; Sterelny 2000; Caporael, Griesemer, and Wimsatt 2013; Godfrey-Smith 1996). Organisms exhibit a whole range of cognitive, or at least apparently cognitive<sup>1</sup>, behaviours: they sense changes in the environment, are able to process this information and select a response from a repertoire of responses. Far from being a late-stage development in evolutionary history, we see these types of behaviours in bacteria, which can undertake evasive action upon detecting predators (Pérez et al. 2016), or swim to a food source upon detecting sucrose gradients (Auletta 2013).

---

<sup>1</sup> The application of the term cognition, as well as other terms such as communication or memory, to organisms such as bacteria remains a controversial point. See discussion in (Lyon 2015)

These considerations – the importance of developmental processes, phenotypic plasticity, proto-cognition – are some of the most important motivations to introduce the concept of organismic agency, which refers to the capacity of an organism to respond to changes in its environment in a way that (apparently) is goal-driven. Organismic agency is thus understood as much broader concept than the agency traditionally ascribed to human, rational subjects, which is typically characterized by means of some mental state, like an intention (cf. Schlosser 2015). The approach to organismic agency in the biological sciences, by contrast, blackboxes whatever cognitive processing may or may not be going on. In this sense organismic agency is best understood as an ecological property (cf. Walsh 2015), namely, a property of the interaction between organism and environment.

## 2. Agential Explanations

As indicated above, the basic intuition underlying the concept of organismic agency is that it is the capacity of an organism to maintain itself in response to environmental perturbation. However, this is unsatisfactory as a definition, since such a vaguely-defined capacity could also be attributed to all sorts of dissipative structures, like Bénard convection cells. Bénard convection cells appear spontaneously when the temperature gradient in a liquid is large enough, and maintain their organization even in face of perturbation in their environment, such as movement of the container walls (Manneville 2006). Such an application of the concept of agency to include such nonlinear physical systems is widely seen as too broad (Moreno and Mossio 2015). Where then to place the boundary between agential and non-agential complex systems? This is a complex question, and a great number of rival accounts have been given, pointing to various factors such as internal organization, or control of environmental constraints (Moreno and Mossio 2015; Barandiaran, Di Paolo, and Rohde 2009; Skewes and Hooker 2009; Shani 2013; Burge 2009; Horibe, Hanczyc, and Ikegami 2011)<sup>2</sup>.

For the purpose of this paper, we will operate with a rather abstract definition of agency:

An system is an agent when (1) it possesses a certain purpose  $P$ , where  $P$  is a particular state of the system, (2) it maximizes the realization of  $P$  in response to environmental change, (3) the system itself is a cause of the realization of  $P$ .

---

<sup>2</sup> The literature on naturalized agency is interdisciplinary to a high degree, with contributors with backgrounds ranging from biology or nonlinear physics to artificial intelligence, robotics, or cybernetics. A systematization of all various contributions and approaches is still lacking.

Questions about where these purposes come from (are they products of evolution, or are they laws of nature, etc.), or how precisely an organism goes about maximizing the realization of a purpose – all these questions are bracketed. Furthermore, since this definition does not specify which types of system can be considered purpose-maximizing systems, it does not pre-empt the debate on how complex systems are to be divided into agents and non-agents. A scientist could attribute agency to a Bénard convection cell, or to self-propelling oil droplets (Hanczyc and Ikegami 2010 attribute autopoiesis and autonomy, which themselves have been taken as models for agency by e.g. S. Kauffman and Clayton 2006), but only as long as then purposes are ascribed to the droplets and the locus of causation of the process of maximization is to placed with the droplet-system as a whole. Whether this can actually be done is a separate question.

Some further explanation of the elements of this definition may be helpful. (1) This definition simply equates a purpose with a particular *state*. For organisms, purposes may refer to developmental states, physiological states, or behavioural states. (2) Goal-directedness is interpreted as a maximization or optimization. This equation of purposefulness with some type of optimization is common across the sciences. (3) The system itself must be considered to be a cause of the maximization behaviour to exclude clear non-agent cases of optimization, such as the marble rolling down into the middle of the bowl (minimizing gravitational potential energy). Here the marble is not considered the cause of its own behaviour. By contrast, an organism that modifies its phenotype in order to be more adaptive to a new environment is considered to be a cause of that modification.

With this operational definition of agency in place, we can introduce ‘agency explanations’ as scientific explanations that explain in virtue of reference to a system’s agency:

- |              |   |
|--------------|---|
| Explanandum: | In response to environmental change $E_1 \rightarrow E_2$ , the system undergoes the change $S_1 \rightarrow S_2$ .   |
| Explanans:   | (1) The system has purpose $P$ ,<br>(2) $S_2$ maximizes the realization of purpose $P$ in environment $E_2$ ,<br>(3) the system itself is a cause of the realization of $P$ . |

As to the explanatory relation between explanandum and explanans, this may be considered causal or non-causal, depending on one’s view of the causal nature of explanation. An agency-explanation is structurally similar to *optimality explanation*: it explains the explanandum as some optimal state of affairs, given certain conditions (i.e., the purpose of the system), and just

as optimality explanations have been argued to be non-causal (Rice 2015, Huneman 2018) as well as causal (Strevens 2008), agential explanations could likewise be argued to be causal or non-causal. After all, the relation does possess a probabilistic counterfactual structure: if the purpose  $P$  were replaced with some other purpose  $P'$ , then the observed behaviour would be (much) less likely. Whether this is sufficient for deeming the relation causal, depends largely on one's definition of causation (cf. Desmond, Huneman XX).

It is clear that, as stated, the relation is not one of causal production. There are clearly some causal processes causing the system's change of state  $S_1 \rightarrow S_2$  (e.g., neurological processes causing behavioural change), but an agency explanation, at least as stated above, does not refer to such causal processes. It explains the behaviour in terms of a purpose, and a condition linking that purpose to concrete conditions in reality (i.e.,  $S_2$ ).

As a final remark, this abstract way of formulating agency explanations allows them to be discerned in various domains, according to how the purpose is defined. For instance, when the purpose defined minimally as a simple utility measure such as fitness, an agency explanation takes the form of the *phenotypic gambit* (Grafen 1984, 2014), which is one of the main working hypotheses underlying the field of behavioural ecology. The phenotypic gambit is the assumption that organisms behave in such a way that maximizes their fitness (expected number of offspring), and seeks to explain organismic behaviour as such.<sup>3</sup>

Other purposes can also be ascribed to organisms: keeping the physiological states within certain ranges, viability, or reproduction. In principle, agency explanations, as defined above, can also be extended to rational agents, where the purpose is defined as value or a general utility measure. Such explanations are commonplace in economics, which often (and controversially) assume that economic actors are utility-maximizing agents (which is of course unrealistic, cf. Tversky and Kahneman 1974).

## 2.1 Agential vs. functional explanations

The shift in contemporary biology towards the agential metaphor suggests that agential explanations warrant closer philosophical attention. Yet, when it comes to the problems of teleology, philosophy of biology has been overwhelmingly focused on the teleological language inherent to the artefact metaphor, namely functional statements and functional explanations, e.g., “the heartbeat in vertebrates has the function of circulating blood through

Commented [HD1]: This section doesn't work quite yet... Maybe it's not even really useful for the paper... On the other hand, some contrast with functional explanation may be needed to avoid confusion.

<sup>3</sup> Not surprisingly, this has also been dubbed the ‘maximizing agent analogy’ (Grafen 2002)

the organism” (Hempel 1959). Organism-level goals have also been used to offer accounts of functional explanations (for a discussion of the problems associated with these accounts, see Lewens 2007; Huneman forthcoming). It may therefore be useful to contrast agential explanations with functional explanations.

Initially philosophers of biology attempted to account for teleological concepts by specifying their empirical content; so for instance, Hempel sought to show that ascribing a function to some structure is to hold that under so-and-so conditions, so-and-so effects can be expected (Hempel 1959: 306). In other words, a functional explanation shows how the functional behaviour (e.g. the beating of the heart) is nomologically expected given the function of the heart and some conditions of application.

The debate took a very different turn with Wright (1973, 1976), when the concept of explanation was explicitly integrated into the definition of functions. This turn is still dominant, and today there are two main approaches to functional explanations. The first is the selected effects account, which roughly holds a function is what explains why some structure was selected for in the past (Wright 1973, 1976; Millikan 1984; Neander 1991). The second is the causal role account of functions, which roughly holds that a function is what contributes to some ‘capacity’ of a larger complex system that contains it (Cummins 1975). While most support has gone to the selected effects account, many hold that the debate remains at an impasse (Godfrey-Smith 1993)<sup>4</sup>.

Ascribing agency to an organism involves a different type of teleological statement, which involves a purpose of the whole organism rather than simply the purpose of one part of the organism. Thus the main difference between functional and agential explanations, at first sight, seems to be that functional explanations attribute a purpose (function) to a trait of an organism, whereas agential explanations attribute a purpose to the organism as a whole. Functions are attributed to traits of organisms, whereas agency is attributed to the organisms themselves.

Yet, the relation between functional and agential explanations is more subtle. As mentioned above, organism-level goals can be used to ground trait-level functions (e.g., the organizational account, Mossio, Saborido, and Moreno 2009). Arguably, the causal role account of function can be said to do the inverse, where the capacities of the organism are analysed as a whole in terms of the capacities of the individual traits.

---

<sup>4</sup> Also in this context some have appealed to the attractiveness of a Kantian approach (Lewens 2007; Zammito 2006; Huneman 2006).

Most importantly, however, agential and functional explanations seem to be complementary: the same phenomenon can receive a functional as well as an agential explanation. Consider the agential explanation, “in order to gain access to mates, the stags locked antlers”. This could be redescribed as “the function of the stag’s behaviour of locking antlers is to gain access to mates”. This point lies in continuity to Dennett’s distinction between the ‘intentional stance’ and ‘design stance’: a system that can be analysed as ‘intentional’ (‘agential’ in context of this paper) can also be analysed as ‘designed’ (by natural selection). Thus agential and functional explanations target slightly different explananda, which are themselves slight redescriptions of the same phenomenon (cf. Andersen 2016, Desmond forthcoming).

### 3. The Ontology and Explanatory Dispensability of Agency

Does organismic agency actually exist, or is it, as is often claimed of teleological concepts, a convenient heuristic at best and a lazy cognitive bias at worst? One typical way to analyse this question is to examine how precisely agential explanations explain, and whether they can be replaced by equivalent non-agential explanations. If this replacement can be done, this is seen as a good argument that agency is dispensable – definitely if the replacements give more detail about the causal processes involved in agential behaviour. This then, in turn, is viewed as supporting the *non-ontic view* of agential explanation. By contrast, showing that agency is explanatorily indispensable supports the ontic view of agential explanation.<sup>5</sup>

It is clear that many phenomena do not ask for agential explanations. If during a storm, a branch of a tree cracks and falls to the ground, and we seek to explain the change in the tree’s state, we spontaneously tend not to appeal to any type of ‘agency’ of the tree. A property of the tree as a whole could be explanatorily relevant – for example, a disposition such as brittleness could be referred to in order to explain why oaks tend to crack more than willows during storms. Nonetheless, we tend not to explain this tree ‘behaviour’ in terms of the purposes of the tree. Rather, given certain forces created by the wind, and perhaps given certain structural properties of the tree, the outcome of the branch cracking was determined. No agency or even teleology is involved.

**Commented [HD2]:** Purpose of this section: introduce how alternative types of explanation are possible of ‘apparently agential behaviour’, and connect it with ontic/non-ontic view of agential explanation. If done successfully, I think it could be a great introduction then to Kantian perspective.

---

<sup>5</sup> I’m trying to connect all this discussion of alternative explanations for ‘apparently agential behaviour’ to the ontic and non-ontic view of explanation. This paragraph is pretty crucial, but needs some serious reworking.

Non-agential explanations can also be given of typical examples agential behaviour, such as bacteria swimming up sucrose gradients. Instead of referring to the goals of the bacterial cell, one could refer to the mechanism of chemotaxis, and the conditions under which it operates. Given a certain environmental input into the mechanism of chemotaxis, a certain output (swimming behaviour) is to be expected. There seems to be no need to attribute any causal role to the organism as a whole. The mechanism does all the causal work, not the organism.

When we look at nonlinear physical systems, such as oil droplets moving through a fluid, or convection cells, that seem to portray behaviours that are at least reminiscent of agential behaviour (Horibe, Hanczyc, and Ikegami 2011), another type of non-teleological explanation is often used. It can be best described as a *statistical explanation*, which abstracts away from most causal details to reveal the explanandum to be a purely statistical phenomenon. This type of explanation was applied to thermodynamic phenomena, such as the flow of heat from hot to cold. In statistical-mechanical analysis, the molecules in a gas or liquid are fluctuating randomly, but after some time, it is likely that the faster moving molecules do not remain bunched up in one area of the container, but spread out over the whole container (either by direct displacement, or indirectly, by transferring momentum through collisions). (CITE)

This type of explanation is non-causal in the sense that it relies only on principles of combinatorics together with some boundary conditions. A uniform temperature is a vastly more likely outcome than any other since it corresponds to a much greater number of possible microstates, or ways of distributing molecular speeds among the molecules in the container (the ‘method of the most probable distribution’: Schrodinger [1946] 2013). It has also been argued that natural selection itself is this type of statistical explanation, simply reflecting facts of differential growth rates that are equally applicable to bank accounts with different interest rates as well as subgroups with different growth rates in a population (Matthen and Ariew 2002; Walsh, Lewins, and Ariew 2002).

To what extent such statistical explanations are possible of organismic behaviour is still very much an open question. There have been attempts at deriving some fundamental aspects of organismic behaviour – such as rate of bacterial replication – in this way (England 2013), but partially because the number of microstates with which organismic behaviour corresponds is so intractably large, this remains a research programme (i.e., promissory work-in-progress) rather than a set of explanatory achievements.

A notable type of explanation that adopts some semblance of the methods of statistical-thermodynamic explanations is the various optimality principles (or ‘extremal principles’)

attributed to dissipative structures and self-organizing organisms. Ilya Prigogine, a pioneer in this field, proposed the principle of minimal entropy production: i.e., systems in far-from-equilibrium conditions organize themselves as to minimize the increase of entropy (Prigogine 1947). The principle of maximal entropy production has also been proposed (Paltridge 1979).

Recently, partially through the work of Karl Friston in theoretical neuroscience, the principle of free energy minimization has been proposed: organisms, including humans, interact with their environments in such a way as to minimize free energy, which can be represented as the quantity of ‘surprise’ in an organism’s environment (Friston 2010). However, there is a crucial difference with, for instance, the second law of thermodynamics (i.e., isolated systems tend to maximize entropy), since the latter, at least in microcanonical ensembles, is actually given an explanation (in terms of the ‘method of the most probable distribution’: Schrodinger [1946] 2013). By contrast, in Friston’s theories, free energy minimization is taken as axiomatic for the theory: this is tantamount to introducing organismic purposes as axiomatic. In this way, a fully non-teleological theory of organismic behaviour is, as far as we know, still unavailable in contemporary science.<sup>6</sup>

The best candidate for non-teleological explanation of organismic behaviour is, and remains, the theory of natural selection. This basically analyses organismic behaviour as the product of a blind, algorithmic like process. Organisms exhibit seemingly goal-directed behaviour simply because this behaviour happened to allow their predecessors to reproduce more. A key assumption here is that the goal-directed behaviour is *heritable*, and so the offspring of organism with the successful behaviour B will also exhibit behaviour B. While this is true of many organisms in the natural world (e.g., social insects such as ants or wasps), it does not seem particularly explanatory of other organisms where the behaviour is not mediated by genes or other units of inheritance, but rather by organism-level systems, such as the nervous system. A complete non-teleological design explanation of organismic behaviour is also unavailable.<sup>7</sup>

So we are left with the rather peculiar situation where there is no actual scientific theory that does not posit organism-level purposes in some way, but where the philosophical problem yet concerns its possibility, or rather, the necessity of positing organism-level purposes. Because, despite the apparent explanatory indispensability of purposes (or, alternatively,

---

<sup>6</sup> Paragraph needs to be double-checked. Also, maybe too much attention given to nonlinear systems?

<sup>7</sup> Also this paragraph needs double-checking.

extremal principles that are taken as axioms), the tendency is still not to imbue them as with a mind-independent existence the way we would do with protons, or gravity.

#### 4. The Kantian Approach to Purposiveness

Kant's work on teleology can offer an interesting perspective in that he considered a closely related problem – apparently incompatible ways of viewing biological organisms – but resolved it in a way that cuts across the ontic/non-ontic dichotomy sketched above. We will thereby be focusing on an aspect of Kant that is not quite the same as often focused upon by the loose cluster of theories including 'systems biology' (Kitano 2001), 'auto-poiesis' (Varela 1979), 'self-organization' (S. A. Kauffman 1993, 2000) – all that set themselves apart from the classical framework of evolutionary theory (the Modern Synthesis).

A passage that is often quoted as particularly relevant is the following where Kant introduces the term of self-organization:

In such a product of nature each part is conceived as if it exists only **through** all the others, thus as if existing **for the sake of the others** and **on account of** the whole, i.e., as an instrument (organ), which is, however, not sufficient (for it could also be an instrument of art, and thus represented as possible at all only as an end); rather it must be thought of as an organ that **produces** the other parts (consequently each produces the others reciprocally), which cannot be the case in any instrument of art, but only of nature, which provides all the matter for instruments (even those of art): only then and on that account can such a product, as an **organized** and **self-organizing** being, be called a **natural end**. (Original emphasis. Kant [1790] 2001, 274; 5:374)

What Kant is arguing here is that organisms are not simply artefacts, where each part is designed to contribute to the whole, but where the artisan is the cause of the production and maintenance of each part of the artefact. The minute hand of a watch is produced by the artisan and not by any other part of the watch; by contrast, the various anatomical and physiological traits of an organism are produced by processes internal to the organism. Thus, an essential property of organisms is that the parts also cause the production and maintenance of the other parts.

However, this passage is most appropriate relative to the developmental perspective on organisms (as opposed to the design perspective).<sup>8</sup> However, organismic agency is not purely

---

<sup>8</sup> For a discussion of these different perspectives, and how the Kantian approach is relevant to contemporary debates in evolutionary biology, see Huneman 2017.

developmental phenomenon<sup>9</sup>, so we will be drawing on different resources in Kant's work. In particular, we will look in more detail at his general idea of purposiveness, and at his general treatment of the antinomy of teleological judgment, which concerns the apparent clash between 'mechanistic' and 'teleological' approaches to the organism.

## 2.1 The Antinomy of Teleological Judgment

In his *Critique of the Power of Judgment*, Kant posits the following two conflicting maxims concerning 'generation' (a contemporary close-equivalent: development) and 'mechanical laws' (a contemporary close-equivalent: causal processes):

*Thesis*: All generation of material things is possible in accordance with merely mechanical laws.

*Antithesis*: Some generation of such things is not possible in accordance with merely mechanical laws. (KU, AA, V: 387)<sup>10</sup>

In particular, Kant had biological organisms in mind as possible entities which are not generated merely according to mechanical laws. This thesis-antithesis pair is simply a contradiction, leading to mutually incompatible views with no prospect of reconciliation.

Kant's first step then is to make explicit that such pronouncements about the nature of reality are actually *judgements* that are necessarily relative to our *cognition* of reality. Hence he proposes the following thesis-antithesis pair:

The *first maxim* of the power of judgement is the *thesis*: All generation of material things and their forms must be judged as possible in accordance with merely mechanical laws.

The *second maxim* is the *antithesis*: Some products of material nature cannot be judged as possible according to merely mechanical laws (judging them requires an entirely different law of causality, namely that of final causes). (KU, AA, V: 387)

---

<sup>9</sup> Development typically refers to morphological changes (cell differentiation, growth, etc.) that are relatively irreversible and slow in comparison to physiological changes (metabolism) or behavioural changes (movement through space). In this way, since agency explanations target changes in 'state' of the organism, such changes may be changes in developmental state but also changes in physiological states, or in behavioural states. Furthermore, attributing agency to an organism involves attributing some causal power to the organism itself. Not all developmental changes would meet this criterion: for instance, the stunting of growth of a plant due to poor nutrition is a change of state in response to a change in the environment, but no purpose is involved, nor can the plant as such be said to be the cause of the change.

<sup>10</sup> The volume and page number refer to the Akademie edition. The translation of the original text is taken from Immanuel Kant, *Critique of Judgment*, translated by Paul Guyer and Eric Matthews (Cambridge, 2000),

This is the antinomy of teleological judgment. The motivation underlying the antithesis draws on the idea that mechanical laws do not seem to adequately account for the organization that can be found in biological organisms. In particular, Kant writes:

‘Nature, considered as a mere mechanism, could have formed itself in a thousand different ways without hitting precisely upon the unity’ (KU, AA, V: 360)

The mechanical laws do not privilege any particular organization over another; hence, if the organization were to be explained with merely mechanical laws, the organization of organisms could only be judged to be the result of *chance* (see Huneman 2014).

**In this way, Kant is relying on a concept of purposiveness that can be described as the ‘lawfulness of the contingent’.**

This reformulation of the conflict shifts the focus from different possible objective processes to different possible types of *judgment* of those processes. (In more contemporary terminology, this shifts the discussion from discussions of the actual causal nature of organisms to the *explanations* of organisms.) This transcendental turn plays a central role in Kant’s philosophy, and is worth briefly situating in his philosophy as a whole. Kant’s transcendental turn, as is well-known, was motivated as a response to Hume’s skeptical argument that no knowledge of the empirical spatio-temporal world was possible through purely analytic judgments. Thus, for example, it was never contradictory to deny that causal relations actually obtained in the spatio-temporal world. In response, Kant argued principles such as space-time, causality, subsistence and community are properties of our cognition which necessarily characterize every possible experience. Knowledge of the empirical world was possible through *synthetic a priori* judgments – judgments about the conditions of possibility of our experience. Since causality is one aspect of those conditions of possibility, when we see two billiard balls colliding and we interpret the collision as a causal relation, there is a sense in which we have no other option but to see a causal relation. Even though it is logically possible to assert that the second ball moves simply by chance when the first ball seems to touch it, that is not how our experience works – causality *constitutes* our experience.

Kant used such *a priori* principles to ground Newtonian dynamics (in the *Metaphysical Foundations of Natural Science*), and also attempted to derive the inverse-squared law of gravitation from the *a priori* cognition of three-dimensional Euclidean space. However, Kant recognized that many empirical objects could not be fully accounted for in this way. In particular, biological organisms cannot be fully explained by the *a priori* principles of causality, subsistence and community. Organisms are not merely clumps of dead matter, but seem to have properties and regularities that are not readily derivable from the principles of

mechanics. Hence the importance of the *regulative* use of reason, which did not result in *a priori* principles that constituted every possible experience, but rather in empirical laws which perhaps had less universal validity but were able to cover specific aspects of the objects under investigation. In other words, there is considerable *contingency* involved in experience beyond what is grounded in the *a priori*. Furthermore, since we are unsatisfied with the appearance of contingency, we extend the use of reason to discover empirical laws under which to subsume our contingent experiences.

The question then arises as to what *guidelines* should be used for this regulative use of reason. In particular, the teleological and mechanical mode of understanding organisms seem to be equally intuitive, but yet are in conflict. Kant gives the example of the bird, where the different anatomical parts seem to be organized in very specific ways to enable flight: ‘the structure of a bird, the hollowness of its bones, the placement of its wings for movement and of its tail for steering, etc.’ (KU, AA, V: 360). And yet, it is possible to view an organism as a clump of dead matter, obeying the laws of mechanics.<sup>11</sup>

**As indicated above**, the problem with viewing an organism *merely* as dead matter is that so much of the organism’s organization then can only be explained *by chance*. Kant rejects this chance explanation, and instead introduces purposivity as a *regulative* principle in the sense that we are justified in judging an organization to be purposive, even though we can *also* perceive the organism as composed of dead matter. Purposes are merely *as if* and are to be understood as categories of our judgments of organisms, but not a constitutive category that necessary structures all possible experience of nature.

How precisely Kant resolves the antinomy is subject to much debate in Kant scholarship, and we do not want to take sides with any one particular interpretation. We note just one in particular: Breitenbach has argued that the apparent conflict between teleological and mechanical judgments can be resolved by recognizing that each represents distinct explanatory modes with *complementary* views of nature (Breitenbach 2008). The mechanical view targets the processes that go on within an organism, whereas the teleological view allows us to pick out the organism in the first place (Breitenbach 2008: 367). The different explanatory modes thus target different explananda – different aspects of a same phenomenon – and in this way the antinomy can be resolved.

---

<sup>11</sup> Note that the judging there to be a causal relation between two objects or events (like two billiard balls colliding) is a constitutive use of reason (understanding), whereas judging according to mechanical laws is a regulative use of reason, even though mechanical laws are clearly closely related to causality as an *a priori* principle. Disentangling how precisely Kant understood the relation between causality and mechanical laws is subject to some debate in Kant scholarship. See for example Allison 1991.

## 5. Organismic Agency as a Demand of Reason

What is, we believe, a crucial aspect of the Kantian approach is that viewing organisms as natural purposes – and by extension, organismic agents – is how Kant emphasizes that it is *unavoidable* given our rational nature. To state that agential and non-agential explanations are complementary could suggest to some that there is some choice involved: a scientist could choose not to use agential explanations. Likewise, interpreting agency as an investigative heuristic or as a predictive tool likewise suggests that, once our scientific knowledge is sufficiently advanced, agential explanations are no longer needed – they are a ladder that is climbed only then to be kicked away.

Yet Kant links purposiveness to the ‘demands’ of (regulative) reason for the ‘unconditioned’, **which can be linked to the idea of purposiveness**. As Grier notes,

The demand for the unconditioned is essentially a demand for ultimate explanation, and links up with the rational prescription to secure systematic unity and completeness of knowledge. Reason, in short, is in the business of ultimately accounting for all things. (...) the demand for the unconditioned is inherent in the very nature of our reason, [and] is unavoidable and indispensably necessary... (Grier 2018)

Kant even takes this demand of reason to deliver some kind of impossibility result for the possibility of a non-purposive explanation of organismic development (and by extension, the same could be said of organismic agential-like behaviour):

it would be absurd for humans even to make such an attempt or to hope that there may yet arise a Newton who could make comprehensible even the generation of a blade of grass according to natural laws that no intention has ordered; rather, we must absolutely deny this insight to human beings. (Kant [1790] 2001, 271; V: 400)

And sometime later:

Absolutely no human reason (or even any **finite reason that is similar to ours in quality, no matter how much it exceeds it in degree**) can ever hope to understand the generation of even a little blade of grass from merely mechanical causes. (our emphasis, Kant [1790] 2001, 279; V: 410)

Some have taken this to refer to the limited cognitive capacity of human minds to represent a great multitude of causes (CITE). However, this is mistaken: in denying the insight to humans, Kant is not referring to properties of what could be called ‘evolved human nature’ but rather to our rational capacity. Any finite reason, even if it would be as computationally powerful as the

largest supercomputer, would not be able to understand organismic purposes only in terms of causal mechanisms, or statistical patterns.

Others have taken this passage as a support for a ontic view of agential explanation, where “*organisms are subjects having purposes according to values encountered in the making of their living*” (Weber and Varela 2002, 102)<sup>12</sup> This ontic interpretation of Kant goes against his overarching transcendental framework

The Kantian perspective has also been associated with projectivism (e.g. Lewens 2007: 544-5), which holds that goals and functions are in fact anthropomorphic projections onto the world. They may be useful heuristics, but do not reveal anything objectively real about the world, and teleological explanations are to be replaced by mechanistic or law-based explanations whenever possible.

Kant conceives of purposiveness as reflecting the workings (and needs) of reason confronted with a purposeless Newtonian external world, but Kant explicitly separates reason from human empirical nature. Even if our empirical nature were very different – for instance, if we had evolved very different cognitive heuristics for understanding the world – as long as we are endowed with reason then for Kant we would still employ teleological concepts. Teleology is not like a form of superstition that can be dispelled by onwards march of scientific reason; it is inherent to reason itself, and is not a ladder that can ever be kicked away.

In this way the Kantian approach cuts across the ontic/non-ontic view of agential explanations. Viewing organisms as agents is a heuristic – it allows organisms to be identified in the first place (cf. Breitenbach 2008) – but it is not merely a heuristic: it is unavoidable for a *rational understanding* of nature. Similarly, agential explanations may be predictive tools – they may accurately summarize complex patterns of behaviour, and allow us to predict how organisms will respond to environmental inputs – but they are more than mere predictive tools, because if they would be mere predictive tools, agential explanations would be replaceable by an explanation that integrates a mass of complex causal detail. Even though the latter may be predictively equivalent or even superior to an agential explanation, it does not afford *understanding* to a rational being like ourselves.

---

<sup>12</sup> Needs to be double checked whether my criticism of them is fair.

## Conclusion

The shift in contemporary biology towards the agential metaphor suggests that agential explanations warrant closer philosophical attention. Yet agential explanations are still today interpreted along the lines of the ontic/non-ontic dichotomy. In this paper we elucidated the Kantian approach in terms of the demand of reason: viewing organisms as agents is not like a form of superstition that can be dispelled by onwards march of scientific reason; it is inherent to reason itself, and is not a ladder that can ever be kicked away.

## References

- Andersen, Holly. 2016. "Complements, Not Competitors: Causal and Mathematical Explanations." *The British Journal for the Philosophy of Science*, August, axw023. <https://doi.org/10.1093/bjps/axw023>.
- Atran, Scott. 2002. *In Gods We Trust: The Evolutionary Landscape of Religion*. Oxford, UK: Oxford University Press.
- Auletta, Gennaro. 2013. "Information and Metabolism in Bacterial Chemotaxis." *Entropy* 15 (1): 311–26. <https://doi.org/10.3390/e15010311>.
- Barandiaran, Xabier E., Ezequiel Di Paolo, and Marieke Rohde. 2009. "Defining Agency: Individuality, Normativity, Asymmetry, and Spatio-Temporality in Action." *Adaptive Behavior* 17 (5): 367–86. <https://doi.org/10.1177/1059712309343819>.
- Barrett, Justin L. 2000. "Exploring the Natural Foundations of Religion." *Trends in Cognitive Sciences* 4 (1): 29–34. [https://doi.org/10.1016/S1364-6613\(99\)01419-9](https://doi.org/10.1016/S1364-6613(99)01419-9).
- Bradshaw, A.D. 1965. "Evolutionary Significance of Phenotypic Plasticity in Plants." In *Advances in Genetics*, 13:115–55. Elsevier. [https://doi.org/10.1016/S0065-2660\(08\)60048-6](https://doi.org/10.1016/S0065-2660(08)60048-6).
- Breitenbach, Angela. 2008. "Two Views on Nature: A Solution to Kant's Antinomy of Mechanism and Teleology." *British Journal for the History of Philosophy* 16 (2): 351–69. <https://doi.org/10.1080/09608780801969167>.
- Burge, Tyler. 2009. "Primitive Agency and Natural Norms\*." *Philosophy and Phenomenological Research* 79 (2): 251–78. <https://doi.org/10.1111/j.1933-1592.2009.00278.x>.
- Calvo Garzón, Paco, and Fred Keijzer. 2011. "Plants: Adaptive Behavior, Root-Brains, and Minimal Cognition." *Adaptive Behavior* 19 (3): 155–71. <https://doi.org/10.1177/1059712311409446>.
- Caporael, Linnda R., James R. Griesemer, and William C. Wimsatt. 2013. *Developing Scaffolds in Evolution, Culture, and Cognition*. MIT Press.
- Craver, Carl F. 2014. "The Ontic Account of Scientific Explanation." In *Explanation in the Special Sciences*, edited by Marie I. Kaiser, Oliver R. Scholz, Daniel Plenge, and Andreas Hüttemann, 27–52. Dordrecht: Springer Netherlands. [https://doi.org/10.1007/978-94-007-7563-3\\_2](https://doi.org/10.1007/978-94-007-7563-3_2).
- Cummins, Robert. 1975. "Functional Analysis." *The Journal of Philosophy* 72 (20): 741–65.
- Dawkins, Richard. 1976. *The Selfish Gene*. Oxford University Press.

- . 1982. *The Extended Phenotype: The Long Reach of the Gene*. Oxford University Press.
- Dennett, Daniel C. 1971. "Intentional Systems." *Journal of Philosophy* 68 (4): 87–106. <https://doi.org/10.2307/2025382>.
- . 1995. *Darwin's Dangerous Idea*. London: Penguin Books.
- Duijn, Marc van, Fred Keijzer, and Daan Franken. 2006. "Principles of Minimal Cognition: Casting Cognition as Sensorimotor Coordination." *Adaptive Behavior* 14 (2): 157–70. <https://doi.org/10.1177/105971230601400207>.
- England, Jeremy L. 2013. "Statistical Physics of Self-Replication." *The Journal of Chemical Physics* 139 (12): 121923. <https://doi.org/10.1063/1.4818538>.
- Fisher, R. A. 1919. "The Correlation between Relatives on the Supposition of Mendelian Inheritance." *Transactions of the Royal Society of Edinburgh* 52 (02): 399–433. <https://doi.org/10.1017/S0080456800012163>.
- Friston, Karl. 2010. "The Free-Energy Principle: A Unified Brain Theory?" *Nature Reviews Neuroscience* 11 (2): 127–38. <https://doi.org/10.1038/nrn2787>.
- Godfrey-Smith, Peter. 1993. "Functions: Consensus without Unity." *Pacific Philosophical Quarterly* 74 (3): 196–208. <https://doi.org/10.1111/j.1468-0114.1993.tb00358.x>.
- . 1996. *Complexity and the Function of Mind in Nature*. Cambridge Studies in Philosophy and Biology. Cambridge, UK: Cambridge University Press.
- Grafen, Alan. 1984. "Natural Selection, Kin Selection and Group Selection." In *Behavioural Ecology*, edited by John Richard Krebs and Nicholas Barry Davies, 62–84. Oxford, UK: Blackwell.
- . 2014. "The Formal Darwinism Project in Outline." *Biology & Philosophy* 29 (2): 155–74. <https://doi.org/10.1007/s10539-013-9414-y>.
- Grier, Michelle. 2018. "Kant's Critique of Metaphysics." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Summer 2018. Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/sum2018/entries/kant-metaphysics/>.
- Hamilton, Matthew B. 2009. *Population Genetics*. Chichester, UK ; Hoboken, NJ: Wiley-Blackwell.
- Hanczyc, Martin M., and Takashi Ikegami. 2010. "Chemical Basis for Minimal Cognition." *Artificial Life* 16 (3): 233–43. [https://doi.org/10.1162/artl\\_a\\_00002](https://doi.org/10.1162/artl_a_00002).
- Hempel, Carl G. 1959. "The Logic of Functional Analysis." In *Symposium on Sociological Theory*, edited by Llewellyn Gross, 271–87. New York: Harper and Row.
- Horibe, Naoto, Martin M. Hanczyc, and Takashi Ikegami. 2011. "Mode Switching and Collective Behavior in Chemical Oil Droplets." *Entropy* 13 (3): 709–19. <https://doi.org/10.3390/e13030709>.
- Huneman, Philippe. forthcoming. "Revisiting Darwinian Teleology SHPB Revised.Docx."
- . 2006. "Naturalising Purpose: From Comparative Anatomy to the 'Adventure of Reason.'" *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 37 (4): 649–74. <https://doi.org/10.1016/j.shpsc.2006.09.004>.
- . 2018. "Outlines of a Theory of Structural Explanations." *Philosophical Studies* 175 (3): 665–702. <https://doi.org/10.1007/s11098-017-0887-4>.
- Huxley, Julian. (1942) 1974. *Evolution: The Modern Synthesis*. 3rd ed. Allen and Unwin.
- Kant, Immanuel. (1790) 2001. *Critique of the Power of Judgment*. Cambridge University Press.

- Kauffman, Stuart A. 1993. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press.
- . 2000. *Investigations*. Oxford University Press.
- Kauffman, Stuart A., and Philip Clayton. 2006. "On Emergence, Agency, and Organization." *Biology and Philosophy* 21 (4): 501–21. <https://doi.org/10.1007/s10539-005-9003-9>.
- Kitano, Hiroaki, ed. 2001. *Foundations of Systems Biology*. Cambridge, Mass: MIT Press.
- Lange, M. 2013. "What Makes a Scientific Explanation Distinctively Mathematical?" *The British Journal for the Philosophy of Science* 64 (3): 485–511. <https://doi.org/10.1093/bjps/axs012>.
- Lewens, Tim. 2005. *Organisms and Artifacts: Design in Nature and Elsewhere*. MIT Press.
- . 2007. "Function." In *Handbook of the Philosophy of Science*, edited by Mohan Matthen and Christopher Stephens, 525–47. Amsterdam: Elsevier.
- Lyon, Pamela. 2015. "The Cognitive Cell: Bacterial Behavior Reconsidered." *Frontiers in Microbiology* 6 (April). <https://doi.org/10.3389/fmicb.2015.00264>.
- . 2017. "Environmental Complexity, Adaptability and Bacterial Cognition: Godfrey-Smith's Hypothesis under the Microscope." *Biology & Philosophy* 32 (3): 443–65. <https://doi.org/10.1007/s10539-017-9567-1>.
- Manneville, Paul. 2006. "Rayleigh-Bénard Convection: Thirty Years of Experimental, Theoretical, and Modeling Work." In *Dynamics of Spatio-Temporal Cellular Structures*, edited by Innocent Mutabazi, José Eduardo Wesfreid, and Etienne Guyon, 207:41–65. New York, NY: Springer New York. [https://doi.org/10.1007/978-0-387-25111-0\\_3](https://doi.org/10.1007/978-0-387-25111-0_3).
- Matthen, Mohan, and André Ariew. 2002. "Two Ways of Thinking About Fitness and Natural Selection." *Journal of Philosophy* 99 (2): 55–83. <https://doi.org/10.2307/3655552>.
- Mayr, Ernst. 1961. "Cause and Effect in Biology." *Science, New Series* 134 (3489): 1501–6.
- Millikan, Ruth Garrett. 1984. *Language, Thought, and Other Biological Categories: New Foundations for Realism*. MIT Press.
- Moreno, Alvaro, and Matteo Mossio. 2015. *Biological Autonomy*. Vol. 12. History, Philosophy and Theory of the Life Sciences. Dordrecht: Springer Netherlands. <https://doi.org/10.1007/978-94-017-9837-2>.
- Mossio, M., C. Saborido, and A. Moreno. 2009. "An Organizational Account of Biological Functions." *The British Journal for the Philosophy of Science* 60 (4): 813–41. <https://doi.org/10.1093/bjps/axp036>.
- Müller, Gerd, and Stuart Newman, eds. 2003. *Origination of Organismal Form: Beyond the Gene in Developmental and Evolutionary Biology*. The Vienna Series in Theoretical Biology. Cambridge, Mass: MIT Press.
- Neander, Karen. 1991. "Functions as Selected Effects: The Conceptual Analyst's Defense." *Philosophy of Science* 58 (2): 168–84. <https://doi.org/10.1086/289610>.
- Okasha, Samir. 2018. *Agents and Goals in Evolution*. Oxford University Press.
- Paltridge, Garth W. 1979. "Climate and Thermodynamic Systems of Maximum Dissipation." *Nature* 279 (5714): 630–31. <https://doi.org/10.1038/279630a0>.
- Pérez, Juana, Aurelio Moraleda-Muñoz, Francisco Javier Marcos-Torres, and José Muñoz-Dorado. 2016. "Bacterial Predation: 75 Years and Counting!" *Environmental Microbiology* 18 (3): 766–79. <https://doi.org/10.1111/1462-2920.13171>.
- Pittendrigh, Colin S. 1958. "Adaptation, Natural Selection, and Behavior." In *Behavior and Evolution*, edited by A. Roe and George Gaylord Simpson, 360–416. New Haven: Yale University Press.

- Prigogine, Ilya. 1947. *Étude thermodynamique des phénomènes irréversibles: (...paru dans la collection des thèses d'agrégation de l'enseignement supérieur de l'Université libre de Bruxelles)*. Dunod.
- Rice, Collin. 2015. "Moving Beyond Causes: Optimality Models and Scientific Explanation." *Nous* 49 (3): 589–615. <https://doi.org/10.1111/nous.12042>.
- Ruse, Michael. 2000. "Teleology: Yesterday, Today, and Tomorrow?" *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 31 (1): 213–32. [https://doi.org/10.1016/S1369-8486\(99\)00046-1](https://doi.org/10.1016/S1369-8486(99)00046-1).
- Schlosser, Markus. 2015. "Agency." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Fall 2015. Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/fall2015/entries/agency/>.
- Schrodinger, Erwin. (1946) 2013. *Statistical Thermodynamics*. Courier Corporation.
- Shani, Itay. 2013. "Setting the Bar for Cognitive Agency: Or, How Minimally Autonomous Can an Autonomous Agent Be?" *New Ideas in Psychology* 31 (2): 151–65. <https://doi.org/10.1016/j.newideapsych.2012.09.003>.
- Skewes, J. C., and C. A. Hooker. 2009. "Bio-Agency and the Problem of Action." *Biology & Philosophy* 24 (3): 283–300. <https://doi.org/10.1007/s10539-008-9135-9>.
- Sterelny, Kim. 2000. *The Evolution of Agency and Other Essays*. Cambridge Studies in Philosophy and Biology. Cambridge University Press.
- Strevens, Michael. 2008. *Depth: An Account of Scientific Explanation*. Cambridge, MA: Harvard University Press.
- Tversky, Amos, and Daniel Kahneman. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science* 185 (4157): 1124–31. <https://doi.org/10.1126/science.185.4157.1124>.
- Varela, Francisco J. 1979. *Principles of Biological Autonomy*. North Holland.
- Walsh, Denis. 2012. "Mechanism and Purpose: A Case for Natural Teleology." *Studies in History and Philosophy of Biological and Biomedical Sciences* 43 (1): 173–81. <https://doi.org/10.1016/j.shpsc.2011.05.016>.
- . 2015. *Organisms, Agency, and Evolution*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781316402719>.
- Walsh, Denis, Tim Lewens, and André Ariew. 2002. "The Trials of Life: Natural Selection and Random Drift." *Philosophy of Science* 69 (3): 429–46. <https://doi.org/10.1086/342454>.
- Weber, Andreas, and Francisco J. Varela. 2002. "Life after Kant: Natural Purposes and the Autopoietic Foundations of Biological Individuality." *Phenomenology and the Cognitive Sciences* 1 (2): 97–125. <https://doi.org/10.1023/A:1020368120174>.
- Woodward, James. 2003. *Making Things Happen: A Theory of Causal Explanation*. Oxford, UK: Oxford University Press.
- Wright, Larry. 1973. "Functions." *The Philosophical Review* 82 (2): 139–68.
- . 1976. *Teleological Explanations*. Berkeley, California: University of California Press.
- Zammito, John. 2006. "Teleology Then and Now: The Question of Kant's Relevance for Contemporary Controversies over Function in Biology." *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 37 (4): 748–70. <https://doi.org/10.1016/j.shpsc.2006.09.008>.