What is a Minimal Model of Cognition?

Abstract:

Active materials are self-propelled non-living entities which, in some circumstances, exhibit a number of cognitively interesting behaviors such as gradient-following, avoiding obstacles, signaling and group coordination. A live proposal across both scientific and philosophical discussions is that this may make them useful as minimal models of cognition (Hanczyc 2014, McGivern 2019). Batterman and Rice (2014) have argued that what makes a minimal model explanatory is that the model is ultimately in the same universality class as the target system, which underpins why it exhibits the same macrobehavior. We appeal to recent research in basal cognition (Lyon et al. 2021) to establish appropriate target systems and essential features of cognition as a target of modeling. Looking at self-propelled oil droplets, a type of active material, we do not find that organization alone indicates that these systems exhibit the essential features of cognition. We then examine the specific behaviors of oil droplets but also fail to find that these demonstrate the essential features of cognition. Because cognitive behaviors are not a universality class, Batterman & Rice's account of the explanatory power of minimal models simply does not apply. However, we also want to stress that it is not intended to; cognition is not the same type of behavioral phenomena as those found in physics. We then look to the minimal cognition methodology of Beer (1996, 2020) to show how active materials can be explanatorily valuable regardless of their cognitive status because they engage in specific behaviors that have traditionally been expected to involve internal representational dynamics. It is not because these model systems are also cognitive that they can be explanatory, but because they can reveal misconceptions about the cognitive underpinnings of certain, specific behaviors in target systems where such behaviors are cognitive.

Keywords: active materials, minimal cognition, minimal models, universality class, cognitive processes

1 Introduction

Active materials are self-propelled synthetic entities which, in some circumstances, exhibit a number of interesting behaviors such as gradient-following, avoiding obstacles, signaling and group coordination (Hagan and Baskaran 2016; Needleman and Dogic 2017; Horibe et al. 2011). What sets active materials apart from other synthetic materials is that they are created and studied primarily "for the targeted manipulation of [their] active behaviors" (Bursten 2020, p. 2011-12). The behaviors of these materials have been generating increasing scientific and philosophical interest because of their similarity to behaviors thought to indicate cognition in basal organisms. For instance, most readers might be familiar with chemotactic behavior in bacteria, but be unaware that analog chemotactic behaviors can be generated in some kinds of engineered oil droplets (Hanczyc 2011). Without prior knowledge that these are not living systems, we might think that they are undertaking the same kinds of activities, such as chasing sustenance or avoiding harm.

Because of the resemblances between some of active materials' behaviors and those we find in living systems, it has been suggested that these active materials can provide useful models of cognition in more complex systems (Bich and Moreno 2016; McGivern 2019; Hanczyc and Ikegami 2010). One might take these as fairly innocuous and general epistemic claims. As Godfrey-Smith (2006) would call it, *the strategy of model-based science* is to use models in order to build theories--and as active materials research is still in a very early stage, this seems highly appropriate. Godfrey-Smith's starting point is Giere's (1988) account of modeling, in which models are built for understanding the world through resemblance relations, where "[t]he modeler's strategy is to gain understanding of a complex real-world system via an understanding of a simpler, hypothetical system that resembles it in relevant respects" (Godfrey-Smith 2006, p. 726).

However, why and whether it is that we are able to gain understanding in this way has been brought under scrutiny by philosophers of science. The theory of minimal models developed by Batterman and Rice (2014) would suggest that for active materials to be minimal models of cognition, they would need to exhibit the *essential features* of the cognitive behaviors of a target system (McGivern 2019). Batterman and Rice argue that minimal models are explanatory because they produce the same universality class of macroscale behaviors by isolating the features that facilitate the behavior under consideration. To have explanatory power, then, the model need not accurately "represent" all of the microscale phenomena; it only needs to generate the same kind of macrobehavior by employing a minimal number of mechanisms or processes rather than involving all of the microscale phenomena that produce the behavior in the target system. This stands in contrast with accounts of the explanatory power of mechanistic models, which are a kind of model in which explanation involves a strict accuracy condition between the model's causal mechanism and the causal mechanism of the target phenomenon. Mechanistic models aim to describe the complete set of causal relations between components that make up a mechanism-in-the-world, and then show how said mechanism gives rise to the phenomenon of interest (Bechtel and Abrahamsen 2006; Craver 2007; Bechtel and Abrahamsen 2011). On Batterman and Rice's account, a minimal model is explanatory through replication of the macrobehavior even though the microdetails, such as specific mechanisms, may differ.

Batterman and Rice claim that in "common features accounts" of the explanatory power of minimal models, it is a common causal, topological, mathematical, or mechanistic feature that makes the model able to accurately explain the target phenomenon. They argue in response that accurately capturing a specific feature of the phenomenon is not at the basis of its explanatory power; instead, it is in virtue of the macrobehavior of both the target system and model system belonging in the same universality class. Minimal models are not concerned with isolating underlying causes under some criteria of accuracy and completeness; rather, we can identify the commonalities associated with a macroscale behavior in the target system without needing to involve many of the microdetails of the target system. Looking at models used in physics, Batterman and Rice (2014) argue that the reason these minimal models are epistemically successful is because the models themselves are a member of the same universality class (Goldenfeld 1992/2018) as the phenomena they seek to model. The universality class is the set that exhibits the phenomenon regardless of the physical details of the instantiation of that phenomenon. The essential features of the behavior belong to all systems in that same universality class–and the behavior found in the model itself shares these essential features rather than simply representing them. Using minimal models of fluid flow as an example (2014), they show that these models include a number of lattice nodes, far less than the actual number of particles that might be flowing in fluid, in order to demonstrate how fluid will flow in different environmental conditions. Both the model system and the fluid itself have the "essential features" of flow (locality, conservation, symmetry) needed in order to qualify as members of the universality class, though the model contains a minimal number of nodes (rather than all of the involved particles) to instantiate the behavior. After all, a minimal model captures the essential features of a behavior in the most economical fashion (Goldenfeld 1992). The minimal model is able to exhibit the macrobehavior precisely because it has the same essential features, thus making the model *itself* an instance of the type of thing that exhibits the behavior.

In evaluating claims that active materials might be thought of as minimal models of cognition, we think some clarification is needed on the essential features of cognition and their relation to the kinds of macrobehaviors in which active materials can engage. An oil droplet engaging in chemotactic behavior, for example, might be thought to be a minimal model of bacterial chemotaxis because it engages in the behavior though it has only a few chemical components (Hanczyc 2014; Hanczyc and Ikegami 2010). If this is a cognitive behavior in bacteria, and an active material can perform the same behavior, then we might be tempted to treat these as belonging in the same universality class of cognitive behaviors. McGivern (2019) has pointed out that if one subscribes to the Batterman & Rice view of minimal models, it leads to a dilemma in this case: "either these non-living systems are valuable as minimal models [of cognition] or they are not. If they are valuable as minimal models, then they must be instances of cognition; if they are not instances of cognition, then they are not valuable minimal models of cognition either" (McGivern 2019, p. 10). McGivern's dilemma offers a challenge for those that consider active materials to provide useful minimal models of cognition: either we must have a principle of exclusion that specifies why these minimal models are useful despite not being cognitive, or bite the bullet and take active materials to be instances of cognition.

But is the similarity between these behaviors enough to propose that active materials can act as minimal models of *cognition*? We argue that it is still an open question how we should understand the resemblance between these systems and living systems, but that it is not one that relies on both the model and target system being cognitive. Even if we take claims about the possibility that they can act as minimal models of cognition to be saying that the creation and manipulation of behaviors in active materials can offer us some understanding of the behaviors of cognitive systems, we still require some clarification on the specific relationship between these *behaviors* and cognition itself. What we argue here is that the minimal models account proposed by Batterman and Rice (2014) does not work for models of cognition. We point instead to the minimal modeling methodology of Randall Beer (1996) to show how minimal models of cognition can be explanatorily valuable without belonging in the same universality class as cognitive systems.

In order to support the first of these claims, we examine the steps from establishing a resemblance relation between cognitive behaviors and active materials to positing that they are minimal models of cognition. In the following section, we appeal to recent research in basal cognition (Lyon et al. 2021) to establish appropriate target systems and essential features of cognition. In section three, we provide an analysis of the organization of self-propelled oil droplets, but do not find that organization alone indicates that the systems exhibit the essential features of cognition. Section four then evaluates whether instead we can look at the behaviors of motile oil drops as providing minimal models of specific cognitive capacities, but we fail to find that these behaviors demonstrate the essential features of cognition. Because cognitive behaviors are not a universality class, Batterman & Rice's account of the explanatory power of minimal models simply does not apply. However, we also want to stress that it is not intended to; cognition is not the same type of behavioral phenomena as fluid flow. We then look in section five to the minimal cognition methodology of Beer (1996, 2020) to show how active materials can be minimally cognitively interesting regardless of their cognitive status because they engage in specific behaviors that have traditionally been expected to involve internal representational dynamics. It is not because these model systems are also cognitive that they can be explanatory, but that they can reveal misconceptions about the cognitive underpinnings of certain, specific behaviors in target systems where such behaviors *are* cognitive.

2 Does Cognition Have Essential Features?

In order to assess whether active materials can be minimal models of cognition, we first need a clear understanding of the essential features of cognition. This is, to put it mildly, tricky. As Pamela Lyon explains it, cognition has no agreed upon definition precisely because there is no agreement on its function (2015). Lyon points out that unlike a definition for respiration, which we can articulate from its biological function and which is precisely what enables us to do careful research on diverse implementations of respiration, definitions of cognition seem to flexibly adapt to researchers' goals, theoretical commitments, and intuitions. If cognitive science hasn't narrowed in as a community on the function of cognition, how are we to know which behaviors are indicative of cognition, let alone determine cognition's essential features?

As it stands, cognition can be defined in reference to a number of kinds of processes across several frameworks. In cases where active materials are proposed as possible minimal models of cognition, cognition is generally being used in an inclusive sense that applies to fundamental biological processes across a wide (if not the complete) range of biological individuals. For instance, recent research on *basal cognition* discusses the substrate independence of cognitive operations, which include “sensing, information processing, memory, valence, decision making, learning, anticipation, problem solving, generalization and goal directedness” (Lyon et al. 2021, p.1; see also Levin et al. 2021). Rather than defining cognition by pointing to specific underpinnings or operations (and those alone), they propose we look at cognition as serving functional roles for serving the existential needs of the organism (Lyon et al. 2021; Levin et al. 2021; Abramson and Levin 2021). Basal cognition researchers argue that we ought to think of cognition as these kinds of features of organisms which serve biological functions and which can be diversely, even non-neurally, implemented.

Here we will treat the basal approach as a means of establishing what cognition *does* while avoiding further concerns and debates around the necessary underpinnings of cognition; in other words, the *approach* to cognition comes prior to any commitments to a particular *framework* for understanding cognition. One benefit to this approach to cognition is that it enables us to investigate the workings of cognition as the integrated biological operations performed for the system in a number of ways, giving researchers an agreed-upon explanandum without a commitment to a single cognitive framework as necessary for providing explanantia. That is, researchers can discuss a wide variety of underpinnings (e.g. mechanisms, processes, sensorimotor contingencies, networks, and/or circuits) across the biological spectrum as enabling cognitive operations, giving a basis for comparison and pluralism rather than preemptively championing any particular one of these components as *the* single source of cognition. The basal cognition approach, meant to provide an anchor for investigating cognition through biological principles, also offers a three-step approach for comparison between systems: “[1] start with the smallest and simplest organisms that display the phenomenon of interest (the function, the mechanism)... [2] in those organisms identify principles from observed and measured patterns of genetic, epigenetic and behavioural interactions… [3] scale up to more complex organisms and observe where the similarities and differences actually lie, not simply where we think they must lie” (Lyon et al. 2021, p. 1).

For those approaching explanations of cognition from this angle, chemotaxis serves as a paradigmatic example of a behavior serving a biologically-based functional role for the organism. In brief, chemotaxis is directed movement up toward a gradient of chemical chemoattractants or away from chemorepellants (Webre et al. 2003). This seemingly simple behavior is said to involve a suite of capacities for bacteria: chemical sensing, a memory mechanism (for comparing current and past chemical saturation in the environment), and the ability to switch between kinds of motor activity (swimming and tumbling). In other words, chemotactic behavior involves sensorimotor coordination, which has been argued to be the primary domain of cognition (Van Duijn et al. 2006). This is well supported within the sciences: “If a brain is an organ that uses sensory information to control motor activity, then the bacterial nanobrain would fit the definition” (Webre et al. 2003). Establishing a comparative baseline along these organizational and functional lines, established through basic biological principles instead of philosophical conjecture, provides an explanatorily valuable and methodologically tractable way of investigating a behavior.

The use of cognitive terminology in the case of bacterial chemotaxis is not without serious contention, though. There is a long history of anthropocentric arguments for reserving 'cognition' to a strict subset of folk psychological phenomena for which representational explanations are said to be required (for example, see Adams 2018, Fulda 2017, Adams and Aizawa 2008, see Lyon 2015 and Sims 2021 for responses). Given the functional approach we take here, we take the onus to be on these theorists to demonstrate that (a) contentful representations exist and (b) that they carve off a unique, non-overlapping subset of processes serving a functional role from those appearing in the list above, in order to justify a label of "cognition" distinct from (and more deserving) than these other phenomena. Conversely, if the grounds for dismissing a behavior as non-cognitive are simply that they are not underpinned by the right sort of thing (representations, contentful intentionality, etc.), then the bigger problem might be that pinpointing the right kind of things doesn't actually tell us anything important about what function a cognitive behavior is achieving *for* an organism. That is, if we assume that cognitive behaviors are only those that involve certain kinds of underpinnings, and we treat the investigation of behaviors as looking for those specific underpinnings, then we ought to be concerned that any empirical analysis of those behaviors might be an exercise in reinforcement of folk psychological intuitions rather than explanation (Buzsáki 2019).

We might then think that minimal cognition instead denotes minimal instantiation dynamics or mechanisms behind cognition in certain systems. However, given that the empirical study of cognitive behaviors is focused largely on establishing the mechanisms or processes that implement sensorimotor coordination, calling any systems minimal raises a host of further questions about how to establish minimality. We will set the adjective of "minimal" aside here in relation to cognition for a few reasons. First, we don't want to imply that minimal cognition picks out a particular category, or that it exists in relation to some other "fully-fledged" category (Lyon 2020). Second, we don't want to confuse the term 'minimal cognition' with the shorthand used for Randall Beer's modeling methodology (1996), which we discuss in a further section. Third, we aren't sure how to establish that a living system achieves these behaviors in extremely simple ways. For example, the receptor and signal transduction systems in bacteria are quite complex, especially given their size. A bacterium's chemical detection network contains an estimated 8,000 receptors, made up of a combination of five different types of transmembrane proteins (Jung et al. 2018). Their signal transduction system contains two feedback cycles operating at different timescales, which are used to compare levels of chemical attractants and repellents at the receptors (Blair 1995). We might be able to say that a *model* achieves a behavior in a simple way, but we are not fully convinced that any living system undertaking a *functional behavior* does–and we have some hesitance saying so about models, depending on the type of model used. The chemical reactions involved in active materials are quite complex, as are the underlying processes involved in computational models.

With this in mind, we can propose understanding cognitive behaviors as those that are supported by sensorimotor coupling and that serve a function for the system, at least for the purposes of evaluating minimal models. These are not framework-specific commitments, as they can be cashed out in terms of computation or information-processing mechanisms, dynamically, hierarchically, or in other ways. What is important is having firm ground for doing the concrete empirical work of establishing the systems and mechanisms that support the sensorimotor activity, and understanding what that activity achieves for the system in terms of meeting existential goals (Lyon et al. 2021). This latter aspect ensures that mere motion can unproblematically be explained by appeal to external factors acting upon a system—it does not require a cognitive explanation. A feather in the wind moves but does not behave, since there is no way in which the feather contributes to being airborne in an ongoing fashion, nor does being airborne contribute to maintenance of the feather.

To summarize, rather than advocating for any specific framework for cognition, we use a functional definition of cognition as sensorimotor coupling aimed at meeting the existential needs of the system. Surveying further debates about what frameworks are best suited for establishing the exact mechanisms or processes at work, or considerations about some baseline threshold at which we can separate different kinds of cognition (or at which we can locate "real" cognition because it involves a certain kind of underpinning), are avoided.

With this understanding of the essential features of cognition, does a similarity between the behaviors of a cognitive system and the behaviors of an active material establish that the latter can be a minimal model of the former? Not yet. First, we would have to establish what kinds of behaviors qualify as those exhibiting these essential features, and what it is that makes the resemblance between a bacterium engaging in chemotactic behavior and a self-propelled oil droplet engaging in chemotactic behavior *cognitive*. Resemblance alone doesn't exactly tell us which behaviors we ought to consider cognitive, as they have to be playing a specific function for the system. In the following two sections, we evaluate what kinds of organization or activities might lead researchers to think that the activities of a kind of active material, motile oil droplets, could make them minimal models of cognition.

3 Minimal Models of Cognitive Organization?

One type of active material often considered to engage in cognitively interesting behaviors is a self-catalyzing oil droplet (Hanczyc 2011). Motile oil droplets form when oil is placed into a mixture of water and surfactant. If a Marangoni instability occurs (a flow along the interface due to an imbalance in surface tension), the droplet can form a convective cell. In repeated studies on nitrobenzene oil in a high alkaline solution (pH=12), Martin Hanzyc and others (Hanczyc and Ikegami 2010, Horibe et al. 2011, Hanczyc 2014, Čejková et al. 2017) have been able to induce a kind of chemotactic process with these motile droplets using oleic anhydrine, which converts to oleic acid on contact with water. Due to this reaction and accompanying waste dispersion, the droplets will travel through their liquid environment autonomously (Marchant 2011).

This is possible because of the self-organization of the chemicals into the convective cell, a process which Hanczyc & Ikegami (2010) compare to the autopoietic organization of living cells (Maturana and Varela 1980). For a system to be autopoietic, as all living systems are, they must be self-producing and self-maintaining, autonomously seeking out energy from their environment in order to maintain their own boundary and identity. The mutual processes of selective openness to the environment and boundary maintenance are seen by Maturana & Varela (and subsequently Hanczyc & Ikegami) as providing a good reason to believe that all living beings must be cognitive, since they are required to be selective about when to be closed to the environment (to avoid dissipation or harmful elements) and open to the environment (to take in energy), which we can observe through their sensorimotor behaviors.

Because of the convection within the cell, new fuel supplies are continuously brought to the surface of the droplet, meaning it can maintain motility until the fuel is exhausted. However, if fuel is put into the surrounding medium, the droplets can also absorb new supplies of fuel. For living systems, being able to obtain energy from the environment is a necessity for avoiding equilibrium and is a result of millions of years of co-evolution between organisms and their environment, so as not completely determined by the organism itself. Hanczyc and Ikegami point out that a similar co-determining system-environment relation is also the case for their motile droplets: "Even after the autonomous droplet emerges, it is still controlled by the environment and its own temporal changes. This is what we consider to be the congruent regularity of the droplet motion, which is the product of both droplet and environment" (Hanczyc and Ikegami 2010, p. 238).

We might think, then, that for those committed to the autopoietic framework, the self-organization and self-maintaining activity of the droplets might make them minimal models of cognition. However, for autopoietic theorists, it is not simply the organization, but the *way* it generates organism-environment relations that is key. Though autopoietic theorists tend to eschew functionalist language for various reasons (see Di Paolo et al. 2017), it is nonetheless important that the behaviors of the system be in service of self-maintenance. For example, one of the key features of an autopoietic system is its selectivity. This is why, for instance, Benard convection cells are not autopoietic, even though they are self-organizing and are often said to have an autopoietic structure (Iniguez 2001; Collier 2004). Moreover, autopoietic selectivity is guided by the existential needs of the autopoietic system: damaging features of the environment are (to the best of the system's ability) avoided, waste products are allowed passage out, and things that can be consumed for energy are allowed in. We find no such selectivity in service of self-maintenance in the behavior of motile oil droplets.

As Batterman and Rice (2014) argue, minimal models are explanatory in virtue of their ability to instantiate the essential features of a macrobehavior belonging to a universality class. A structure, though, is not a behavior. We therefore need more than structural resemblance to establish the essential features of cognition which a minimal model could instantiate. Fortunately, there are a multitude of examples of active droplets engaging in behaviors similar to those behaviors we see in cognitive systems. One of the most oft-pointed to examples is that of gradient-following, which resembles chemotaxis in bacteria. In motile oil droplets, the oleic anhydrine to acid reaction lowers the pH at the droplet/medium interface, so a droplet can follow a pH gradient, and "is therefore capable of chemotaxis as found previously only in living systems" (Hanczyc 2014, p. 1041). Because the oil droplet's motion is generated by convection internal to the system, and because of the ability of the droplet to follow a pH gradient, Hanczyc argues "that the droplet has an interface that can sense its local chemical environment and an internal convective flow acting as a motor. Therefore, the system possesses a primitive form of sensory–motor coupling" (Hanczyc 2011, p. 2886; Hanczyc & Ikegami 2010). The motor is considered to be the convection which brings new fuel to the interface, and the droplet senses local changes in pH by means of an interface imbalance. The gradient-following behavior even allows any droplets capable of following a pH gradient to find the shortest path through a maze, given that a chemoattractant has been placed at the end. The chemoattractant diffuses into the medium, and the droplet will follow the strongest gradient (shortest path) through the maze (Lagzi et al. 2010).

Now we seem to have a clear case of a behavior oftentimes associated with cognition, chemotaxis, but what determines that we ought to think of the droplet's gradient-following as falling under the universality class of *cognitive* behavior? In bacteria, the function of chemotaxis is to link the metabolic needs of a given bacterium with the external environment. Droplet chemotaxis is not likewise fulfilling any existential needs of the oil droplet system. The chemoattractant is not attractive *because* it is an energy source, but because of a chemical reaction that in no way contributes to the maintenance of the system. This point is highlighted by Bich and Moreno (Bich and Moreno 2016) in their discussion of why regulation is a necessary component of a cognitive system. They explain that while self-propelled oil droplets exhibit gradient-following behaviors, these behaviors are unregulated from the perspective of the system itself. Though the convection dynamics cause the movement, the environment alone determines the directionality of the system, which is qualitatively different from bacterial chemotaxis in a way that distances it from anything we might consider cognitive:

> "[T]he direction of droplets taxis is directly controlled by external conditions (pH gradients). It is a very interesting case of physical dynamic stability, realised through the direct coupling between the droplet dynamics and its environmental gradients. While droplets's movement is to an important degree co-determined by the environment, and its direction is governed by the external gradient, this is not the case for bacteria, which are 'intrinsically active' (Bechtel 2008) due to their capability of self-maintenance. Their movement, in fact, is internally generated by the same regime that produces and maintains the system, and it is also 'inherently goal-oriented' (Barham 2012). Using Kauffman's expression, bacteria are autonomous systems because they 'act on their own behalf' (Kauffman 2000)" (Bich and Moreno 2016, p. 16)

So, while self-propelled oil droplets exhibit a similarity to bacterial chemotaxis, which also involves gradient-following, this gradient-following does not involve meeting the existential needs of the system. Or, in more autopoietic terms, the "intrinsic goal that produces the movement is the maintenance of the organism" (Bich and Moreno 2016, p. 16) in the case of bacterial chemotaxis, but not in the case of the oil droplet. The droplets might offer minimal models of gradient-following activity, but the leap from this to being a minimal model of cognition seems like a big one.

Egbert and Di Paolo (2009) have made the broader point that a lack of behavior is also a problem for computational models of autopoiesis, which they say can self-maintain but not behave:

> "Generally the simulated autopoietic entities exist in an environment which requires no organism-scale action to continue to exist (e.g. McMullin 2004; Varela, Maturana, & Uribe, 1974). A few more recent models have demonstrated agents performing a slightly extended autopoiesis; extensions such as incorporating a simple behavior such as osmotic crisis avoidance (Ruiz-Mirazo and Mavelli 2007) or chemotaxis (Suzuki and Ikegami 2009). In these cases, the added behaviors are actually extensions of the mechanisms of autopoiesis—they are inseparable from the autopoiesis. To stop the mechanism of behavior is to stop the mechanism of autopoiesis. However, this is not the case for the majority of behaviors observed in nature that stop and start while autopoiesis continues." (Egbert and Di Paolo 2009, p. 388).

In other words, the activity that we find generated in models of autopoiesis is an extension of autopoietic processes; stop the behavior, and autopoiesis itself stops. The behaviors do not involve the kind of selectivity required for the maintenance of an autopoietic system; they are merely contingently attached to these autopoietic processes. This is not the case for living systems, for which autopoietic organization and the behavior of the system can be distinguished, though they are mutually supporting.

This is noted in the more recent work of Matthew Egbert (2021) on the self-preserving behaviors of oil droplets. In evaluating the possible metabolic basis of the activity of these droplets, he is clear to specify that it is the chemical reactions "taking place on the surface of the droplet" that cause it to follow gradients "towards conditions that facilitate or extend the life of those very same reactions" (Egbert 2021, p. 2). Where bacterial chemotaxis relies on assessment through the organism's metabolism, in which case the metabolic needs of the systems underlie its movement, the movement of the oil droplet has no whole-system regulatory aspect. An oil droplet can not be said to be undertaking metabolism-dependent movement, so Egbert specifically points to the dissipative system comprised of "(i) the chemistry on the surface of the droplet; (ii) the marangoni flow that it creates; and (iii) motility driven by the marangoni flow" as possible engaging in a viability-based response to existential needs (*ibid*, p. 4). Egbert links this to minimal forms of cognition:

> "The rate of hydrolysis increases with alkalinity, and Marangoni flow is such that the reaction drives the droplet toward local environmental conditions that accelerate that

> reaction. This is an interesting example of what has been called metabolism-based behaviour—a precarious dissipative structure which regulates its environment in response to its own metabolic health (Egbert et al. 2010[1] ; Egbert and Pérez-Mercader 2016) which has been considered a basic form of cognition…"

Elsewhere, Egbert also makes this connection, saying that motile oil droplets may be "minimal examples of *autonomous agents*--construed as precarious self-maintaining systems that act to satisfy their own needs" (Egbert 2020, p. 1). Note that because of commitments to the autopoietic framework, their definition of autonomous agency is nearly synonymous with the functional account of cognition we've provided.

We think the claim that motile oil droplets might be minimal examples of autonomous agents might be true, if, in conflict with the autopoietic account, agency and cognition are construed as separate phenomena (Brancazio & McGivern, under review[2])–but that would also be on the condition that acting to satisfy the system's needs could be purely accidental. We have no evidence that oil droplets act in a way that actually does contribute to the maintenance of the system unless the droplet is following a course explicitly set for such a purpose by an outside force. As Egbert has pointed out, the activity of motile oil droplets is generated purely by reaction at the boundary of the Marangoni effect rather than being generated by the system proper, and without aiming to fulfill the existential needs of the droplet us a whole autopoietic system. Likewise, we have no reason to think that other kinds of active materials are not also acting in ways that are purely accidentally contributing or harming their well-being.

Motile oil droplets do seem able to provide possible minimal models of specific types of behaviors that, for some systems, would be considered cognitive when they are furthering the existential goals of the system, as is the case with gradient-following and chemotaxis. This would not qualify them as minimal models in the sense Batterman and Rice propose, as they do not have explanatory power in virtue of belonging to the universality class of cognition or cognitive behaviors. Perhaps, though, rather than being minimal models of cognition *proper,* we can make the case that they provide models of some particular cognitive capacities (memory, coordination). That is, rather than looking at the relationship between behaviors in a universality class, we might still think of oil droplets as providing minimal models of specific cognitive features. We evaluate this possibility in the next section.

4. Minimal Models of Cognitive Capacities?

We have raised some concerns with the idea that active materials are minimal models of cognition in the sense proposed by Batterman and Rice (2014). Using motile oil droplets as our example, we have shown that they do not, in general, exhibit sensorimotor coordination in behaviors serving the integrated existential needs of the system. If cognition itself is not being modeled directly, this may still leave open the possibility that particular *cognitive capacities* are being modeled, such as "sensing, information processing, memory, valence, decision making, learning, anticipation, problem solving, generalization and goal directedness" to again cite the basal cognition toolkit (Lyon et al. 2021). If minimal models are expected to pick out only one particular capacity of cognitive system, then it might seem that active materials could display the essential features of just that single capacity.

For example, the motile oil droplet is hypothesized to have several ways to engage in sensory-motor coupling, where the coupling mode changes depending on the context. For droplets, these modes are flexible because there are both internal factors (convection flow pattern inertia and internal distribution of chemicals) and external factors (trails and accumulation of chemicals), both of which can be thought of as memory dynamics (Ikegami et al. 2015). As the droplets expel chemical waste, they leave trails through their medium that, as they continue to follow the chemical gradients in the system, affect their movement. This has been proposed to be a primitive kind of memory because it functions as an external device that alters the ongoing behavior of the system through chemical changes the system itself has made in the environment (Hanczyc and Ikegami 2010; Ikegami et al. 2015). These changes can affect the behavior of the system, not just its movement trajectory, by altering the modes of sensori-motor coupling which are dominant at a given time (e.g. inertial dynamics or gradient-following and avoidance of waste trails).

Let's compare this to an organism that the oil droplet might be thought of as modeling: the slime mould *Physarum Polycephalum*. Slime moulds are acellular organisms that produce extracellular trails as they traverse their environments, which alter their future movement. They avoid crossing these extracellular trails, as they mark areas in which the mould has already recently foraged (Smith-Ferguson and Beekman 2020). Endorsing Baluška & Levin's (2016) definition of memory as "experience-dependent modification of internal structure, in a stimulus specific manner that alters the way the system will respond to a stimulus in the future as a function of its past" (p. 902), Sims & Kiverstein argue that we should consider the slime mould's production of extracellular slime as also fitting this definition, though external rather than internal to the system, in a "generalised biological memory" (2022, p. 1). They justify this through reference to the extended mind hypothesis (Clark and Chalmers 1998), which proposes that we should consider external memory devices–specifically those created by the system itself–as constitutive of cognitive systems in the case that these devices make a functional contribution to the operation of that particular cognitive capacity. The internal/external divide is irrelevant to the operation of memory itself, as these differences do not matter functionally in the way that they shape the system's behavior. This is in line with researching scientists' approach to understanding the function of extracellular slime (Smith-Ferguson and Beekman 2020; Reid et al. 2012).

As we've argued above, resemblance relations alone cannot establish that the behaviors of these different systems belong in the same universality class. The behavioral similarities are clear: both systems leave trails which influence their own future movements through their environment. However, when the slime mould leaves its trails of extracellular slime, this behaviour signals something to the system that it is able to act on depending on its existential needs. Slime moulds feed via extending and contracting their body to slowly "crawl" over large areas, avoiding danger (for instance, bright sunlight or harmful chemicals like salt) and efficiently searching its environment for nutrients before moving on to a new locale (Nakagaki et al. 2004). Ruling out areas that have already been searched is necessary for the slime mould to conserve energy for further exploration; though this is described by researchers as a *choice* "because when no previously unexplored territory is available, the slime mold no longer avoids extracellular slime" (Reid et al. 2012, p. 17490). For slime moulds, then, external memory is used to conserve the energy of the system by marking previous paths, and in a way that is sensitive to (and driven by) the existential needs of the system.

The memory dynamics observed in oil droplets are behaviorally similar to the extracellular memory of the slime mould, but without being flexible to the needs of the system. We know that the oil droplet's behavior is not guided by metabolic needs, so the avoidance of trails serves no purpose for the system–it is a contingent facet of ongoing chemical reactions (Ikegami et al. 2015). The oil droplets studied by Ikegami et al. (2015) do not decouple themselves from this environmental signal when context demands it, unlike the slime mould. We can see a clear difference between a mould and a droplet here: the mould's behaviours are part of (and to some extent subordinate to) a broader scheme of self-maintenance, while the behaviour of oil droplets serves no function for the system.

Again, we want to stress that researchers working on motile oil droplets do not say that these systems are cognitive, nor do they claim that the explanatory value of using these systems as minimal models comes from belonging in the same universality class as cognitive systems (or capacities). In fact, Ikegami et al. (2015) are clear about what they see as being the fundamental differences between cognitive (living) and non-cognitive (non-living) systems: "Intentional movement does not make sense on the molecular or chemical level but on the cognitive level where we can investigate the properties of [sensori-motor contingency] selection." (p. 353). What we want to make clear is that we don't think that sensori-motor contingency selection can be thought of as cognitive, nor as minimally modeling cognition, without the system having existential needs which the selection serves to fulfill–otherwise there simply is no *selection* happening that makes use of dynamics or processes internal to the system, just forces external to the system acting upon it with differing gravity.

What we have argued so far demonstrates that these active materials cannot be thought of as minimal models of cognition in the way proposed by Batterman and Rice (2014). If these systems are in fact valuable as minimal models of cognition or of cognitive processes in some other way, then that could be justified by some other means of cashing out their explanatory power, which we will do in the following section. Before we do that, we want to propose one further possible way of understanding the claim that active materials can be minimal models of cognition, one that uses the explanatory strategy employed as its justification. After discussion of the resemblance between the behavior of motile oil droplets and the behaviors of cognitive systems, Hanczyc and Ikegami ask a set of questions about the necessity of explanations that involve positing cognition: "So the question is, how can we derive sophisticated intelligence from a merely thermodynamic system? In other words, when is it necessary to use the intentional stance (Dennett 1989) to describe a system's behavior—for example, by using sensing, or cognition, instead of reaction, or hysteresis?" (Hanczyc & Ikegami 2010, p. 233). The intentional stance derived from Dennett (1989) involves using language usually reserved for cognizing systems (perceiving, acting, thinking, believing, knowing, etc) when necessary for crafting explanations about said systems (though it need not involve any real attribution of cognition to these systems). We gain no explanatory value from describing a rock as *desiring* to roll down a hill; we use the language of the physical sciences to explain its descent. Explaining the activities of a bird, however, might require the language of *desiring* food, *perceiving* danger, and so on.

While adopting the intentional stance when thinking about active materials is useful, it is important not to read too much into this investigative maneuver. Adopting the intentional stance or using intentional language is not a response to findings, where we discover that a system is cognitive and thereafter decide to use the appropriate language when describing it. Rather, it is an acknowledgement (and codification) of a series of assumptions and (to borrow Dennett's wording) *hunches* about whether we should try using cognition-laden language to gain purchase on understanding a system. In the case of oil droplet memory trails, the use of intentional concepts like memory is clearly useful to researchers, but should not be taken to be an admission that beneath the macrobehavior lies cognition.

Further on this point, Batterman & Rice's (2014) account specifies that its explanatory target is the macrobehavior of a system, via appeal to the ineliminable underlying features that produce that behavior. When observing fluid flow it is fairly straightforward that the macrobehavior of flow is underpinned by a fluid possessing certain features. Physics has a good handle on what counts and does not count as a fluid. So, in Batterman & Rice's account, the claim that the LGA model explains fluid flow transitions into the understanding that this tells us something about fluids themselves–the LGA model allows us "to investigate and to understand the actual behavior of *real fluids"* (Batterman & Rice 2014 pg. 359, emphasis added).

Not so easy to do the same for *real cognition*. The elision from macrobehavior to underlying cognitive capacity is a trickier business. The sorts of behaviors we might be inclined to approach via the intentional stance seem quite different in this respect to Batterman & Rice's (2014) examples. While an explanation of fluid flow comfortably transitions into a better understanding of fluids (whatever the actual physical implementation of this may be), intentional behaviors are macrobehaviors without universal underlying features we can draw conclusions about. Cognition is not obviously "made of" anything in the way that systems exhibiting fluid flow are. Existential needs do not have essential components playing the same functional role wherever we find a type of behavior.

The critical problem is that it takes a commitment to a specific cognitive framework (and only that framework) to establish that there are common underlying features, and thus a universality class, for all cognition. To draw these kinds of conclusions, one must already be committed to a particular conception of cognition. The system being regarded as intentional does not speak for itself in this regard; the adoption of a particular theoretical framework and associated concepts is a necessary step that bridges the gap between observed behaviours (gradient-following, sensorimotor coupling) and explanations in terms of underlying cognitive capacities producing those behaviours. Therefore the notion that minimal models might reveal to us what cognition is by demonstrating what its minimal features are, is a non-starter. In order to move from intentional macrobehavior to cognitive capacity (or cognition), there needs to be a more substantial theoretical framework in place, specifically one that interprets specific relationships between behaviours and cognitive capacities. This could be representationalist, information-based, biosemiotic, and so on. The problem is that taking this step circumvents establishing the function of that behavior *for* the system, which we have argued above is needed to establish that the behavior is cognitive in the first place. This is because such a move preemptively limits the study of cognition to focusing on systems that satisfy requirements on what the underpinnings of cognition are assumed to be on a theoretical basis rather than focusing more practically on behaviours which serve a functional role for organisms.

We therefore conclude that Batterman & Rice's (2014) account of the explanatory justification of minimal models cannot apply to minimal models of cognition, because there simply is no universality class of either cognition or cognitive behaviors that meets the same "minimal model" criteria as fluid flow. This also dissolves McGivern's dilemma (2019): if a system is a useful minimal model of cognition (on the Batterman & Rice view), then it is also cognitive itself via the universality class; but if that system is not itself cognitive, then it is unclear why it is useful as a minimal model of cognition. However, in rejecting the universality class view of minimal model explanations of cognition, we do not reject the idea that minimal models such as active materials can be explanatorily useful. What we hope to have demonstrated so far is that cognitively-interesting behaviors are not necessarily cognitive, and that the reasoning for this need not lie in the commitment to a cognitive framework which specifies that cognition have particular underpinnings. In the section below, we will go into more detail on why it is that minimal models of cognitively interesting behaviors can still tell us a good deal about cognition, and are explanatorily justified.

5. Minimally Cognitively Interesting Agents

We have shown in the previous sections that though active materials mimic behaviors which, in some living systems, are cognitive, that does not make them part of the same universality class. According to Batterman and Rice (2014), minimal models have explanatory power in virtue of belonging in the same universality class as the target phenomena. Conversely, this means that if the behaviors of active materials are not cognitive, then it is not clear how they can offer explanatory value towards understanding cognition. However, we still think there is another way of thinking of the value of these models, one which offers explanatory value through an alternative methodological approach. We argued in section 3 that based on the basal cognition approach, a cognitive behavior involves sensorimotor coupling that functions to support the existential needs of the organism. We demonstrated that (1) active materials are not minimal models of cognition based on cognitive organization alone, that (2) active materials are not minimal models of cognition based on behavior alone, and that (3) even if organized as cognitive systems are, and even if engaging in a behavior that in organisms would be considered cognitive, an active material is not engaging in a cognitive behavior unless it is doing so in order to fulfill its existential needs; thus, there is no universality class of cognitive behaviors. We now go to Beer's methodology of minimally cognitive behaviors to demonstrate that minimal models can still help us understand and explain cognition, though these explanations are not grounded in membership in a common universality class.

The term 'minimal cognition' has often been adopted as shorthand for work on minimally cognitive agents, as popularized by Randall Beer (1996, 2020). Beer's framework originated as a response to Clark & Toribio's (1994) claim that the domain of interest to cognitive science was that which tackled "sufficiently 'representation-hungry'" problems (p. 418), or those that traditionally seemed to require positing a representational entity in their explanation (language, group coordination, memory, etc). The program developed by Beer, drawing on contemporaries like van Gelder (1995) and Brooks (1991), instead shows that minimal components can often evoke "the simplest behavior that raises cognitively interesting issues" (Beer 1996, p. 422). These minimal components often involve resources from an expanded cognitive science toolkit (dynamical systems theory, embodiment, cognitive offloading, robotics, information theory, and so on) to achieve behaviors without complex computational or representational processing within the system.

Beer's interest at the time was in developing artificial minimally cognitive agents, such as artificial (robotic) insects, and in building a mathematically based research methodology (2020). However, following Pamela Lyon (2019), he warns about the drifting usage of 'minimal cognition' from a methodological program to an ontological category. That is, the terminology was not intended to imply that the models instantiate the essential features of cognition, that the models point to the essential features of cognition in target systems, or even that there is such a possibility. Lyon's (2019) worry about this usage of the term 'minimal cognition' is that theorists take its implication to be drawing a line between the cognitive capacities of minimal systems and the cognitive capacities of more complex systems in a robust ontological sense. If this is the case, then we would be committing ourselves to the existence of a distinct category of cognition, and further, a category defined in relation to some other "fully-fledged" category. Treating minimal cognition as an ontological category gets away from the initial spirit of Beer's project, as it was not meant to define cognition in any particular way, establish boundaries, or "to propose specific criteria that demarcate the cognitive from the noncognitive" (Beer 2020, p. 3). The methodology of minimally cognitive behaviors, hereafter MMCB to avoid this confusion, is intended to provide a concrete means of investigating and predicting behavior, not establishing the ontological boundaries of what cognition is.

What makes it such that a system raises cognitively interesting issues? In Beer's initial work on the topic (1996), he demonstrates that a simple two-dimensional model agent with an eye (array of distance sensors), two motors, an "arm", and a "hand" can engage in a number of cognitively interesting activities. The agent, controlled by a dynamical neural network and simple feedforward circuits, but without any representational capacities, could orient itself around, navigate, and distinguish between objects. It could sense which gaps between objects it could fit through and which it could not. The agent could also manipulate objects and build simple structures. It was also supposed that the agent might be able to engage in simple cooperative tasks with other agents of its kind.

Beer takes these to be cognitively interesting behaviors in the sense that they warrant further investigation and explanation concerning the minimal internal dynamics or mechanisms that enable these behaviors. For example, circle/diamond discrimination in the simple 2D models leads Beer to ask "Can we identify 'circle' and 'diamond' (or 'smooth' and 'pointy') detectors in these circuits?" (1996, p. 8), where these did not involve the representational capacities argued by Clark & Toribio (1994) to be the lone resource and focal point of cognitive science (Beer 2020). As a behavior-first methodology, the idea here is, through the use of a minimal model, to create a behavior that cannot be explained without appeal to the internal dynamics of the system, however those might be cashed out. No specific class of behaviors is needed, or posited, to ground the methodology. Rather, the methodology takes its point of departure from our desire to explain certain phenomena *as* cognitive, with no commitments to what this might entail. The MMCB provides a way of investigating phenomena we already want to investigate while being quite open about what might underpin those behaviors.

This ongoing work has led to the development of a specific research program for analysing the capacities of minimally cognitively interesting systems and positing empirically testable alternatives to representational explanations within the purview of the cognitive sciences:

> "As mathematical theories, IT [information theory] and DST [dynamical systems theory] can be applied to any system that takes the proper form to meet their defining requirements; they intrinsically make no scientific claim as to "what's really going on." … we utilize the methodology of minimally cognitive behavior, in which evolutionary algorithms are used to evolve model agents that can exhibit some cognitively interesting behavior, and these evolved agents are then subjected to a mathematical analysis to understand the mechanisms underlying their behavior (Beer, 1996, 2003; Harvey, Di Paolo, Wood, Quinn, & Tuci, 2005)." (Beer and Williams 2015, p. 2)

In brief, the MMCB involves the development of a model minimal system capable of achieving some or more basic behaviors often taken to need a representational cognitive explanation. The model system is evolved, and then the minimal elements of instantiation are pinpointed. Thus, we can ask more broadly if it might instead "be most appropriate to view these circuits as merely instantiating dynamics that, when coupled to the dynamics of their bodies and environments, give rise to effective performance of the tasks for which they were selected?" (p. 8).

Importantly, the models can demonstrate what may *not* be needed in an explanation of a particular behavior. That is, they show how assumptions about common underpinnings for cognitive behaviors might be wrong, or overinflated. They allow us to narrow down and refine our understanding of the internal dynamics or mechanisms necessary for a system to engage in a behavior without necessarily adhering to any particular cognitive framework, while facilitating interpretation and comparison between competing frameworks invoked in the explanations of those behaviors. The models help "to minimize the impact of our a priori preconceptions about how an agent that successfully performs a given task ought to work, thus making the analysis of how it actually does work an interesting and insightful exercise" (Beer and Williams 2015, p. 3). In this way, these models can be explanatorily valuable by revealing ways in which cognitive science has built in assumptions about what *will* be needed in an explanation of behavior.

To reiterate, the methodology does not presuppose that any particular behaviors are cognitive, nor that there are distinct classes of cognitive behaviors. On this Beer says:

> "On the one hand, the capabilities and behavior of the model agents that we study must be rich and sophisticated enough to be cognitively interesting, so that they raise the sorts of issues that we would like to explore. …[T]hese model agents must be simple enough to be computationally and analytically tractable, so that we have some hope of evolving and analyzing them using techniques that are at most an incremental step beyond what is currently known to be feasible." (Beer 1996, p. 2)

The possible underpinnings or specifics about the internal dynamics are not pre-determined within the methodology. The methodology is thus not framework-specific in this sense; it offers resources for testing different kinds of theory-specific explanations.

The MMBC can be used not just to examine behaviors, but also to examine ideas about cognitive organization. For example, Beer has utilized minimal models, such as gliders in the Game of Life (2004), to examine and compare framework-specific explanations of cognition that have their basis in the autonomous self-maintenance of a system. Beer treats gliders (2004, 2014, 2018, 2020) as autopoietic entities, based on the biological theory of Maturana and Varela (1980) which asserts that self-production and self-individuation are the hallmarks of life and cognition. On this framework, the ability of an entity to withstand perturbations without disintegrating is possible because of the entity's autopoietic processes, and this domain of interactions is considered its "cognitive domain." Beer demonstrates that gliders offer a minimal model of this process, as they can withstand perturbations based on previous states of the system much in the same way other autopoietic entities can. For instance, Beer offers examples of two identical gliders that undergo a series of perturbations (Beer 2004). Though both start with the same structure, the series of perturbations are different for each, leaving them in different states when they encounter the next perturbation. On encountering the *same* perturbation after encountering *different* perturbations, each glider is differently prepared through modifications resulting from the first encounter, leading one to be able to withstand the second perturbation while the other is destroyed. Thus, Beer shows that "each perturbation that a unity experiences, as well as the structural changes that it undergoes even in the absence of perturbations, influences its sensitivity and response to subsequent perturbations" (Beer 2004, p. 316). Beer offers a series of such examples in order to demonstrate that gliders can instantiate many of the ideas of autopoietic theory in a concrete model which can be used to map out a single entity's cognitive domain (Beer 2014) as determined through this specific theoretical framework. Beer does not make any claims about the implications of this for what we might attribute to the model itself, but he is clear that it does not demonstrate anything along the lines of what it would be to *satisfy the essential features of cognition* (Beer 2020b, Lyon 2019). The MMCB can be used to test claims and concepts of a particular framework, but it does not tell us about the essential features of cognition–nor does that model's behaviors or organization tell us that the model itself exhibits those essential features.

Rather than taking the idea that active materials can be 'minimal models of cognition' off the table, we think Beer's methodology shows that the universality class understanding of the explanatory power of minimal models is not appropriate for *all* minimal models. In actuality, we think this is in line with what is meant by those studying active materials when they say that active materials may be used as minimal models of cognition. The MMCB shows us what interesting behaviors can be achieved without the use of representational components, for example, but further steps would be needed to say that these behaviors are indicative of the essential features of cognition. Beer has said that it would be a mistake to think that his intention was to model *cognition* proper, rather than offering a testable model of a putative cognitive phenomena (e.g. perturbation dynamics, structural coupling) (Beer 2020a, b). However, it does not provide a basis for claiming that a behavior being replicated within a model is a cognitive behavior in all systems. As we've argued above, that would require demonstrating that in each case, for each system, that behavior is functionally aimed towards meeting the existential needs of that very system (though the sensorimotor activity might be instantiated through fewer, greater, or quite different components altogether).

What we have argued might seem to be in line with what Batterman & Rice have said about establishing the essential features of a behavior: "[Minimal] models are explanatory in virtue of [there] being a story about why large classes of features are irrelevant to the explanandum phenomena" (Batterman and Rice 2014, p. 356). We don't disagree, as this would certainly be the case for a large number of phenomena that are the domain of other sciences, such as physics and chemistry. But our argument holds against the possibility of there being an explanandum phenomena in the case of cognition that instantiates as a behavior across the universality class of cognitive systems. We can investigate the minimal possible underpinnings of certain behaviors associated with some cognitive systems, such as chemotaxis, with the understanding that for a cognitive system, these behaviors will be playing a functional role for the system that requires a further explanation in terms of its underpinnings, and that mimicry of this behavior does not place the behavior in the same 'cognitive' universality class. This point is similarly made by Beer: "a behavioral characterization can often be performed in relative isolation from an account of the constitutive processes that underlie [these] behavioral dynamics…However, if we wish to understand why an agent's behavioral dynamics has the particular structure that it does or we wish to probe an agent's behavior at the limits of its viability, then we must consider at least some aspects of its underlying constitution." (Beer 2014, p. 204). Beer's work demonstrates how minimals models of cognitively interesting behaviors can lend explanatory power toward our understanding of cognition, though these minimal models do not belong in the same universality class as genuinely cognitive systems. Whether or not we want to call these models, or active materials, 'minimal models of cognition' can be decided by the reader.

6. Conclusion

We've shown that the account of the explanatory power of minimal models, as provided by Batterman and Rice (2014), does not work for cognition. On the functional understanding that cognitive behaviors involve sensorimotor activity in support of the existential needs of a system (drawn from Lyon et al. 2020), cognition is not indicated simply in the behavior's tokening of a type. This rules out the idea that cognitive behaviors have a universality class. It should be no surprise that minimal model explanations, like virtually any given explanatory strategy, have their limits. Taking a survey of historical and contemporary accounts of explanation - covering law, mechanistic, unificatory, for instance - it is often the case that there are areas where they thrive, and others where they get no traction. Though a popular option, mechanistic models may explain well in some sub-domains of biology and cognitive science, but poorly in others (Meyer 2020). Minimal model explanations are, as we see it, no exception here. While models dealing with physical and chemical, or even neuroscientific phenomena (Ross 2015) may be well-treated by a minimal model explanation per Batterman & Rice (2014), the idiosyncrasies of cognitive science preclude its use there.

However, the use of active materials as minimal models to study cognition can be explanatorily justified in ways other than through appeal to a universality class. The MMCB of Beer (1996, 2020), shows how minimal models can help explain cognition through the engineering of simple sensorimotor dynamics that facilitate theory testing and comparison of framework-specific explanations of what is happening at the microscale. Active materials, our target in this paper, similarly engage in cognitively interesting behaviors even though they are not themselves cognitive in any sense we are interested in here. What active materials represent for scientists and philosophers is an almost unique situation: systems that exhibit cognitively interesting behaviours but in the absence of cognition. Here new interventions in the spirit of Beer's work on autonomous artificial agents and Game of Life automata can (and are) being developed. Where this will take us in our understanding of cognition, and which frameworks will enter the picture to scaffold this work, remains to be seen. What is beyond doubt is that active materials provide a promising avenue for enhancing that understanding.

References:

Abramson, C.I., Levin, M., 2021. Behaviorist approaches to investigating memory and learning: A primer for synthetic biology and bioengineering. Commun. Integr. Biol. 14, 230–247.

Adams, F., 2018. Cognition wars. Stud. Hist. Philos. Sci. B Stud. Hist. Philos. Modern Phys. 68, 20–30.

Adams, F., Aizawa, K., 2008. The Bounds of Cognition. Wiley-Blackwell.

Agmon, E., and Beer, R., 2014. The Evolution and Analysis of Action Switching in Embodied Agents. Adaptive Behavior 22 (1): 3–20.

Baluška, F., Levin, M., 2016. On Having No Head: Cognition throughout Biological Systems. Front. Psychol. 7.

Barham, J., 2012. Normativity, agency, and life. Stud. Hist. Philos. Biol. Biomed. Sci. 43, 92–103.

Batterman, R.W., Rice, C.C., 2014. Minimal model explanations. Philos. Sci. 81, 349–376.

Bechtel, W., 2008. Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience. Psychology Press.

Bechtel, W., Abrahamsen, A., 2006. Phenomena and mechanisms: Putting the symbolic, connectionist, and dynamical systems debate in broader perspective. Contemporary debates in cognitive science. Oxford: Basil Blackwell.

Bechtel, W., Abrahamsen, A., 2011. Complex biological mechanisms: Cyclic, oscillatory, and autonomous. Philosophy of complex systems. Handbook of the philosophy of science 10, 257–285.

Beer, R., 1996. Toward the Evolution of Dynamical Neural Networks for Minimally Cognitive Behavior. In: Maes, P., Mataric, M., Meyer, J., Pollack, J., Wilson, S. (Eds.), From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior. MIT Press, Cambridge, MA, pp. 421–429.

Beer, R.D., 06/2004. Autopoiesis and Cognition in the Game of Life. Artif. Life 10, 309–326.

Beer, R.D., 04/2014. The Cognitive Domain of a Glider in the Game of Life. Artif. Life 20, 183–206.

Beer, R.D., 08/2020. Bittorio revisited: structural coupling in the Game of Life. Adapt. Behav. 28, 197–212.

Beer, R.D., 2018. On the Origin of Gliders. In: The 2018 Conference on Artificial Life.

Presented at the The 2018 Conference on Artificial Life, MIT Press, Tokyo, Japan, pp. 67–74.

Beer, R.D., 2020. Some historical context for minimal cognition. Adapt. Behav. 105971232093159.

Beer, R.D., Williams, P.L., 01/2015. Information Processing and Dynamics in Minimally Cognitive Agents. Cogn. Sci. 39, 1–38.

Bich, L., Moreno, A., 10/2016. The role of regulation in the origin and synthetic modelling of minimal cognition. Biosystems. 148, 12–21.

Blair, D.F., 1995. How bacteria sense and swim. Annu. Rev. Microbiol. 49, 489–522.

Brooks, R.A., 1991. Intelligence without representation. Artif. Intell. 47, 139–159.

Bursten, J.R.S., 2020. Classifying and characterizing active materials. Synthese.

Buzsáki, G., 2019. The Brain from Inside Out. Oxford University Press.

Čejková, J., Banno, T., Hanczyc, M.M., Štěpánek, F., 2017. Droplets As Liquid Robots. Artif. Life 23, 528–549.

Clark, A., Chalmers, D.J., 1998. The extended mind. Analysis 58.

Clark, A., Toribio, J., 1994. Doing without representing? Synthese 101, 401–431.

Collier, J., 2004. Self-organization, Individuation and Identity. Rev. Int. Philos. 228, 151–172.

Craver, C.F., 2007. Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience. Clarendon Press.

Dennett, D.C., 1987. The Intentional Stance. MIT Press.

Egbert, M., 2020. Marangoni Based Motile Oil-Droplets in Simulated Artificial Chemistry. In: The 2020 Conference on Artificial Life. Presented at the The 2020 Conference on Artificial Life, MIT Press, Online, pp. 260–262.

Egbert, M., 2021. Self-preserving mechanisms in motile oil droplets: a computational model of abiological self-preservation. R Soc Open Sci 8, 210534.

Egbert, M.D., Barandiaran, X.E., Di Paolo, E.A., 12/2011. Behavioral Metabolution: The Adaptive and Evolutionary Potential of Metabolism-Based Chemotaxis. Artif. Life 18, 1–25.

Egbert, M.D., Di Paolo, E., 2009. Integrating Autopoiesis and Behavior: An Exploration in

Computational Chemo-ethology. Adapt. Behav. 17, 387–401.

Egbert, M.D., Pérez-Mercader, J., 2016. Adapting to Adaptations: Behavioural Strategies that are Robust to Mutations and Other Organisational-Transformations. Sci. Rep. 6, 18963.

Fulda, F.C., 2017. Natural Agency: The Case of Bacterial Cognition. Journal of the American Philosophical Association 3, 69–90.

Gierre, R. N. 1988. Explaining Science. Univ. Chicago Press, Chicago.

Godfrey-Smith, P., 2006. The strategy of model-based science. Biology and Philosophy 21, 725–740.

Goldenfeld, N., 2018. Lectures on Phase Transitions and the Renormalization Group, 1st ed. CRC Press.

Hagan, M.F., Baskaran, A., 02/2016. Emergent self-organization in active materials. Curr. Opin. Cell Biol. 38, 74–80.

Hanczyc, M.M., 2011. Metabolism and motility in prebiotic structures. Philos. Trans. R. Soc. Lond. B Biol. Sci. 366, 2885–2893.

Hanczyc, M.M., 2014. Droplets: unconventional protocell model with life-like dynamics and room to grow. Life 4, 1038–1049.

Hanczyc, M.M., Ikegami, T., 2010. Chemical basis for minimal cognition. Artif. Life 16, 233–243.

Horibe, N., Hanczyc, M.M., Ikegami, T., 2011. Mode Switching and Collective Behavior in Chemical Oil Droplets. Entropy 13, 709–719.

Ikegami, T., Horibe, N., Hanczyc, M.M., 2015. Potential Memory Effects in Self-Moving Oil Droplets. International Journal of Unconventional Computing 11.

Iniguez, J., 2001. Rayleigh-Benard Convection: A Negentropic Approach. Science Direct Working Paper.

Jung, K., Fabiani, F., Hoyer, E., Lassak, J., 2018. Bacterial transmembrane signalling systems and their engineering for biosensing. Open Biol. 8.

Kauffman, S.A., 2000. Investigations. Oxford University Press.

Lagzi, I., Soh, S., Wesson, P.J., Browne, K.P., Grzybowski, B.A., 2010. Maze solving by chemotactic droplets. J. Am. Chem. Soc. 132, 1198–1199.

Levin, M., Keijzer, F., Lyon, P., Arendt, D., 2021. Uncovering cognitive similarities and differences, conservation and innovation. Philos. Trans. R. Soc. Lond. B Biol. Sci. 376, 20200458.

Lyon, P., 04/2020. The persistent enigma of motivation: a commentary on Nathaniel Barrett's "On the nature and origins of cognition as a form of motivated activity." Adapt. Behav. 28, 109–112.

Lyon, P., 2015. The cognitive cell: bacterial behavior reconsidered. Front. Microbiol. 6.

Lyon, P., Keijzer, F., Arendt, D., Levin, M., 2021. Reframing cognition: getting down to biological basics. Philos. Trans. R. Soc. Lond. B Biol. Sci. 376, 20190750.

Marchant, J., 2011. Oil droplets mimic early life. Nature.

Maturana, H.R., Varela, F.J., 1980. Autopoiesis and Cognition—The Realization of the Living, ser. Boston Studies on the Philosophy of Science. Dordrecht, Holland: D.

McGivern, P., 2019. Active materials: minimal models of cognition? Adapt. Behav. 105971231989174.

McMullin, B., 2004. Thirty years of computational autopoiesis: a review. Artif. Life 10, 277–295.

Meyer, R., 2020. Dynamical Causes. Biology & Philosophy 35 (5): 48.

Nakagaki, T., Kobayashi, R., Nishiura, Y., Ueda, T., 2004. Obtaining multiple separate food sources: Behavioural intelligence in the Physarum plasmodium. Proceedings of the Royal Society B: Biological Sciences 271, 2305–2310.

Needleman, D., Dogic, Z., 2017. Active matter at the interface between materials science and cell biology. Nature Reviews Materials 2, 17048.

Paolo, E.D., Buhrmann, T., Barandiaran, X., 2017. Sensorimotor Life: An Enactive Proposal. Oxford University Press, Oxford, New York.

Reid, C.R., Latty, T., Dussutour, A., Beekman, M., 2012. Slime mold uses an externalized spatial "memory" to navigate in complex environments. Proceedings of the National Academy of Sciences 109, 17490–17494.

Ross, L.N., 2015. Dynamical Models and Explanation in Neuroscience. Philos. Sci. 82, 32–54.

Ruiz-Mirazo, K., Mavelli, F., 2007. Simulation Model for Functionalized Vesicles: Lipid-Peptide Integration in Minimal Protocells. In: Advances in Artificial Life. Springer Berlin Heidelberg, pp. 32–41.

Sims, M., 2021. A continuum of intentionality: linking the biogenic and anthropogenic approaches to cognition. Biol. Philos. 36, 51.

Sims, M., Kiverstein, J., 2022. Externalized memory in slime mould and the extended (non-neuronal) mind. Cogn. Syst. Res. 73, 26–35.

Smith-Ferguson, J., Beekman, M., 2020. Who needs a brain? Slime moulds, behavioural ecology and minimal cognition. Adapt. Behav. 28, 465–478.

Suzuki, K., Ikegami, T., 2009. Shapes and self-movement in protocell systems. Artif. Life 15, 59–70.

Van Duijn, M., Keijzer, F., Franken, D., 2006. Principles of minimal cognition: Casting cognition as sensorimotor coordination. Adapt. Behav. 14, 157–170.

Van Gelder, T., Journal of Philosophy Inc., 1995. What Might Cognition Be, If Not Computation?: J. Philos. 92, 345–381.

Varela, F.G., Maturana, H.R., Uribe, R., 1974. Autopoiesis: the organization of living systems, its characterization and a model. Curr. Mod. Biol. 5, 187–196.

Webre, D.J., Wolanin, P.M., Stock, J.B., 2003. Bacterial chemotaxis. Curr. Biol. 13, R47–9.