**Modeling Multiscale Patterns:**

**Active Matter, Minimal Models, and Explanatory Autonomy**

**Citation:** Rice, C. Modeling multiscale patterns: active matter, minimal models, and explanatory autonomy. *Synthese* **200**, 432 (2022). https://doi.org/10.1007/s11229-022-03885-7

**Please Cite Published Version!**

**Abstract.** Both ecologists and statistical physicists use a variety of highly idealized models to study active matter and self-organizing critical phenomena. In this paper, I show how universality classes play a crucial role in justifying the application of highly idealized ‘minimal’ models to explain and understand the critical behaviors of active matter systems across a wide range of scales and scientific fields. Appealing to universality enables us to see why the same minimal models can be used to explain and understand behaviors across these different systems despite drastic differences in the causes and mechanisms responsible for the behaviors of interest. After analyzing these cases in detail, I argue that accounts that focus on identifying common causes or mechanisms in order to explain patterns are unable to accommodate these cases. In contrast, I argue that the justification for using these minimal models is that they are within the same universality class as real systems whose causes and mechanisms are known to be different. I also use these cases to identify several different kinds of explanatory autonomy that have important implications for how scientists ought to approach the modeling of multiscale phenomena.

**Acknowledgments:** This paper is dedicated to Margaret Morrison for all the kindness she showed me over the years and for her work that has inspired so many. Thanks to Robert Batterman, Julia Bursten, Jennifer Jhun, and three anonymous reviewers for helpful feedback on previous versions of these ideas and drafts.

**1. Introduction**

Instead of looking at logical relations between theories, several philosophers of science have recently looked to cases of modeling multiscale phenomena in order to extract insights about how we ought to think about explanation, reduction, emergence, and autonomy (Batterman 2013, 2017; Bursten 2018; Green and Batterman 2017; Morrison 2018; Winsberg 2006, 2010). In particular, these philosophers have argued that detailed analysis of the case-specific ways that scientific modelers confront multiscale phenomena can help clarify how these concepts are deployed within particular modeling contexts. Following this approach, in this paper, I use examples from scientific practice to argue that, instead of trying to model common causes or mechanisms across different systems, often modelers explain multiscale patterns by constructing minimal models that are within the same universality class as a wide range of real (and possible) systems that involve very different causes and mechanisms. A universality class is a set of systems that will display similar behaviors despite (perhaps drastic) differences in their physical features (Batterman 2002; Batterman and Rice 2014; Kadanoff 2000, 2013; Morrison 2015; Rice 2017, 2018). I will argue that often the goal of scientific modelers is to find minimal models within these universality classes whose universal behaviors are largely autonomous of the causes or mechanisms of the system—even if we consider abstract descriptions of causes and mechanisms at macroscales. In addition to illustrating something important about how certain types of patterns ought to be modeled and explained, I will argue that consideration of these cases yields important lessons for how various kinds of explanatory autonomy ought to influence the methods used to model multiscale phenomena. More specifically, I will use these cases to distinguish several different ways of conceptualizing explanatory autonomy that provide normative guidance for how scientists ought to approach the modeling of multiscale phenomena.

A particularly interesting example of attempting to model multiscale phenomena are attempts to model systems of *active matter*. In active matter systems, self-driven particles (i.e. units) continuously convert stored energy into motion, which repeatedly drives the system out of equilibrium. As a result, these systems display novel types of complex behaviors that present distinctive modeling challenges. One example are attempts to model the ways that these self-driven systems can organized themselves into critical states, what scientists have come to call *self-organized criticality* (SOC) (Frigg 2003). In these self-organized critical states, *all the scales of the system become relevant to their large-scale behaviors*. In addition to targeting phenomena that lack a characteristic scale, another striking feature of these cases is that similar types of critical behaviors have been found across systems that are extremely different in terms of their physical features. As a result, scientists from very different fields—e.g. ecology and statistical physics—*are* *often able to use the same models* to explain and understand complex behaviors. In other words, the same models are often used to model systems whose causes and mechanisms are known to be drastically different from one another. We would like to understand what makes these highly idealized models “safe” (Batterman 2017) for purposes of explanation across very different scales and fields despite their ignoring almost all of the features of the causes and mechanisms found in real systems.

In what follows, I focus my discussion on the *justifications* that ought to be given for using these extremely minimal models to explain and understand real patterns despite their ignoring virtually all of the physical features of real systems. While many philosophers have suggested that repeatable patterns ought to be explained (or understood) by identifying the common causes or mechanisms that give rise to the phenomenon across each of the systems (Craver 2006, 2007; Glennan 2017; Potochnik 2017; Strevens 2008; Woodward 2003), I will argue that such an approach is at odds with the modeling techniques used, and justifications offered, by scientists attempting to explain the patterns of behavior displayed by these active matter systems. In other words, I will argue that the approach to pattern explanation adopted by several of the most prominent accounts of explanation is unable to capture these cases.

As an alternative, I propose that the explanatory use of these minimal models is accounted for by noting that certain behaviors are universalacross classes of real, possible, and model systems whose causal and mechanistic features are drastically different. Moreover, I will argue that it is their being within these universality classes that justifies the use of these minimal models to explain and understand the complex behaviors of interest. In other words, the justification for using these minimal models to extract explanatory information is provided by showing that numerous systems with very different causes and mechanisms (across a variety of scales and fields) will all display certain universal patterns of behavior that are largely *autonomous* of the physical features of the system. This isn’t to say that scientists never explain patterns by identifying common causes or mechanisms—they often do. Nor do I intend to suggest that the universality view defended here is the only available alternative to a common causal pattern approach. Indeed, I suspect that alternative methods of using idealized models to explain patterns will be uncovered by analysis of other cases. However, the universality view I defend here is based on the justifications many scientific modelers give for using extremely minimal models to explain and understand multiscale patterns observed across systems whose physical features are extremely heterogeneous (Rice 2020, 2021). Moreover, I will argue that in these cases there is no common causal or mechanical pattern that could be cited in order to explain the occurrence of the explanandum across these very different systems. Therefore, the common causal pattern approach needs to be supplemented by the universality account (and other methods of justification uncovered by looking at additional ways scientists use models to explain multiscale patterns).

After arguing that this universality account better captures the justifications that are (and ought to be) offered in these cases, I draw on the late Margaret Morrison’s (2018) discussion of the relationship between multiscale modeling and emergence to argue that these examples allow us to distinguish various kinds of explanatory autonomy. In particular, I will argue that these cases display various ontological, epistemic, and pragmatic forms of explanatory autonomy and explore some of their normative implications for practicing multiscale modelers. The two most novel forms of explanatory autonomy introduced here are: (1) explanatory autonomy from the system’s causes or mechanisms *at any scale* and (2) explanatory autonomy from particular kinds of modeling strategies. The first kind of explanatory autonomy is demonstrated by showing that not only are microconstituents sometimes irrelevant to macroscale behaviors (as Morrison discusses), but sometimes the causes and mechanisms at all scales of the system(s) are irrelevant to the system’s universal behaviors. Consequently, I will argue that it isn’t just that these patterns need not be explained by appealing to common causes or mechanisms across each of the systems, but that there are strong reasons for thinking that they *cannot* be explained by appealing to those features. The second kind of explanatory autonomy focuses on the limitations of certain types of modeling techniques for accounting for certain types of explananda. In particular, I argue that the independence of these multiscale phenomena from the causes or mechanisms operating at any particular scale and the explanandum’s lack of a characteristic scale shows why these universal behaviors typically cannot be explained by modeling techniques that employ casual-mechanical models that are tied to a narrow range of characteristic scales.

While this paper certainly builds on the views presented by Batterman (2002), Batterman and Rice (2014), and Rice (2017, 2018, 2020, 2021), it goes beyond that work in three important ways:

(1) It provides a detailed analysis of two new cases studies that expand the set of cases that support/motivate that kind of view and clarify its scope of application.[[1]](#footnote-1) These cases show that the justifications offered by scientists studying self-organized critically are often in line with the universality account. Moreover, I will argue that there are reasons for thinking that this justification is offered by these scientists because certain patterns cannot be explained by appealing to common causes or mechanisms. That is, this paper offers further analysis of the contextual features that explain *why* these scientists adopt the universality approach.

(2) This paper more directly critiques the ways in which causal accounts of explanation *have aimed to explain patterns*. This is important since, while some of the issues for causal approaches to explanation arise solely because of their focus on causes (or accurate representation), additional issues arise due to the ‘abstraction to common macrofeatures’ idea that underlies the particular way many of these views try to capture the explanations of patterns (or regularities).

(3) This paper engages with (and expands on) Morrison (2018)’s recent discussion of the relationships between multiscale modeling and emergence by using the case studies to distinguish several different uses of ‘explanatory autonomy’ and the normative guidance those different uses provide for practicing multiscale modelers. As I mentioned above, I will specifically argue that some multiscale patterns are autonomous of the causes and mechanisms operating in the system at any scale and are explanatorily autonomous of particular types of modeling strategies. These additional types of explanatory autonomy expand the discussion beyond the relationships between micro and macro scales that are commonplace in the reduction-emergence literature.

The next two sections describe the development of two types of minimal models that are widely used to explain and understand the complex behaviors of active matter systems. Then, in Section 4, I argue that scientists’ uses of these minimal models to explain are at odds with how many philosophical accounts have suggested that patterns across different systems ought to be explained. As an alternative, I will follow the suggestions made by the scientific modelers themselves that we should instead understand these cases as attempts to discover minimal models within the same universality class as the real systems in which the behaviors of interest occur. Section 5 then distinguishes several different forms of explanatory autonomy (including some novel forms revealed by these case studies) and discusses the normative implications of those kinds of autonomy for scientists’ attempts to model multiscale phenomena. The final section concludes.

**2. Example 1: Developing a minimal model of active matter**

As I mentioned above, active matter is found in systems in which self-driven particles convert stored energy into motion. This process often gives rise to large-scale patterns of behavior that are sometimes referred to as ‘collective motion’ (Viseck and Zafeiris 2012). Common examples of such systems are flocks of birds and schools of fish, but examples have been found outside of biology as well. To physicists, these systems are particularly interesting because they display phase transitions (and other kinds of complex behaviors) when they are *far from equilibrium*. In addition, the ubiquity of these phenomena across a wide range of scales, “strongly hints at the existence of some universal features, possibly shared among the many different situations” (Ginelli 2016, 2099). This belief that there may be universal features across very different cases has led to the construction of several “minimal models…, that is models stripped of as many details as possible and only equipped with the basic features that we believe characterize the problem” (Ginelli 2016, 2100).

One of the most famous models used to understand these kinds of behaviors is the Vicsek model proposed by Vicsek et al. (1995). This model gained popularity because it is extremely minimaland yet is still able to display certain phase transitions that have been observed in a wide range of real systems (Chaté et al. 2008, 451). The original model consists of (polarized) point particles in two dimensions. These particles are self-propelled and can move in any direction at some constant velocity. The time steps of the model involve having a particle ‘reorient’ its direction according to the average direction of its neighboring particles, where the size of the particle’s neighborhood is set by what is called ‘the interaction radius’. This radius is similar to the correlation length used in other areas of physics. In short, “The only rule of the model is *at each time step a given particle driven with a constant absolute velocity assumes the average direction of motion of the particles in its neighborhood of radius r with some random perturbation added*.” (Vicsek et al. 1995, 1226, original emphasis).[[2]](#footnote-2)

Specifically, in a Vicsek model of N particles on a two-dimensional lattice of size L, the *ith* particle starts at constant speed *v*0 at a random direction of *θi*(*t*). In each of the time steps, the *ith* particlereorients to the average direction of the particles within the interaction radius *r*. As a result, the updated direction of the *ith* particle will be (Vicsek et al. 1995, 1227):

*θi*(*t* + 1) = 〈*θi*(*t*)〉*r* + Δ*θ*

Where Δ*θ*is random noise. After assuming this orientation, the position of the *ith*particle is then updated as follows:

*xi* (*t* + 1) = *xi*(*t*) + *vi*(*t*) Δ*t*

The only free parameters of the model are the density of particles, the constant magnitude (but not direction) of the velocity of the particles in each time step, and the amount of noise. Vicsek et al.’s original simulations showed that the behavior of the model is stable for a wide range of constant velocity magnitudes (0.003 < *v* < 0.3). As a result, they focused on how changing the density and noise of the system influenced the complex transport behaviors of the system. The results showed that for small densities and noise, the particles tend to form groups that move together in random directions. At higher densities and noise, the particles tend to move randomly with limited correlation. However, when the density is large and the noise is small, the overall motion of the system “becomes ordered on a macroscopic scale and all of the particles tend to *move in the same spontaneously selected direction*” (Vicsek et al. 1995, 1227). In short, the modeling results show that the system’s density of particles and noise are crucial for determining the macroscale behaviors of the system.

These modelers go on to show that the average velocity of the overall system, *Va*, can be thought of as an *order parameter* that can be used to characterize this phase transition within the system (Vicsek et al. 1995, 1228).[[3]](#footnote-3) Specifically, the absolute value of this order parameter is approximately zero in the first phase of randomly distributed particles, but the parameter is approximately one in the second phase when there is coherent collective movement of the particles. As a result, the phase transition between random and collective behavior in the system can be tracked by looking at the average (directional) velocity of the particles in the model.[[4]](#footnote-4)

 One of the crucial features of this extremely minimal model is that it is able to display phase transitions similar to those found in real active matter systems (see Figure 1).

****

Figure 1: Velocities of the particles are represented by arrows for various values of density and noise. (a) shows an initial random distribution of velocities at *t* = 0. (b) shows that for small densities and noise the particles tend to form groups that move together in random directions. (c) shows that after some time higher densities and noise results in particles that move randomly with some limited correlation. (d) shows that for high densities with small noise the motion of the particles becomes ordered. (Figure from Vicsek et al. 1995, 1227).

As Vicsek et al. put it, “In spite of its simplicity, this model results in rich, realistic dynamics, including a kinetic phase transition from no transport to finite net transport through spontaneous symmetry breaking of the rotational symmetry” (Vicsek et al. 1995, 1226). Indeed, despite its being so minimal, the model is able to reproduce the main observable patterns of interest concerning active matter systems. What is more, like phase transitions in equilibrium systems, these critical behaviors are characterized by a few *critical exponents*, *β* and *δ* ,of the equations that describe the system’s behavior in the thermodynamic limit (Vicsek et al. 1995, 1228):

*Va* ~[ηc(ρ) - η]*β* and *Va* ~ [ρ - ρc(η)]*δ*

Where ηc(ρ) and ρc(η) are the critical noise and the critical density in the limit L → ∞. In the original Vicsek model, the critical exponents are *β* = .45 ± .07 and *δ* = .35 ± .06 (Vicsek et al. 1995, 1228). Later investigations of the model showed that, for smaller densities, the system also follows a scaling law: ηc ~ ρ 1/*d*, where *d* is the dimension of the space in which the particles move (Chaté et al. 2008, 452). In other words, these modeling results not only show how the critical behaviors of the system depend on the system’s noise and density of particles, but also how the critical noise (at smaller densities) depends on the density of particles and the dimension of the space in which the particles move.

Vicsek et al. go on to argue that their minimal model “is interesting because of possible applications in a wide range of biological systems involving clustering and migration” (Vicsek et al. 1995, 1226). Later investigations showed that possible applications include modeling macromolecules, bacteria colonies, amoeba, cells, insects, fish, birds, mammals, and humans (Viseck and Zafeiris 2012, 78).[[5]](#footnote-5) What is more, the complex behaviors captured by the Vicsek model are not tied to any specific scale of the system and “there is an amazing variety of systems made of such units bridging over many orders of magnitude in size” (Vicsek and Zafeiris 2012, 73). As a result, “Collective motion is everywhere and at every scale, from herds of large mammals to amoeba and bacteria colonies, down to the cooperative behavior of molecular motors in the cell” (Chaté et al. 2008, 451). Indeed, the interest in active matter systems is largely due to these behaviors being “a ubiquitous phenomenon, observed in a wide array of different living systems and on an even wider range of scales” (Ginelli 2016, 2099). In other words, the behaviors of interest are autonomous of most of the features of the particular physical instantiation of the model’s minimal features and *lack any kind of characteristic scale at which they typically occur.* Indeed, as with critical behaviors more generally, this lack of a characteristic scale is a hallmark feature of these complex behaviors. As a recent review puts it, “There is no doubt that the studies of the emergent dynamics of real complex systems from biology, geophysics, astrophysics, economics and more… keep identifying behavior which [involve] the lack of one characteristic scale in time and space, large fluctuations and no need for specific external tuning” (Palmieri and Jensen 2020, 7).

More recent investigations have shown that the model’s wide range of application is possible because the model captures a number of features of self-driven systems that are universalin terms of being largely independent of the particular causes or mechanisms that underlie the minimal features described in the model. For example, “the source of energy making the motion possible (the ways the units gain momentum) and the conditions ensuring the similar absolute velocities *are not relevant*” (Vicsek and Zafeiris 2012, 72, my emphasis). That is, it does not matter *how* the motion is causally produced or what mechanisms ensure that the velocities of the agents align with those of their neighbors. In addition, Chaté et al. (2008) showed that, “both a topological model with a fixed number of neighbors and the metric Vicsek model belong to the same universality class, which in turn seems to be quite robust” (Chaté et al. 2008, 12139). This result leads these modelers “to conclude that the Vicsek model defines a robust universality class that is independent of the type of interactions used” (Chaté et al. 2008, 12139). In other words, investigation of these minimal models and the universality class to which they belong has also shown that most of the features of the system—and which scales of the system they are found at—are *irrelevant* to the system’s macroscale critical behaviors. This information about the irrelevance of those features and their independence of any particular scale helps scientists explain and understand why these patterns are so stable across very different physical systems.

While most of the changes to the interactions of the system fail to change the critical behaviors displayed by the Vicsek model, a few key features have been identified as essential to being a member of the Vicsek universality class: (1) spontaneous symmetry breaking to polar order, (2) some kind of self-propulsion and local alignment interactions, and (3) the conservation of the number of particles (Ginelli 2016, 2100). Changing these minimal features of the system results in different scaling behaviors near criticality; i.e. these are the key features on which the macroscale critical behaviors (the explananda) depend.[[6]](#footnote-6)

In sum, investigations of the Vicsek model (and other closely related models) have aimed to establish the range of the universality of the behaviors of various minimal models across various changes in order to justify the application of these extremely minimal models to a wide range of real phenomena that occur across very different physical systems and a wide range of scales. Indeed, as Ginelli (2016) summarizes the situation, “it is important to understand that all physical systems and models sharing the same basic features with this [universality] class, will also display the same asymptotic properties. The only way to escape this is to alter some fundamental property of the system, like changing the broken symmetry” (2100). In short, the universal critical behaviors of interest (the explananda) depend on just a few minimal features that characterize/distinguish the universality class and are independent of the ways in which those minimal features are implemented in terms of the particular scales, causes, or mechanical features of the system.

By showing that the Vicsek model is within this universality class, these modelers are then justified in using the model to make explanatory inferences about how the explanandum (i.e. the universal critical behaviors we observe) depends on the minimal features of their highly idealized models. In other words, the explanans offered in these cases will reference those few features that characterize/distinguish the systems within the universality class and are included in the minimal model (e.g. spontaneous symmetry breaking and conservation of the number of particles). Moreover, we know that all members of this universality class will display certain macroscale patterns of behavior (e.g., phase transitions). Thus, scientists can use the minimal model(s) within that universality class to learn about how those patterns of behavior depend on the minimal features of the model (e.g. the density, noise, or spatial dimension of the system).

Another crucial part of the explanation of the stability of these patterns is the demonstration that the vast majority of the other features of the system are irrelevant to displaying the explanandum. That is, in addition to showing how the pattern depends on a few minimal features, by investigating the systems and models within the universality class these modelers are also able to show why the pattern of interest is stable across very different real, possible, and model systems. This is an important part of the explanans when scientists want to explain why a certain pattern is stable *despite the very different physical features of the systems that display that pattern.*

In sum, as Vicsek and Zafeiris conclude after their extensive review of various kinds of collective behaviors and models: (1) most patterns of collective motion are universal (the same patterns occur in very different systems), (2) simple models can reproduce this behavior, and (3) these universal patterns depend on the simple features included in these minimal models (e.g. the system’s noise, density, or conservation of number of particles) (Vicsek and Zafeiris 2012, 134). In other words, there are a few universality classes that capture a wide range of systems that display the same collective patterns and those universality classes include very simple minimal models that reproduce those universal behaviors. These results are, in turn, what justifies appealing to these extremely minimal models to explain and understand various kinds of collective motion that arise in active matter systems. It is in virtue of showing us how the explanandum (i.e. the stable patterns of behavior) is dependent on changes to the minimal features that characterize the universality class and is independent of the various differences among the systems within the universality class that these modelers are able to provide an explanation for why the patterns occur across precisely this range of systems despite drastic differences in most of their physical features.[[7]](#footnote-7)

**3. Example 2: Developing minimal models of self-organized criticality**

Another interesting type of complex behavior that can arise in these self-driven systems is what has come to be called *self-organized criticality* (SOC) (Frigg 2003)*.* In 1987, Bak, Tang, and Wiesenfeld (BTW) proposed that, under very general conditions, complex systems will *organize themselves* into critical states. These self-organized critical behaviors occur when the interaction effects on a particle’s ‘nearest neighbors’ spread throughout the entire system such that all the scales of the system become relevant to the system’s overall behavior—i.e. the correlation length (or interaction radius) of the system diverges to infinity. In other words, “The ‘criticality’ in SOC refers to the case in which there are no characteristic scales…so that scaling extends from some short scale cutoff (e.g., the lattice spacing) to a cutoff set by the system size” (Carlson and Swindle 1995, 6712). Indeed, criticality is defined as a state in which the system no longer has a characteristic scale because all the scales of the system become relevant.

However, while most critical systems require some kind of *tuning* from the outside (e.g. increasing temperature), SOC systems display these critical behaviors without external manipulation or control (Bak et al. 1987; Nicolis 1989). Instead of external tuning, in SOC systems there is some kind of internal ‘driving’ process that repeatedly drives the system into a critical state. What is striking is that SOC behaviors have been found across a diverse range of scientific fields and systems:

Phenomena in very diverse fields of science have been claimed to exhibit SOC behavior. It started out with sandpiles, earthquakes, and forest fires. Next came electric breakdown, motion of magnetic flux lines in superconductors, water droplets on surfaces, dynamics of magnetic domains, and growing interfaces. The idea was soon suggested to apply to economics, and SOC models have more recently been proposed as ways of understanding biological evolution. (Jensen 1998, 2)

One of the key features of systems that will display SOC is a *separation of time scales* (Jensen 1998, 3; Watkins et al. 2017, 21). Specifically, “the threshold to instability is slowly approached under driving, and the subsequent reconfiguration is fast. Implicitly there is a separation of timescales—slow driving and fast relaxation” (Watkins et al. 2017, 7). A simple example is an earthquake in which the forces that lead to the critical behavior are built up over many years, but the release of those stresses takes place over just a few seconds.[[8]](#footnote-8) Similar time scale separations can be found in ecosystems since:

Many small, fast processes repeatedly interact to produce a larger, slower structure that constrains the behavior of the small processes in such a way that they mutually reinforce one another. Such emergent processes are self-organized. They are not created by some outside force, but rather from the mutual reinforcement of their component processes. (Peterson 2008, 20).

While these emergent patterns at larger scales are often more stable than patterns at smaller scales, when they do collapse, they often do so abruptly. That is, these self-organized systems often give rise to large-scale discontinuities that produce critical behaviors. For example, lakes can abruptly shift from a state of being oligotropic (being low nutrient, low in plant production, and relatively clear water) to a state of being eutrophic (having high nutrients, high plant production and murky water) (Peterson 2008; Scheffer and Carpenter 2003). What is more, similar types of discontinuities have been found “across all ecologically relevant scales” (Peterson 2008, 23).[[9]](#footnote-9)

As we will see below, the complex behaviors of these systems can be described by power laws and some of the exponents of those power laws are “identical for systems that appear to be different from a microscopic perspective” (Jensen 1998, 1). This discovery is crucial for two reasons: (1) power laws are typically *scale-invariant* relationships and (2) the stability of these power laws enables the same models to be applied to a range of systems that are heterogeneous in their physical features.[[10]](#footnote-10) Consequently, the challenge for these scientific modelers is to understand how similar scale-invariant power laws (and other SOC behaviors) can occur across very different types of physical systems at very different scales.

Although several models have been developed for SOC behaviors, for reasons of space, I will focus on the development of various sandpile models, which have become the paradigm type of model for SOC (Bak and Chen 1991, 46). The first thing to notice about these sandpile models is that they are *extremely* *minimal*: they include only a few features and use relatively simple updating algorithms. What is striking is that these models have been widely used to understand various SOC behaviors that have been observed across a wide range of scientific fields despite their being completely silent about the realistic causal or mechanical features of the system that are known to be relevant to the behaviors of those systems. For example, as Zinck and Grimm (2009) explain in the case of a SOC model used to study the spread of forest fires, the model “is acknowledged by landscape and fire ecologists as being ecologically significant to some degree, although it *ignores virtually all details that are discussed in more detailed forest fire models*, for example, topography, fuel and soil moisture, wind directions, weather, species composition, and individual trees” (Zinck and Grimm 2009, E177, my emphasis). Yet, despite ignoring almost all the features of the real system(s) discussed in other areas of the field, these minimal models exhibit the separation of time scales required to give rise to SOC behaviors. What is more, as we will see below, the use of these models is typically justified by showing that they are members of universality classes that include various real-world systems in which SOC has been observed. That is, the goal of these modelers is explicitly to discover minimal models and the universality classes they are members of rather than attempting to accurately model the causes or mechanisms that are known to influence these behaviors in real-world system(s).

 The original BTW sandpile model is a one-dimensional lattice in which each ‘site’ carries some number of ‘grains’ (see Figure 1). Grains are slowly added to the sites by a driving force that places one grain at a time on a randomly selected site. If the addition of this grain results in the number of grains on this site crossing a particular ‘toppling’ threshold, then one grain is moved from that site (*hi*) to its neighboring site (*hi+1*). The toppling threshold for these models is typically either a specific number of grains on the site (e.g. four grains or more) or a slope/difference between the number of grains on the site and its neighboring site (e.g. a difference of two grains).[[11]](#footnote-11) This basic updating algorithm is then cascaded throughout the rest of the sites of the model using the same threshold for each site. When multiple sites pass the toppling threshold, ‘avalanches’ occur in the model in which the effects of adding a single grain of sand to one site are carried throughout the model until none of the sites exceeds the toppling threshold. The size of these avalanches can be measured by looking at the number of units that exceed the threshold for toppling, ζ (Watkins et al. 2017, 16). When ζ is high, relatively large avalanches occur that result in greater dissipation of the grains thereby greatly reducing ζ. When ζ is smaller, the avalanches will be smaller and so they will reduce ζ less. In short, ζ is the control variable for these avalanches, but it is changed over time by the basic drive of the system that slowly adds one grain of sand at a time; i.e. no external tuning of the system is required to drive the system to criticality.



Figure 1: The original BTW sandpile model (Watkins 2017, 15).

One of the crucial features of the BTW sandpile model is that the avalanches can occur across a wide range of spatial (and temporal) scales. Indeed, some of the avalanches will be small and quick, others will spread out and take place over much longer time scales. This means that the critical behaviors (in this case avalanches) modeled here lack a characteristic spatial or temporal scale (Bak et al. 1987, 382). Another crucial feature of the model is that these avalanches occur because “the system fully stabilizes before another grain is added, so that there is a complete *separation of time scales* between addition and relaxation events. This is designed to mimic situations where the driving rate is much slower than the relaxation rate in real systems” (Carlson and Swindle 1995, 6713). That is, the key feature that this minimal model shares with real systems is a separation of time scales between the internal driving process and the dissipation/relaxation of the system.

The properties of these ‘avalanches’ can be studied numerically with the most interesting result being that the size and time distributions of the avalanches *follow power laws that are scale invariant*. In particular, the size (*s*) and time (*t*) distributions of the avalanches obey the following power laws:

 *D*(*s)* ~ *S -τ*

*D*(*t*) ~ *t* -γ

What is striking is that these power laws and the values of their critical exponents (*τ* andγ)are universal (i.e. stable) across many different ways in which the basic sandpile might be changed; e.g. using very different details for the driving or relaxation processes. In addition, investigation of these minimal models has shown that these power laws (and their critical exponents) are independent of the initial conditions of the system; i.e. where the original grains are placed.

In order to investigate the universality of these behaviors further, modelers interested in SOC behaviors later developed the Manna sandpile model (Manna 1991). The key feature of the Manna sandpile model (MSM) is that its driving process is stochastic rather than deterministic (Manna 1991, L365). This difference is due to the fact that, in the MSM, sites with two or more grains are considered to be above the toppling threshold and the model’s toppling algorithm distributes two grains to randomly selected neighboring sites. According to Manna, “In spite of these differences we believe that our model should be in the same universality class with the [BTW] sandpile model” (Manna 1991, L368). However, while the BTW model and the Manna model were originally thought to belong to the same universality class, when additional critical exponents are considered “The Manna two-state model is found to belong to a universality class of random relaxation models which is distinct from the BTW universality class” (Ben-Hur and Biham 1995, 1317).

Later, additional models were also found to be within this new Manna universality class. For example, Watkins et al. note that, “The Oslo Model is, in fact, a representative of an entire universality class…often referred to as the Manna universality class” (Watkins et al. 2017, 27). In addition, modelers have constructed ‘quasideterministic’ rotational sandpile models in which the toppling algorithm sends one grain in the direction of the recently received grain (from another site) and one grain adjacent to that grain in a clockwise or counterclockwise fashion (Lee 2017). Interestingly, despite fundamental differences in the features of the driving mechanism described by these models, several of these non-deterministic sandpile models are in the same universality class as the Manna model and display similar values of the critical exponents (*τ* andγ) within the power laws that describe their avalanches(see Table 1).



Table 1: Estimates of the critical exponents that characterize the avalanche properties of the BTW, Manna, rotational, and three step quasi-deterministic sandpile models. (Lee 2017, 6).

For example, for the first three sandpile models represented in Table 1, “The critical exponents characterizing the PDFs of avalanche properties were found to be consistent, to within statistical errors, across all models, suggesting that these three models belong to the same universality class” (Lee 2017, 7). Put differently, when multiple models all display power laws with similar slope (i.e. critical exponents) this suggests that each of the models is within the same universality class (Figure 2).



Figure 2: The probability distribution of the avalanche size for three different sandpile models: the manna sandpile model (MSM), the rotational sandpile model (RSM) and a three step quasideterministic sandpile model (3QDSM). The dashed line on each plot shows the power law fit for each of the models. The data for the 3QDSM have been shifted down and the data for the RSM shifted up for viewing purposes. The similarity of the slope of these power laws shows that their critical exponents are similar. (figure from Lee 2017, p. 3).

As was the case with the Vicsek model, what we see here is the construction of multiple highly idealized minimal models that each display some kind of critical behaviors. The task for these scientific modelers is then to use the critical exponents of the scale-invariant power laws obeyed by these models to establish the existence of various universality classes that include those minimal models and (hopefully) some real-world systems that display SOC behaviors. Then, by investigating these models, the goal is to show that various minimal features of the system are important for membership in various universality classes. The modelers then use the minimal models to determine how those minimal features are responsible for the system displaying the observed patterns of critical behavior they would like to explain. For example, investigation of these minimal models has resulted in the identification of a small set of minimal features that are characteristic of systems that display SOC behaviors (Watkins et al. 2017, 22)[[12]](#footnote-12):

1. Non-linear interaction, normally in the form of thresholds
2. Avalanching (intermittently occurring due to driving)
3. A separation of time scales

By identifying these minimal features that characterize the universality class and investigating minimal models within that class, these modelers can discover how the observed patterns of behavior depend on just a few minimal features that characterize the universality class (and is independent of other heterogeneous features of those systems). Indeed, these modelers go on to argue that “every system that simultaneously fulfills these three conditions will display SOC” (Watkins et al. 2016, 22). However, there remains some debate about whether or not the above list of features is complete (or sufficient) on its own (Watkins et al. 2016, 23). Thus, more carefully, investigations of these minimal models have shown that the above minimal conditions (cited in the explanans) make a difference to the systems’ displaying the following macroscale patterns of behavior (the SOC behaviors we would like to explain) (Watkins et al. 2017, 20):

1. Non-trivial scaling (finite size scaling; no dependence on a control parameter)

2. Spatio-temporal power law correlations

3. Apparent self-tuning to the critical point (of a possibly identified, underlying continuous order phase transition).

Finally, the hope is that by understanding the features that distinguish the universality classes these models are members of, scientists can justifiably use these minimal models to explain the critical behaviors observed in real systems that are within the same universality class. Indeed, “The primary motivation for the extensive efforts directed toward self-organizing automata is the hope that an understanding of the scaling properties in models will shed light upon similar features in real systems” (Carlson and Swindle 1995, 6712). In particular, the universal power laws displayed by these minimal models are interesting because “Similar SOC phenomena occur ubiquitously in nature, such as the crack propagation of earthquakes (Chen et al. 1991), regularity of biological extinction and punctuated equilibria (Bak and Spenppen 1993), sediment deposition, and Himalayan avalanches (Bak 1996).” (Lee 2017, 1). In other words, the power laws and critical exponents displayed by these sandpile models appear to be universal across many real systems as well. As a result, by identifying various universality classes, the minimal features that characterize (or distinguish) the systems within those classes, and identifying minimal models that satisfy those conditions, scientists can justifiably use highly idealized minimal models to infer how the macroscale patterns of behavior they observe depend on the features of the real, possible and model systems within the universality class.

**4. Two approaches to modeling patterns**

*4.1. Modeling common causes and mechanisms across different systems*

Several philosophical accounts of modeling and explanation suggest that, in order to explain a pattern across numerous systems, a scientific model ought to describe the (difference-making) causes or causal mechanisms that are common across those systems. For example, Michael Strevens adopts what he calls the ‘causal mechanical’ approach to explaining regularities, “according to which the end point for regularity explanation is, as for event explanation, a causal model representing—however abstractly—facts about fundamental-level causal influence” (Strevens 2008, 221). Strevens goes on to argue that explanations of patterns identify a causal mechanism that links some feature(s) *P,* that is common to all of the instances, with the occurrence of the phenomenon. In other words, the explanation of the pattern will just abstractly describe features that each of the individual systems have in common that are sufficient to causally produce the explanandum. For example, the explanation of ‘All ravens are black’ would cite the fact that ‘All normal ravens have *P*’ and then describe a causal mechanism that links *P* with blackness (Strevens 2008, 229). More generally, Strevens concludes that a fundamental theorem of his approach is that “The explanation of a causal generalization and the explanation of any instance of the generalization *invoke the same causal mechanism*” (Strevens 2008, 223). In other words, the goal of this approach is to (perhaps very abstractly) model the causal mechanism(s) that each of the systems have in common that gives rise to the pattern in each of them.

 In a similar way, James Woodward’s account of explanation focuses on identifying patterns of causal dependence that are *invariant* across various changes to the system. The importance of invariance for Woodward is that, “Invariance under at least one testing intervention (on variables figuring in the generalization) is necessary and sufficient for a generalization to represent a causal relationship or to figure in explanations” (Woodward 2003, 250). The point is that explanations should identify causal patterns of dependence that are stable across various interventions on (or changes to) other features of the system. Indeed, for Woodward, “an invariant relationship remains stable or unchanged as various other changes occur” (Woodward 2003, 239). This stability is what enables the causal explanation to answer various what-if-things-had-been-different questions about the (range of) stability of the causal relationships cited in the explanans. In short, patterns (or generalizations) ought to be explained by identifying causal relationships that are invariant (or stable) across different (real and possible) systems (see Woodward 2010 for more discussion).

 More recently, Angela Potochnik has argued for a causal approach to explanation that focuses on what she calls ‘causal patterns’ (Potochnik 2017). On Potochnik’s account, scientific explanations (and understanding) are provided by identifying patterns of manipulability and specifying the range of cases in which those manipulability patterns hold (Potochnik 2017, 136). As Potochnik puts it: “Causal patterns have two features that together constitute their explanatory significance: (1) they show how the phenomenon to be explained causally depends on one or more properties of the world, and (2) they indicate the *scope* of that dependence.” (Potochnik 2017, 139). In short, Potochnik’s account likewise focuses on explaining patterns or regularities by identifying a common set of causal manipulability relations across each of the systems in which the pattern occurs. The goal is to show how the phenomenon depends on the causes of the system (that are of interest to the research program) and show that those causal relationships are found in each of the cases in which the pattern of interest occurs.[[13]](#footnote-13)

Mechanistic accounts similarly suggest that in order to explain a repeated phenomenon (or generalization), a model ought to describe the features of the mechanism(s) that etiologically cause, or constitutively underlie, the explanandum in each case (Craver 2006, 2007; Kaplan and Craver 2011; Glennan 2017).[[14]](#footnote-14) For example, David Kaplan and Carl Craver (2011) have argued for what they call the ‘model-to-mechanism-mapping’ requirement which claims that:

(3M) In successful explanatory models in cognitive and systems neuroscience (a) the variables in the model correspond to components, activities, properties, and organizational features of the target mechanism that produces, maintains, or underlies the phenomenon, and (b) the (perhaps mathematical) dependencies posited among these variables in the model correspond to the (perhaps quantifiable) causal relations among the components of the target mechanism. (Kaplan and Craver 2011, 611).

Given this view of what makes mechanistic models explanatory, it is not surprising that most mechanistic discussions of how to explain recurrent patterns suggest that the model ought to describe the components and activities of the common mechanism(s) that cause or constitute the phenomenon in each of the instances of the pattern’s occurrence. For example, Craver describes cases such as using a mechanistic model of Long-Term Potentiation (LTP) to explain a generalization by describing the various conditions in which such a mechanism will produce the effect of interest (Craver 2007, 66-72). Craver then concludes that, “To say that a generalization is stable is to say that the specified relation between the cause variable and the effect variable holds under a (generally nonuniversal) range of conditions” (Craver 2007, 99). This echoes Strevens’s suggestion that an explanation of a pattern will invoke the same causal mechanism as an explanation of any instance of that pattern. The central idea is that patterns (or generalizations) are to be explained by showing that the same causal relationships (or mechanisms) hold across each of the instances in which the pattern occurs.[[15]](#footnote-15) I will refer to this general approach to explaining patterns as the *common causal pattern* (or CCP) view.

The above cases raise a number of challenges for this CCP approach. In particular, the CCP approach misses the importance of using minimal models in order to explain similar behaviors *that are produced by* *very different causes and mechanisms* across physically distinct systems. Indeed, the goal of these modelers is to account for universal power laws and critical behaviors that occur across systems that are known to be causally and mechanically heterogeneous in how they give rise to the universal behaviors of interest. As a result, these modelers construct models that deliberately ignore and idealize the causes and mechanism of real systems to an extreme degree in order to identify the minimal features (e.g. a separation of time scales) that guarantee that the system will display the pattern(s) of interest independent of the causal or mechanical features of the system that produce, underlie, or give rise to the pattern(s) in any particular case (Batterman 2002; Batterman and Rice 2014; Rice 2021).

The point here is *not* just that these minimal features and patterns occur at macroscales; i.e. these models don’t just eliminate ‘lower-level’ causal details. Instead, the key point is that the features these modelers aim to capture in their minimal models are features that are universal across fundamental changes to the causes and mechanisms operating *at any scale*. In other words, these modelers do not aim to abstractly describe some (macroscale or partial) causal mechanisms that are found in each of the real systems. Instead, these modelers aim to identify minimal features that are essential for membership in the universality class and are autonomous (i.e. independent) of the causes or mechanisms operating in the system. We can also give a clear analysis of why these modelers adopt this alternative modeling methodology. Specifically, I suggest that these modelers pursue modeling techniques that abstract away from the systems’ causes at macro, micro, and meso scales *because there is little reason to think that the features that are stable across these systems are causal or mechanistic features (at any scale).*

This raises a more general issue for attempts to ‘scale up’ causal or mechanistic accounts of explanation from the explanations of particular events to the explanations of patterns. The assumption of many CCP accounts appears to be that if we keep abstracting (e.g. to higher levels of description) we will eventually arrive at a causal pattern (or abstract causal model) that captures features of the causes or mechanisms common to each of the systems in which the pattern occurs. However, we should not assume—at least not without argument—that the extreme causal heterogeneity that we observe in natural systems can always be turned into causal similarity simply by abstracting to higher levels, introducing myriad idealizations, or eliminating various causal details. When the causes and mechanisms that produce instances of a pattern are known to be drastically different, involve entities from very different fields of science, and occur across a wide range of different scales, there is little reason to expect that abstracting away features of those causes and mechanisms will result in a causal-mechanical description that is satisfied by each of the systems in which the pattern occurs.

More generally, I contend that philosophical accounts of pattern explanation should not assume that every instance of a pattern will be generated by the same causes or mechanisms. Perhaps there will always be some kind of similarity between systems if we abstract away enough of their differences. But if all we can say is that, at some level of abstraction, all the systems will share some features, the point is rather trivial. In order for CCP accounts to be informative about how various patterns ought to be explained, it needs to be the case that the similarities between the various systems in which a pattern occurs *are causal or mechanical features* of those systems that are sufficient to explain the occurrence of the pattern. But this is precisely what cannot be assumed in many instances of using minimal models to explain universal behaviors—especially in cases where scientists know the relevant causes and mechanisms of the various systems that display a pattern are radically different, involve different components, and occur at very different spatial and temporal scales. In fact, these cases are interesting to scientific modelers primarily *because* they are patterns that are stable across systems whose causes and mechanisms are known to be very different. We need an account of how these models can be used to explain such patterns that does not assume the systems that display the pattern will always do so as a result of the same causal or mechanical relationships (or interactions).

A possible reply here is to grant that these cases cannot be captured by causal accounts of explanation, but that causal accounts never aimed to have their accounts be able to capture explanations of these kinds of patterns (or at least of their stability).[[16]](#footnote-16) While I think it is clear from the literature that many causal and mechanistic accounts have aimed to capture the explanations of patterns, I also think this reply merely highlights a dilemma for defenders of causal or mechanistic accounts of explanation. Either causal-mechanical accounts of explanation aim to capture explanations of the stability of patterns across different systems, or they don’t. If they do, then their approach to doing so (i.e. the common causal approach) fails to capture the explanations offered by scientists in these cases. Alternatively, if causal accounts don’t intend to capture these kinds of cases, then they are unable to capture a type of explanation that is central to scientific practice since scientists are often interested in explaining the universality of various patterns across different systems. Either way these cases present a problem for causal and mechanistic approaches.

Another possible response here might be to try and salvage the CCP view by simply dropping the requirement that the explanatorily relevant features of the system need to be causal or mechanistic features of the system.[[17]](#footnote-17) In other words, in line with several recent accounts that have expanded causal difference-making accounts to accommodate noncausal explanations (e.g. Bokulich 2011; Rice 2013, 2021; Saatsi and Pexton 2013), perhaps we can accommodate these cases simply by allowing that the difference-making (i.e. explanatorily relevant) features of the system need not be causal features.

While I am sympathetic with this generalized difference making approach to explanation, the problems with how the CCP approach handles these cases go beyond just its focus on causal explanations. This is because CCP views claim more than just that the difference makers for a pattern will be causal or mechanistic features. In particular, CCP approaches suggest that the best way to explain a pattern is to abstract away from the differences among the difference makers in each of the individual systems and then construct a model that accurately describes the remaining common features that are sufficient to produce the pattern in each of the systems. In other words, the CCP approach assumes that the difference makers cited in the explanation of the pattern will just be the common set of difference makers (or their common features) cited in the explanation of each instance of the pattern (but perhaps described more abstractly).

The first problem with this suggestion is that, in many cases, scientists will not be able to identify the explanatorily relevant features (e.g. those features that characterize/distinguish the universality class) just by abstracting away from the differences between the difference makers in each of the systems in which the pattern occurs. This is because the features that explain the stability of a pattern across very different systems often will not be discovered by merely abstractly describing the ways that each of those systems produce the explanandum. While a variety of features (e.g., the system’s causes and mechanisms) will make a difference within each of the individual systems—in the sense that without those features that individual system would not display the pattern of interest—it does not follow that there will be a set of features common to the difference makers within each of those systems that is sufficient to explain the occurrence of the explanandum. In contrast with the above abstraction approach, the whole point of noting the universality of these patterns is to highlight that *the same patterns can be produced in very different ways.* For example, while the minimal features that characterize the universality class are essential for determining the particular macroscale behaviors of real systems, real systems only produce those behaviors when those minimal features are combined with other features that play a role in the specific way that the system produces the macroscale behaviors (e.g. the particular causes and mechanisms that underlie the driving and relaxation processes of the system). Similarly, the minimal features that characterize the universality class are not the only features of the minimal model. Rather the minimal model displays the universal patterns of behavior because it combines those minimal features that determine the particular macroscale behaviors of the system (e.g. a separation of time scales) with other features (e.g. the model’s updating algorithms) that, when combined within the model, are sufficient for the model to produce the universal behaviors we would like to explain (e.g. SOC and phase transitions). As a result, while these minimal models do enable scientists to discover a set of features common to each of the systems that make a difference to their critical behaviors, those features typically *are not just the features that are common to the difference makers that produce each instance of the pattern.* The crucial thing to notice here is that the same feature(s) can make a difference to a particular occurrence of a pattern (in a single system) and not make a difference to displaying the pattern in general (since many other systems with very different features will still display the pattern). Moreover, a set of features might make a difference within a particular system and be insufficient for explaining why a pattern is stable across other systems with very different features. Indeed, the whole point of universality is that there are multiple different waysto produce the same behaviors and there is no requirement that those ways all share some common features that are sufficient to (causally) produce the explanandum in each of them.

Another way to see this point is to note that the features that characterize/distinguish the universality class (i.e. the explanatorily relevant features) are not discovered simply by looking at a set of models that accurately describe the difference makers for each of the individual systems and then identifying common features among them. For example, system one might display the pattern of interest due to having features A and B, whereas system two might display the pattern due to having features C and D, and system three might display the pattern due to having features E and F. This is precisely the kind of situation in which scientific modelers invoke the concept of universality; very different physical systems nonetheless display the same macroscale patterns of behavior. Now, since each of these sets of features is sufficient to produce the phenomenon in a particular case, we could explain the first instance of the pattern by citing features A and B, the second instance by citing features C and D, and the third instance by citing features E and F. However, when it comes to explaining *why the pattern is universal across these very different systems* citing these different ways that they produce (or constitute) the phenomenon is insufficient. Instead, the features that characterize the universality class and show why these very different systems all display the universal behaviors will often not be any of the features A-F, nor any partial commonalities among them, but will instead be importantly different sets of features—e.g. features G and H—that are not just abstract descriptions of the difference makers for each of the instances of the pattern. For example, consider the minimal features that characterize the universality class of the Vicsek model: (1) spontaneous symmetry breaking to polar order, (2) some kind of self-propulsion and local alignment interactions, and (3) the conservation of the number of particles. These features are not simply the common features of the difference makers that produce the behaviors within each of the systems within the universality class. Instead, they are importantly different features of the overall system that are independent of the particular ways in which the systems instantiate, realize, or produce these features. Therefore, in many cases, the difference makers cited in the explanations of particular instances of the pattern (within particular systems) will *not* be the same as the difference makers cited in the explanation of the universality of the pattern across very different systems. Thus, one of the central problems with the CCP approach is that it assumes that the difference makers for explaining a universal pattern will always be identifiable via an abstraction process that identifies the commonalities among the ways that different systems produce the pattern (i.e. discovering the causal pattern that holds across each of the systems). While these explanations certainly identify a set of difference-making and non-difference-making features for the universal behaviors, they do not do this by simply abstractly describing the common features of the difference makers that produce each instance of the pattern. Instead, these modelers investigate minimal models and universality classes to show that the particular difference makers within individual systems (e.g., the causes and mechanisms that produce the behaviors in each case) are largely irrelevant since, as long as the system has the minimal features that characterize the universality class, very different sets of features will still produce the explanandum.

Now, although several defenders of causal accounts have proposed something like the above causal abstraction process for explaining patterns, it is not a requirement that every causal account endorse this abstraction process.[[18]](#footnote-18) However, regardless of the particular methods used in attempting to discover common causal patterns, the more general issue here is that *there is no common set of difference-making causes or mechanisms to discover that is sufficient to explain the pattern’s occurrence across each of the instances.* That is, these abstraction processes fail to yield explanations of these patterns not only due to the means by which they proceed, but also because their end goal of identifying a set of common causal difference makers that can explain the pattern’s occurrence and stability is unattainable in these cases.

Another issue with merely replacing the CCP approach with a more generalized difference-making approach is that a crucial part of the explanation of these universal patterns is showing that most of the features of the systems in the universality class are *irrelevant* to displaying the explanandum (Batterman and Rice 2014). For example, a crucial piece of explanatory information uncovered by investigating the Viscek model is showing that the particular causes and mechanisms that underlie the particles’ constant velocity and the ways the particles align themselves with their neighbors are irrelevant to displaying the universal patterns of behavior at more macroscales. Merely allowing noncausal features of the system to count as difference makers fails to capture the crucial role that identifying these *non-difference-making* features plays in these explanations (Rice 2021). Indeed, simply identifying features on which these universal behaviors depend fails to explain why the pattern is universal across those systems *despite rather drastic differences in their physical features*. While this information about irrelevance isn’t required to explain every pattern, in many cases (e.g. in the cases discussed here) scientists are explicitly interested in this kind of information about which features are irrelevant to displaying the explanandum because it helps to account for the stability of patterns across very heterogeneous systems.

Therefore, beyond its focus on causes, a further issue with the CCP approach is that it assumes that the best way to try and explain a pattern is to abstract away from the details of the heterogenous ways the individual systems produce the pattern (or to more macroscales) and construct models that include the common features left after this abstraction process. While this approach will sometimes be successful, we should not assume that the explanatorily information scientists are interested in will always be discoverable simply by abstracting away from the details of individual systems and looking for common features at more macroscales.In contrast, what these cases show is that often the explanation of stable patterns across very heterogeneous systems requires modeling techniques that systematically identify minimal features that characterize different universality classes and show why most of the systems’ physical features (across multiple scales) are irrelevant to displaying the pattern of interest.

Considering the kinds of minimal features cited in these explanations reveals another reason why CCP approaches will typically be inadequate in the cases described above. One of the main motivations for the use of these minimal models is that the phenomena of interest *lack any characteristic scale.* In the case of self-organized criticality, this isbecause all the scales of the system become relevant near criticality; i.e. the system’s correlation length diverges to infinity*.* However, given that most causal or mechanistic models require there to be a specific set of scales at which the relevant features of the process(es) occur (what is called the ‘characteristic scale’ of the causal processes), such models will be ill-equipped to handle phenomena that lack any characteristic scale. This is where the tyranny of scales becomes problematic (Batterman 2013; Green and Batterman 2017; Oden 2006). The problem is that, while most of the causal and mechanistic modeling techniques available have been designed to capture particular kinds of processes that occur at some characteristic scale(s) of the system, in these cases of modeling critical behaviors, all the scales of the system become relevant to the macroscale behaviors of interest. Indeed, as Sara Green and Robert Batterman note, “multi-scale modeling in both physics and biology show that modelers in both domains must confront the tyranny of scales problem. There is no single approach that can account for all relevant aspects of multi-scale systems” (2017, 32). Therefore, even though there can be causes and mechanisms at multiple different scales of the system, it is difficult to see how in practice one would construct a causal or mechanistic model that would incorporate *all* the scales of the system and all the relevant interactions between them as the correlation length diverges near criticality. Put differently, since the explanatorily relevant features are not part of the systems’ causes/mechanism and lack a characteristic scale, neither identifying common features of the systems’ causes/mechanisms nor abstracting to more macroscales will reveal the explanatory features uncovered by investigating these minimal models and the universality classes to which they belong. Consequently, rather than attempting to identify the appropriate scale(s) (or ‘levels’) at which the explanatorily relevant causes occur, these modelers aim to identify *scale-invariant* relationships that look the same across a wide range of scales. Put differently, while it is possible for there to be causal patterns that involve features across a wide range of scales of the system, it is often extremely difficult for scientific modelers to construct a causal or mechanistic modelof the phenomena because most of the available causal and mechanistic modeling approaches are designed to represent causal factors (or processes) at a relatively narrow range of scales. Therefore, another issue with the CCP approach is that attempting to build causal-mechanical models that capture causal patterns that characteristically occur at some scale(s) of the system will often fail to identify the kinds of scale-invariantproperties that are essential to explaining and understanding the stability of these universal behaviors.

In sum, in studying active matter and SOC behaviors, the CCP approach is largely unhelpful for three main reasons: (1) the patterns of interest range across systems that are known to be extremely different with respect to their relevant causes and mechanisms, (2) we cannot assume that causal (or difference making) heterogeneity will become causal (or difference making) similarity at higher levels of abstraction, and (3) many of the phenomena of interest have no characteristic scale(s) at which the explanatorily relevant features occur. Consequently, it appears that CCP approaches to pattern explanation will be unable to accommodate these cases of modeling active matter and SOC behaviors.

*4.2. Minimal models and universality classes*

As an alternative to CCP views, I contend that the above cases suggest that we develop an account of how to justify the use of minimal models to explain and understand patterns that deliberately appeals to universality classes and the multiscale features that characterize them rather than the representation (or description) of common causes or mechanisms. Instead of aiming to represent causes or mechanisms common to each of the instances of the pattern, these minimal models aim to identify scale-invariant power laws whose universal critical exponents become the stable features of interest. The critical exponents of these power laws are then used to distinguish/characterize various universality classes that contain various minimal models. The next step is to determine whether various real systems that display the behaviors of interest are also in the same universality classes as these minimal models (i.e. whether they display power laws with similar exponents). If so, then scientists are justified in inferring that the dependence relations that hold between the minimal features of the model (e.g. a separation of time scales and conservation of momentum) and its patterns of behavior will also hold in the real systems within that universality class.

The important thing to note about these appeals to universality classes is that they focus on stability across different systems *without* having to assume that the stability is the result of common causes or mechanisms. This is why *appeals to universality classes are not the same as appeals to causal patterns*. Indeed, as the cases described above illustrate, the most striking cases of universality are those that occur across systems (e.g. fluids and magnets) that are known to have very different physical features and causal patterns. As Kadanoff (2013) describes it, “Whenever two systems show an unexpected or deeply rooted identity of behavior they are said to be in the same universality class” (178). In other words, it is precisely when the systems are *not* expected to display similar behaviors—e.g. because they are known to involve very different physical causes and mechanisms—that the concept of a universality class is most useful. In short, universality classes can capture stable patterns of behavior that range over systems with different causal patterns (or causal relationships).

I argue that the goal of the above instances of minimal modeling is to explain and understand the patterns of active matter and SOC behaviors that we observe, not by attempting to uncover common causal patterns or mechanism, but by identifying certain minimal conditions of the system required to be a member of the universality class. The importance of these minimal conditions is that they are largely *autonomous* of the causes and mechanisms (or spatial or temporal scales) of particular systems. This is what enables the minimal models to display the same universal patterns as real-world systems despite ignoring and idealizing virtually all of the causes or mechanisms that are responsible for those behaviors in any real-world case. It is also what enables the same minimal models to be applicable for understanding patterns in fields as different as statistical physics and ecology and across very different scales. That is, the reason these minimal models can be employed in the explanation of universal patterns across such different systems is because those patterns are (largely) autonomous of the particular causal-mechanical features (or scale) of the system. As a result, these models completely ignore those causal-mechanical features in favor of identifying minimal features such as a separation of time scales between internal driving and system-wide relaxation.

Given this difference in modeling goals, we require a very different kind of justificatory story about why such highly idealized minimal models can be used to explain and understand these patterns. As evidence for the need for this alternative view, some modelers of SOC behaviors draw an explicit distinction between two approaches to modeling the physics of the system:

Practically, one can consider two approaches to the physics of such a system. One is to address the detailed plasma physics of an energy release event in isolation. The relevant questions are then: what is the cause of the onset of the instability? What is the specific instability mechanism? ...This approach is best suited to physics which occurs on one specific spatio-temporal scale. (Watkins et al. 2017, 7).

In other words, one approach would be to try and uncover the causes and mechanisms that give rise to these behaviors in particular cases by building a model that would describe how the causal-mechanical features of real systems give rise to the critical behaviors. As Watkins et al. point out, such an approach is best suited to cases where the relevant physics occurs at a specific spatiotemporal scale of the system. The problem is that, in instances of phase transitions in active matter and instances of SOC, all of the scales of the system are relevant and there is no ‘characteristic scale’ at which the causes or mechanisms that produce the behavior are found. As a result, these modelers go on to suggest that the above approach is intractable when attempting to model these behaviors. In other words, it isn’t just that these modelers happen to adopt a different approach, but that the nature of the multiscale phenomena they are interested in modeling makes the CCP strategy impossible (or, at least, largely unhelpful). Thus, they propose an alternative modeling approach:

Alternatively, one can consider the situation where the physics on all scales is equally important, and is, furthermore, strongly coupled…In this case the “bottom up” approach of solving for the dynamics of individual events is intractable. Instead, one can look to the success of the renormalization group approach (Wilson 1971, 1979) in critical phenomena. ... Central to the structure of such a model is self-similar scaling (the system looks the same on all scales subject to a rescaling), leading to power law distributions of (event) sizes and power law (long-range) correlations as the key observable. Importantly, *a broad range of different detailed, microscopic interactions, on coarse-graining, lead to the same collective behavior*, *thus one expects the same essential phenomenology to be ubiquitous*. (Watkins et al. 2017, 8, my emphasis).

In short, attempting to model the causal and mechanistic interactions of these systems will be intractable when the phenomenon involves all of the strongly coupled scales of the system. Fortunately, the concept of universality shows us how mathematical modeling techniques can enable scientists to explain stable collective behaviors (i.e. the same essential phenomenology) across a broad range of systems with very different causal and mechanistic interactions.[[19]](#footnote-19) Once these stable patterns are identified, the alternative modeling approach aims to construct minimal models in which the objects of interest are self-similar scaling and the critical exponents of scale-invariant power laws that are universal (i.e. stable) across a wide range of different causal and mechanistic interactions in the system. Because of this universality across extremely different systems, “Many of the SOC models are highly idealized and *do not even attempt* to capture the basic interactions of natural systems.” (Watkins 2017, 23, my emphasis).

The remaining question, then, is *how we ought to justify* the use of such minimal models that completely ignore the causal and mechanistic interactions that produce these behaviors in their target systems. As I argued above, appealing to common causes or mechanisms doesn’t really help here because (1) the systems of interest are very heterogeneous with respect to their causes and mechanisms, and (2) all the scales of the system are important so it is unlikely that any of the available causal or mechanical models (even at more macroscales) would be able to include all the relevant scales and causal interactions and remain tractable. Here is where the concept of universality (classes), and the ways it is importantly different from a causal pattern, proves extremely useful. Because the minimal features that characterize the universality class are stable across systems that are very different with respect to their causes and mechanisms, all that is required is that the scientists build in the minimal features described above and the model will be within the same universality class as many real-world systems. As a result, the model will display the critical behaviors of interest to these modelers and can reveal how those behaviors depend on the minimal features that characterize the systems in the universality class. Indeed, modelers of SOC repeatedly note that their main goal is to identify universality classes and find minimal models within those classes that display the appropriate critical behaviors. As Preussner explains “The big challenges in SOC on a more technical level thus have remained the same for almost twenty years: on the most basic level, the identification of universality classes containing models that display solid scaling behavior” (Pruessner 2012, 5). The main reason this is important to scientists studying SOC is that: “*Universality justifies the simplified models in SOC*, which ignore all but a few details of what they are modelling” (Pruessner 2012, 19, my emphasis). In other words, the main reason discovering universality classes that contain minimal models is so important to these modelers is that showing that the minimal models are in the same universality class as various real systems is what *justifies* the use of the extremely minimal model to explain and understand the universal critical behaviors of real systems*.* Without this link between these minimal models and the behavior of real systems—whose causes and mechanisms are very different from the algorithms described by the minimal models—there would be no basis for inferring claims about how those real systems operate via the investigation of such simple models.

Once scientists have constructed a minimal model and demonstrated that it is in the same universality class as the real systems of interest, the explanation provided for these patterns involves two kinds of information:

1. The use of the minimal model to show how the observed macroscale pattern (the explanandum) depends on (changes to) the features that characterize/distinguish the universality class (e.g. the noise of the system, separation of time scales, or the system’s dimensionality).
2. The use of the minimal model and other modeling techniques (e.g. renormalization or homogenization) to demonstrate that the remaining heterogeneous features of the systems within the universality class (e.g. the features ignored or idealized by the minimal model) are irrelevant to displaying the universal patterns of behavior.

In combination, this set of dependence and independence information constitutes an explanans that enables scientists to explain the occurrence (and stability) of the universal patterns of behavior we observe across causally heterogeneous systems and across a wide range of spatial and temporal scales.

 At this point one might suggest that while minimal models can be used within explanations of universality, demonstrating that the various heterogeneous features among the systems in the universality class are irrelevant could perhaps be provided by an underlying causal explanation.[[20]](#footnote-20) Consequently, the idea would be that the dependence information provided by these minimal models could be complemented (or deepened) by appealing to causal explanations that underwrite the relevance or irrelevance of various features (e.g. a causal explanation of the system’s scale invariance). While the existence of such a complementary causal explanation is perhaps possible in some cases, it is not necessary to give this additional causal explanation for the above explanations to go through; i.e. the above explanations that appeal to minimal models and various mathematical techniques for delimiting universality classes are sufficient to explain the patterns of interest (and their stability) without having to appeal to underlying causal considerations for why various features are relevant or irrelevant. A good (but not decisive) indication of this is that the scientists who use these minimal models do not appeal to such an underlying causal story in their attempts to model and explain these patterns in active matter systems—philosophers might try to add in such an explanation, but it isn’t offered by the scientists who construct these explanations.

What is more, one of the main points of the above discussion is to show that the CCP approach is more than just the claim that what is cited in the explanation are causes. Crucially, what is claimed is that the explanation needs to cite causes (or mechanisms) that are *common* or *similar* across each of the cases. I am skeptical that this can be done in the cases discussed here. The first reason is that, as I argued above, the causes and mechanisms of the various systems that display the universal behaviors are extremely heterogeneous. Consequently, the causes and mechanisms that underwrite the irrelevance of various features will be very different across the systems that display these universal patterns. As a result, it is unclear how the demonstration of the irrelevance of various features could be provided by citing a common causal-mechanical story because *there are no common causes or mechanisms to cite!* Perhaps more importantly, I have argued that the causes and mechanisms at any scale of these systems are irrelevant to their displaying the universal behaviors of interest. Given this, it is unclear how citing features that are irrelevant to displaying the universal pattern would somehow explain the very irrelevance of those (or other) features. Therefore, even though such an underlying causal explanation is possible in some cases, in many cases we have good reasons for thinking that such a common causal explanation won’t be possible. In the next section, I provide additional reasons for thinking that these universal patterns cannot be explained by citing underlying commonalities between their causes and mechanism by unpacking several distinct forms of explanatory autonomy displayed by these cases.

**5. Universality, explanatory autonomy, and multiscale modeling**

Philosophers have long debated the nature of emergence and what counts as a genuinely emergent phenomenon. Rather than attempting to adjudicate between competing conceptions of emergence (or reduction), in this final section I want to look at the *practical modeling implications* of various kinds of ‘explanatory autonomy’ revealed by the above cases. To reiterate, my goal here isn’t to draw any metaphysical conclusions about how we ought to think about emergence (or reduction). Instead, I want to look at how different types (or conceptions) of explanatory autonomy displayed by these cases ought to impact the way in which scientists approach the modeling of multiscale phenomena. In order to do so, I begin with a useful distinction between two kinds of explanatory autonomy recently made by Margaret Morrison (2018).

Morrison begins by arguing that not all multiscale phenomena ought to be considered instances of emergent behavior (Morrison 2018, 2). In particular, she argues that there is a crucial difference between *ontological* and *epistemic* independence between a macroexplanation and microexplanation of some phenomenon (Morrison 2018, 6). When these two explanations are epistemically independent of one another, we *need not* appeal to the microconstituents of the system in order to explain the macrophenomena. This kind of explanatory autonomy, Morrison argues, is common to many physical systems. However, a stronger kind of explanatory autonomy is found in cases where the macrophenomena *cannot* be explained in terms of the microconstituents of the system. After drawing this distinction, Morrison argues that genuine emergence must involve some kind of ontologicalreason why the macrophenomenoncannot be explained by appealing to the microstructures of the system:

The relation between ontological and epistemic independence is especially important since the latter is a necessary but not a sufficient condition for emergence; the fact that we *need not* appeal to micro phenomena to explain macro processes is a common feature of physical explanation across many systems and levels. Instead, what is truly significant about emergent phenomena is that we supposedly *cannot* appeal to microstructures in explaining or predicting these phenomena, even though they are constituted by them. (Morrison 2018, 6)

Morrison’s key claim here is that epistemic barriers to explanation due to the number of factors involved, or other computational limitations, are not genuine cases of emergence. Instead, genuinely emergent phenomena present ontological barriers to explanation in terms of the system’s microconstituents. In order to capture these ontological barriers to explanation, Morrison goes on to argue that *universality* is the defining feature of emergence because it shows how the macroscale behaviors of interest can be ontologicallyindependent of the microscale features of the system. As Morrison puts it,

…universality, what I want to call the defining feature of emergence, is not simply the fact that we can use the same ‘theory’ to account for many different types of systems, or that properties of a large class of systems are independent of the dynamical details of those systems. Instead…it is the convergence of values for critical exponents in different types of systems that indicate the similarities in their critical behavior; similarities that cannot be explained by resorting to microphysical considerations…Universality classes are different in that they exhibit different properties and behaviors than those explainable in terms of their underlying constituents. (Morrison 2018, 13)

In other words, the distinguishing feature of emergence that we see in cases of universal behavior is that the macroscale phenomena cannot be explained in terms of the constituents of the system.

Now, even on Morrison’s much narrower definition of emergence, these minimal models will count as models of emergent behaviors because they aim to show that *very different physical systems can all give rise to the same universal behaviors.* As Bak and Chen tell us: “Self-organized criticality is a holistic theory: the global features, such as the relative number of large and small events, do not depend on the microscopic mechanisms. Consequently, global features of the system *cannot* be understood by analyzing the parts separately.” (Bak and Chen 1991, 46, my emphasis). This is the mark of universality that Morrison argues distinguishes genuine ontological emergence: the phenomena of interest simply *cannot* be explained by referencing the physical constituents of the system because we want to know why several different configurations of the physical features of the system give rise to the same universal critical behaviors. The practical modeling implication of this kind of ontological explanatory autonomy is that if scientists want to understand the universality of these critical phenomena, they need to move beyond attempting to describe the microconstituents that give rise to (or constitute) these behaviors in real systems and instead look at the use of minimal models that ignore almost all of the features of real systems in order to identify universality classes that range over very heterogeneous real, possible, and model systems. Furthermore, in virtue of this kind of ontological autonomy, these cases of modeling the critical behaviors of active matter systems also illustrate Morrison’s epistemic form of explanatory autonomy: the explanation of the observed macrobehaviors need not cite the microconstituents of the system. This epistemic form of explanatory autonomy further justifies these modelers’ attempts to explain these universal patterns by appealing to minimal models that ignore (or idealize) the microconstituents of the systems within the universality class.

However, I think these cases of multiscale modeling illustrate several additional kinds of explanatory autonomy beyond Morrison’s metaphysical and epistemological conceptions. This is because we need not cast our discussion of explanatory autonomy solely in terms of macrobehaviors and microconstituents; i.e. explanatory autonomy need not hold just between explanations or entities at ‘higher’ and ‘lower’ levels (or scales) of the system. Instead, some phenomena are explanatory autonomous *from particular types of features* regardless of whether those features occur at the micro or macro scales of the system. For example, there can be, and often are, causes and mechanisms at macroscales of the system. Yet, I have argued that these cases show that certain features of the system are autonomous of the causes and mechanisms of the systems of interest *at any scale (or ‘level’).* Therefore, another important kind of explanatory autonomy found in these cases, I argue, is that the behaviors of interest are *stable with respect to wholescale changes of the systems’ causes and mechanisms*. As a result, the minimal models used to explain these universal phenomena *need not* reference the physical causes or mechanisms of the system at any scale. This justifies scientific modelers in constructing minimal models that ignore most (if not all) of the realistic causes and mechanisms of the system in favor of trying to identify minimal features that characterize a universality class. This is certainly a kind of epistemic explanatory autonomy (since the explanation need not reference those kinds of features), but it does not require us to invoke a micro-macro distinction. In short, being able to explain a phenomenon without referencing *any* of the systems’ causal more mechanistic features goes beyond just being able to explain the phenomenon without referencing causes or mechanisms at microscales.

The above cases also reveal other (more pragmatic) kinds of explanatory autonomy that have additional implications for how modelers ought to approach these multiscale phenomena. Specifically, because critical phenomena involve all the scales of the system (i.e. they lack a characteristic scale) and depend on various scale-invariant features, these cases also show a kind of explanatory autonomy *from particular ways of attempting to model a system*. The practical modeling implication of this lack of a characteristic scale is that causal and mechanistic modeling techniques that are tied to particular scales will be insufficient for providing the desired explanation. In other words, I have argued that these universal phenomena *cannot* be accounted for by causal or mechanistic modeling techniques that are designed to capture features at a narrow range of scales for two reasons: (1) the features relevant to displaying the universal behaviors occur across a wide range of scales and (2) the universal behaviors are observed across very different spatial and temporal scales. Hence both characterizing the explanandum and providing the explanans will require us to account for a wide range of spatial and temporal scales. Here the explanatory autonomy in question is due to the limitations of the kinds of modeling techniques used for particular kinds of features. Because the phenomena of interest lack any characteristic scale and all of the scale of the system are important to these critical behaviors, modeling techniques designed to capture the causes and mechanisms of the system at particular scales will be unable to provide the desired explanation. The point is that *certain kinds of phenomena simply* *cannot be explained* *by certain kinds of modeling techniques* (regardless of whether those techniques are applied at micro, meso, or macro scales)*.* In the context of multiscale modeling, scientists often know that the explanandum depends on features across a wide range of scales or levels of the system. As a result, scientists often are not interested solely in macroscale autonomy from microscales constituents because the multiscale phenomenon of interest clearly is not autonomous of many of the features at smaller scales. However, in this context, we can still offer multiscale modelers guidance by being able to identify certain types of phenomena that cannot be explained by certain kinds of multiscale modeling techniques and suggesting other approaches that are more likely to succeed. For example, when investigating multiscale phenomena that lack a characteristic scale, identifying minimal models within a universality class characterized by scale-invariant power laws is far more likely to be successful than attempting to apply causal or mechanistic modeling techniques that target a narrow range of scales of the system. Identifying this kind of pragmatic explanatory autonomy from a particular kind of multiscale modeling strategy is extremely helpful for making modeling decisions even in contexts where questions such as “What is the characteristic scale of the phenomenon?” or “Does the phenomenon occur at the micro or macro scale?” do not make sense.

In summary, these cases enable us to distinguish several different conceptions of explanatory autonomy because the explanations sought in these cases: (1) cannot be provided by appealing to the (causal or mechanistic) microconstituents of real systems, (2), need not reference the (causal or mechanistic) microconstituents of the system, (3) need not reference the causes or mechanisms of the system at any scale and (4) cannot be provided by causal or mechanistic modeling approaches that are tied to particular scales. These ontological, epistemic, and pragmatic conceptions of explanatory autonomy provide clear justifications for why these modelers instead aim to find minimal models within the same universality class as real systems that display the patterns of interest despite ignoring the causes or mechanisms of the model’s target systems.

**6. Conclusion**

I have argued that the use of minimal models to explain and understand active matter and SOC behaviors is (best) accounted for by noting that these behaviors are universalacross classes of real, possible, and model systems whose causal and mechanistic features are drastically different. In addition, following the scientists who construct these models, I have argued that it is their being within the salient universality classes that justifies the use of these minimal models to explain and understand these complex behaviors in real systems. Consequently, the modeling techniques and justifications involved in these cases are at odds with many philosophical accounts that suggest that patterns ought to be explained by abstractly describing causal or mechanistic features common to each instance of the pattern. Instead, the explanation provided in these cases requires the delimiting of various universality classes and showing that various minimal models are within those universality classes. Identifying these universality classes is what shows why the same minimal models can be justifiably used to explain behaviors across a wide range of fields and across a wide range of scales. These cases have also revealed a number of different kinds of explanatory autonomy that can provide practicing scientific modelers guidance about which modeling approaches are likely to be successful at explaining these kinds of multiscale phenomena. As a result, philosophical discussions of the explanation of patterns and the modeling of multiscale phenomena need to account for the various kinds of explanatory autonomy displayed distinguished here and scientists’ continual use of minimal models to explain and understand universal patterns.

**References**

Bak, P. (1996). *How nature works: the science of self-organized criticality*. New York, NY: Springer-Verlag.

Bak, P. and Chen, K. (1991). Self-organized criticality. *Scientific American*, January 1991, 46-53.

Bak, P. and Spenppen, K. (1993). Punctuated equilibrium and criticality in a simple model of evolution. *Physical Review Letters*, *71*, 4083.

Bak, P., Tang, C. and Wiesenfeld, K. (1987). Self-organized criticality: An explanation of 1/*f* noise. *Physical Review Letters*, *59*(4), 381-384.

Bak, P., Chen, K. and Tang, C. (1990). A forest-fire model and some thoughts on turbulence. *Physics Letters A*, *147*, 297-300.

Batterman, R. W. (2002). *The devil in the details: Asymptotic reasoning in explanation, reduction, and emergence*. Oxford: Oxford University Press.

Batterman, R. W. (2013). The tyranny of scales. In *The Oxford Handbook of Philosophy of Physics* (ed.) Robert Batterman. Oxford: Oxford University Press, pp. 255-286.

Batterman, R. W. (2017). Autonomy of theories: An explanatory problem. *Noûs*, doi: 10.1111/nous.12191, 1-16.

Batterman, R. W. (2019). Universality and RG Explanations. *Perspectives on Science,* *27*, 26-47.

Batterman, R. W. and Rice, C. (2014). Minimal model explanations. *Philosophy of Science*, 81(3), 349-376.

Ben-Hur, A. and Biham, O. (1995). Universality in sandpile models. *Physical Review E*, *53*(2), 1317-1321.

Bokulich, A. (2011). How Scientific Models Can Explain. *Synthese*, *180*, 33–45.

Bursten, J. R. (2018). Conceptual strategies and inter-theory relations: The case of nanoscale cracks. *Studies in History and Philosophy of Modern Physics*, *62*, 158-165.

Carlson, J. M. and Swindel, G. H. (1995). Self-organized criticality: Sandpiles, singularities, and scaling. *Proceedings of the National Academy of the Sciences*, *92*, 6712-6719.

Chaté, H., Ginelli, F. Grégoire, G., Peruani, F., and Raynaud, F. (2008). Modeling collective motion: variations on the Vicsek model. *European Physical Journal B*, *64*, 451-456.

Chen, K., Bak, P. and Obukov, S. P. (1991). Self-organized criticality in a crack-propagation model of earthquakes. *Physical Review A*, *43*, 625.

Craver, C. (2006). When mechanistic models explain. *Synthese*, 153, 355-376.

Craver, C. (2007). *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Oxford: Oxford University Press.

Drossel, B. and Schwabl, F. (1992). Self-organized critical forest-fire model. *Physical Review Letters*, *69*(11), 1629-1632.

Franklin, A. (2019). Universality reduced. *Philosophy of* Science *86, 1295-1306.*

Frigg, R. (2003). Self-organised criticality—what it is and what it isn’t. *Studies in History and Philosophy of Science Part A, 34*(3): 613-632.

Glennan, S. (2017). *The New Mechanical Philosophy*. Oxford: Oxford University Press.

Ginelli, F. (2016). The physics of the Vicsek model. *The European Physical Journal Special Topics*, *225*, 2099-2117.

Green, S. and Batterman, R. (2017). Biology meets physics: Reductionism and multi-scale modeling of morphogenesis. *Studies in History and Philosophy of Biological and Biomedical Sciences*, *61*, 20-34.

Jensen, H. J. (1998). *Self-organized criticality: Emergent Complex Behavior in Physical and Biological Systems*. Cambridge: Cambridge University Press.

Kadanoff, L. P. (2000). *Statistical Physics: Statics, Dynamics, and Renormalization*. Singapore: World Scientific.

Kadanoff, L. P. (2013). Theories of matter: Infinities and renormalization. In *The Oxford Handbook of Philosophy of Physics*, (ed.) Robert Batterman. Oxford: Oxford University Press, pp. 141-188.

Kaplan, D. M. and Craver, C. F. (2011). The explanatory force of dynamical and mathematical models in neuroscience: A mechanistic perspective. *Philosophy of Science*, 78, 601-627.

Lee, S. B. (2017). Classification of universality classes for quasideterministic sandpile models. *Physical Review E*, *96*, 012117.

Manna, S. S. (1991). Two-state model of self-organized criticality. *Journal of Physics A General Physics*, *24*, L363-L369.

Malamud, B. D., Morein, G. and Turcotte, D. L. (1998). Forest fires: An example of self-organized critical behavior. *Science,* New Series, *281*(5384), 1840-1842.

Morrison, M. (2015). *Reconstructing Reality: Models, Mathematics, and Simulations*. New York: Oxford University Press.

Morrison, M. (2018). Turbulence, emergence and multi-scale modeling. *Synthese*, https://doi.org/10.1007/s11229-018-1825-5

Oden, J. T. (2006). Simulation based engineering science—an NSF Blue Ribbon Report. www.nsf.gov/pubs/reports/sbes\_final\_report. pdf.

Palmieri, L. and Jensen, H. J. (2020). The forest fire Model: The subtleties of criticality and scale invariance. *Frontiers in Physics*, *8*: 1-8.

Peterson, G. D. (2008). Self-organization and discontinuities in ecosystems. In Allen, C.R. and Holling, C. S. (eds.), *Discontinuities in Ecosystems and Other Complex Systems*. Columbia: Columbia University Press.

Potochnik, A. (2017). *Idealization and the Aims of Science*. Chicago: University of Chicago Press.

Pruessner, G. (2012). *Self-Organised Criticality: Theory, Models and Characterisation*. Cambridge: Cambridge University Press.

Rice, C. (2017). Models Don’t Decompose That Way: A Holistic View of Idealized Models. *The British Journal for the Philosophy of Science*, axx045, https://doi.org/10.1093/bjps/axx045.

Rice, C. (2018). Idealized models, holistic distortions and universality. *Synthese*, *195*(6), 2795-2819.

Rice, C. (2021). *Leveraging Distortions: Explanation, Idealization and Universality in Science*. Cambridge, MA: MIT Press.

Saatsi, J., & Pexton, M. (2013). Reassessing Woodward’s account of explanation: Regularities, counterfactuals, and non-causal explanations. *Philosophy of Science*, *80*, 613–624.

Scheffer, M. and Carpenter, S. (2003). Catastrophic regime shifts in ecosystems: Linking theory to observations. *Trends in Ecology and Evolution,* *18*, 648–656.

Strevens, M. (2008). *Depth: An account of scientific explanation*. Cambridge: Harvard University Press.

Turcotte, D. L. and Malamud, B. D. (2004). Landslides, forest fires and earthquakes: examples of self-organized critical behavior. *Physica A*, *340*, 580-589.

Vicsek, T., Czirók, A., Ben-Jacob, E., Cohen, I. and Schochet, O. (1995). Novel type of phase transition in a system of self-driven particles. *Physical Review Letters*, *75*(6), 1226-1229.

Vicsek, T. and Zafeiris, A. (2012). Collective motion. *Physics Reports*, *517*, 71-140.

Watkins, N. W., Pruessner, G., Chapman, S. C., Crosby, N. B., and Jensen, H. J. (2017). 25 years of self-organized criticality: concepts and controversies. *Space Science Review*, *198*, 3-44.

Winsberg, E. (2006). Handshaking your way to the top: Inconsistency and falsification in intertheoretic reduction. *Philosophy of Science*, *73*: 582-594.

Winsberg, E. (2010). *Science in the Age of Computer Simulation*. Chicago: University of Chicago Press.

Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.

Woodward, J. (2010). Causation in biology: stability, specificity, and the choice of levels of explanation. *Biology and Philosophy*, *25*, 287-318.

Zinck, R. D. and Grimm, V. (2009). Unifying wildfire models from ecology and statistical physics. *American Naturalist*, *174*(5), E170-E185.

1. While the ‘universality account’ has been applied to number of case studies in (Batterman and Rice 2014; Rice 2017, 2018, 2020), it is still routinely claimed that such an account only applies to cases of renormalization. Thus, one aim of this paper is to argue that the universality account is more general than has been assumed and to clarify the scope of its application. That is, I aim to provide further support for the claim that such an account is generally applicable to a wide variety of cases that include instances of self-organized criticality. [↑](#footnote-ref-1)
2. In this way, the model is similar to the alignment of spins used in Ising models of ferromagnetic phase transitions. [↑](#footnote-ref-2)
3. Since all the particles are assumed to have the same magnitude of their velocity, this average is determined solely by the various *directions* of the particles in the system that are continuously reorienting to align with the average direction of the particles in their neighborhood. [↑](#footnote-ref-3)
4. This implies that, “we can assume that in the thermodynamic limit our model exhibits a kinetic phase transition analogous to the continuous phase transition in equilibrium systems” (Vicsek et al. 1995, 1228). [↑](#footnote-ref-4)
5. In addition, these modelers propose that their model will likely be applicable to drastically different physical systems; e.g. the clustering and convection involved in a system of disks floating on air. [↑](#footnote-ref-5)
6. Once these essential minimal features were identified, several recent studies began investigating the different universality classes associated with slightly altered versions of ‘Vicsek-like’ models (Ginelli 2016, 2114). [↑](#footnote-ref-6)
7. This analysis contrasts a bit with Roman Frigg’s (2003) negative evaluation of the explanatory success of these models. While the models are certainly highly idealized and can be fruitful in the ways Frigg suggests, I contend that further investigation of these models over the last twenty years has provided a clearer sense of which features of real systems these behaviors depend on and why many of the features idealized by these models are irrelevant to whether or not the system displays the universal behaviors of interest. Consequently, in contrast with Frigg’s discussion, I will argue that these models can provide a plethora of information about dependencies and independencies that hold in real-world systems without faithfully representing the features of those systems. [↑](#footnote-ref-7)
8. In fact, BTW (1987) originally argued that the hallmark of SOC is the lack of any scale in time as well as space, but the necessity of the lack of spatial scale is debated by some physicists (Jensen 1998). [↑](#footnote-ref-8)
9. It is worth noting here that I agree with Frigg (2003) that many of the claims of the ubiquity of SOC behaviors have been exaggerated beyond what is warranted by the empirical evidence. In particular, I don’t think we have reason to believe that SOC shows us generally ‘how nature works’ or ‘how everything evolves’. However, I do think that scientific modelers have shown that many of these behaviors are quite universal in the sense that they are widely applicable and largely autonomous of the particular features of real (and model) systems. [↑](#footnote-ref-9)
10. For example, Malamud et al. (1998) use a minimal model from statistical physics to explain the frequency-area distributions of actual wildfires and to study the recent fire history in Yellowstone National Park. [↑](#footnote-ref-10)
11. There are several different ways that this toppling threshold can be calculated. Some models assume that simply having enough grains at a particular site will result in toppling (e.g. the site having more than 4 grains). Others, such as the model described here define toppling in terms of the slope/difference between the number of grains at the site and the number of grains at its neighboring sites. Throughout the paper, I’ve tried to present the examples in their simplest form unless the details make a difference to the arguments or conclusions I defend. Thanks to an anonymous reviewer for noting this difference between the ways these models are implemented in different cases. [↑](#footnote-ref-11)
12. Although these features appear to be both necessary and sufficient for the systems in this universality class to display SOC behaviors, the necessary features for inclusion in a universality class will not always be sufficient (on their own) to display the macroscale behaviors characteristic of that class. For example, while several universality classes in physics show that the system’s symmetry of the order parameter and dimensionality are essential to determining the universality class of the system, merely having those features (alone) is not sufficient for the system to display the macroscale behaviors of interest (e.g. phase transitions) since those features only result in the critical behaviors of interest when they are incorporated within particular systems with myriad other features. [↑](#footnote-ref-12)
13. In particular, specifying the scope of the causal pattern shows that the same causal pattern is embodied across different systems. [↑](#footnote-ref-13)
14. Etiological explanations describe the causal history of the explanandum; whereas constitutive explanations describe the mechanism that underlies the phenomenon. [↑](#footnote-ref-14)
15. If there are differences among the causes/mechanisms that produce the different instances of the pattern, these authors suggest that we can explain the pattern by abstracting away from those differences and identifying the common features of the causes or mechanisms that produce each instance of the pattern. [↑](#footnote-ref-15)
16. Thanks to an anonymous reviewer for suggesting this possible reply. [↑](#footnote-ref-16)
17. Thanks to two anonymous reviewers who raised this possible reply and for pushing me to clarify which of the critiques of the CCP view depend on the assumption that all explanations are causal and which are independent of that assumption. [↑](#footnote-ref-17)
18. Thanks to an anonymous reviewer for raising this objection. [↑](#footnote-ref-18)
19. Indeed, like the cases discussed above, applying renormalization techniques often enables scientists to identify extremely minimal models that are within the same universality class as the real systems of interest (Batterman and Rice 2014). [↑](#footnote-ref-19)
20. An anonymous reviewer suggested that this kind of objection is proposed by Alexander Franklin (2019). Franklin’s discussion is specifically tied to RG explanations of universality so it isn’t entirely clear just how those arguments would apply here since RG is not the only way to explain a universal pattern. However, I’ve tried to respond to what I take to be the general line of objection here. Since I don’t have the space to work through all of my responses to Franklin’s view, I will, instead, direct the reader to Robert Batterman’s paper (2019) that provides many of the same responses I would give. [↑](#footnote-ref-20)