Samuel C. Fletcher

Similarity Structure on Scientific Theories

Received 30 Apr. 2017; accepted 19 Feb. 2018

Abstract: I review and amplify on some of the many uses of representing a scientific theory in a particular context as a collection of models endowed with a similarity structure, which encodes the ways in which those models are similar to one another. This structure, which is related to topological structure, proves fruitful in the analysis of a variety of issues central to the philosophy of science. These include intertheoretic reduction, emergent properties, the epistemic connections between modeling and inference, the semantics of counterfactual conditionals, and laws of nature. The morals are twofold: first, the further adoption of formal methods for describing similarity (and related topological) structure has the potential to aid in decisive progress in philosophy of science; and second, the selection and justification of such structure is not a matter of technical convenience, but rather often involves great conceptual and philosophical subtlety. I conclude with various directions for future research.

1 Introduction

Most work on the role of similarity in philosophy of science has focused on analogical (i.e., similarity-based) reasoning (Hesse, 1966; Bartha, 2016) and on scientific representation, i.e., whether and how models are similar to the targets they represent (Giere, 1988; Weisberg, 2013; Frigg and Nguyen, 2016). Yet, many topics in philosophy of science concern or depend on the relationships between scientific models themselves—models that are sometimes freestanding, but often associated with a particular theory or theories—along various aspects: How are they similar? How do they differ? In what ways are they the same? These questions concern not just philosophers of science, but practicing scientists, too (Frigg and Hartmann, 2012). While scientists and philosophers have usually developed *ad hoc*, often informal methods for answering these questions in specific contexts, there is some benefit to pursuing a more unified, systematic approach. Besides the pragmatic aid it could provide in orienting possible answers to analogous questions in new domains, it would also reveal the com-

Samuel C. Fletcher: Department of Philosophy, University of Minnesota, Twin Cities, & Munich Center for Mathematical Philosophy, Ludwig Maximilian University of Munich

mon types and degrees of structure that these questions probe. In doing so, it might also suggest new solutions to old philosophical problems and even novel questions that the potpourri obscured.

My main purpose here is to outline a research program that does just this, using tools adapted from *topology*.[1] While often conceived of as the mathematical study of the properties of spaces preserved under continuous deformation, such as stretching and bending (but not tearing or gluing), topology has in fact a much greater domain of application. Here, the basic insight is an analogy between various sorts of similarity and geometric proximity. If one represents a theory by its models, then the resulting space of models represents the collection of possibilities or states of affairs that the theory affords.[2] In other words, instead of each element representing a point in (a hypothetical) physical space, it represents a way a part of the world could be, according to the theory in question. A similarity structure on this space encodes a particular sort of similarity amongst these models. Just as the topology of a geometrical space determines which curves in that space vary continuously, a similarity structure on the models of a theory determines which families of models represent continuously varying sets of possibilities. And just as the topology of a geometrical space determines which sequences of points converge to another, so too does a similarity structure on the models of a theory determine which sequences of states of affairs eventually become arbitrarily similar to another.

Equipping the space of models with extra structure is necessary for this task because the bare space itself does not provide the expressive resources for these kinds of statements. However, I do not hold that a theory should be

1 I am not the first to suggest adding some sort of topological structure to the models of a scientific theory: one finds it—in particular, the invocation of uniform structure (Willard, 1970, Ch. 9)—in the German structuralist school (Schmidt, 2014), although that approach has tended to assign a single structure to a theory; see Balzer et al. (1987, Ch. 7) and Ludwig and Thurler (2006, Ch. 4.6). (One does find the suggestion of using topology, not uniform structure, in Scheibe (2001, p. 318–9), but this suggestion does not seem to have been pursued systematically.) By contrast, as I discuss later in this section, I take such sorts of structure to accrue only to particular contexts of investigation for a theory; in different contexts, different similarity structure may be determined. See also Mormann (2013) for more general reflections on the history and relations between topology and philosophy of science.

2 I do not assume that a theory merely is a collection of models—see the discussion at the beginning of section 2. Also, I do not take this representation to imply any substantive commitment beyond what is required for the typical uses of a scientific theory common to plausible versions of scientific realism and antirealism. The topological approach to the formalization and application of similarity is intended to be a broad tent.

represented *simpliciter* by a structured space—that is, its set of possibilities endowed with a similarity structure. After all, it seems reasonable to expect the relevant notion of similarity to be contextual, determined by the details of a particular investigation, which would in general preclude associating a single similarity structure with the possibilities a theory affords.[3] Nevertheless, if one can adequately formulate, in a particular context, how a theory's models are to be regarded as relevantly similar, then one can use the structure thereby induced to answer longstanding questions regarding a surprising variety of topics.

In addition to offering a formal, fine-grained representation of similarity concepts, one can also use the framework to discover what our commitments about relevant similarity entail, which could well be surprising and would count as a discovery in the colloquial sense. For instance, in section 5.2, I indicate how according to one way of describing how relativistic spacetimes are similar, every one is arbitrarily similar to one with time travel. There is also an indirect way in which the framework facilitates a process of discovering new similarities, namely through a kind of reflective equilibrium. If one is unsure what the relevant similarities are supposed to be in a certain context, then one can test an answer by seeing what results it entails, then modifying the answer in response if necessary. Through the interplay of describing similarities precisely and the results this entails, one may indeed discover ways in which the models are relevantly similar that one did not before consider.

So, after section 2 introduces the formalism of similarity structures and their relations to various topological structures, the following four sections will discuss a selection from this variety.[4] Section 3 concerns approximative, or limiting-type reduction between different scientific theories concerned with overlapping domains. There, similarity structure on the joint collection of models of both theories helps one explicate the circumstances under which the models of one approximate, or are sufficiently similar to, the other, thereby lending to the explanation of the (typically older, less accurate) one in terms of the other. The natural conceptual counterpart to reduction is emergence, and the topic of section 4 is the emergence of properties, in particular those that arise in some models with respect to others, such as for models that are (in a precise sense) the limit of a sequence of other models. An essential aspect of emergent properties is their comparative novelty or inexplicability, and this section exhibits how this can be defined in terms of dissimilarity. Moving broadly from primarily represen-

3 One can provide more formal arguments to this effect in particular cases. See, e.g., Fletcher (2016) for the case of topological similarity structures for general relativity.

4 Many of these topics are adumbrated in Fletcher (2014, Ch. 5).

tational topics to inferential ones, section 5 focuses on the problem of justified inference from models. When a model of a state of affairs is idealized—a situation that almost always obtains—it contains deliberate distortions or simplifications in its representation thereof.[5] Which properties of a model is one justifying in inferring about its target? I formulate and discuss a necessary condition for such an inference, inspired by Duhem (1954), which requires the inferred property to hold of all sufficiently similar models under consideration. These kinds of conditions have long been closely related to issues involving modality and modal reasoning, so section 6 outlines how similarity structure on a collection of models suffices to provide an ordering semantics for counterfactual conditionals (Lewis, 1981; Kratzer, 1981; Swanson, 2011), with the models as stand-ins for possible worlds. This suggests an interesting new approach to laws of nature that shares affinities with both antirealist and systems accounts thereof, perhaps without some of their shortcomings. Finally, in the concluding section 7, I review some broad advantages to the program, a few shortcomings, and a swath of interesting questions arising from each of the above topics that have yet to be fulled explored.

2 Similarity and Topology

2.1 Theories and Models

What does it mean to represent a theory by its models? I have in mind, for example: representing (pure) General Relativity by the class of Lorentzian geometries (M, g_{ab}),[6] where M is the manifold of spacetime events and g_{ab} the metric tensor; representing the simple theory of predator-prey interactions in ecology by solutions $(x(t), y(t))$ to the Lotka-Volterra equations (Sarkar, 2016, §2.1),

$$\frac{dx}{dt} = \alpha x - \beta xy, \qquad \frac{dy}{dt} = \delta xy - \gamma y,$$

which denote the population of predators x and prey y over time t; or representing the preferences of an agent by a partial order $\preceq$ over states of the world

5 These are, strictly speaking, only one type of idealization—Galilean idealization (McMullin, 1985). For more on the categorization of types of idealization, see, e.g., Frigg and Hartmann (2012) and Weisberg (2013, Ch. 6).

6 Strictly speaking, the class of smooth, connected, four-dimensional Lorentzian manifolds—see Hawking and Ellis (1973, Ch. 3.1) or Malament (2012, p. 119).

A.[7] Thus, I have in mind the notion of model as used by applied mathematicians, rather than logicians, where the model represents a possible state of affairs falling under the domain of the theory.

There are several advantages to this approach to representing theories, including eschewing the complications of providing axiomatizations for them—a form in which real scientific theories are not often provided—and guarding against unintended models. But I want to stress that I do not demand that one *identify* a theory as a class of models, only that one can *represent* it as such (Thompson, 2007, p. 485–6). In doing so, one may well distort a multifaceted theory that has nonformal aspects (Bailer-Jones, 2002; Craver, 2002).[8] Accordingly, I do not intend here to engage with the debate concerning the what the "right" structure for scientific theories is supposed to be—even active defenders of the syntactic view of theories take a concern with models to be consonant with their approach, a shared commitment amongst virtually all approaches (Winther, 2015; Halvorson, 2013, 2016), which is all that similarity structure requires. (Cf. footnote 2.) But, I do confine attention to theories not suffering so much from vagueness that their models can be described as a definite class of objects—usually, mathematical ones. For example, "classical mechanics," despite being obviously a mathematical theory, is on its own too vague to specify a definite class of models—there are many different varieties of classical mechanics (Wilson, 2006, 2009), differing even in whether they are deterministic (Malament, 2008; Fletcher, 2012). Thus, to represent a theory of classical mechanics by its models, one must describe precisely which possible states of affairs it allows. Moreover, in dynamical theories, such as theories of mechanics, these states of affairs are not typically states at a time—indeed, in relativistic theories there is no objective such state—but rather (perhaps partial) histories.

Now, whatever the nature of the states of affairs or phenomena described by a theory, different ones can be similar to one another. For instance, the history of a proton falling in a gravitational field from a certain location in a relatively electrically neutral space will be similar in many respects to that of a neutron falling from approximately the same location. These respects might include the shape and spatial location of the trajectories, but not the type of particle involved or the electromagnetic fields in the vicinity of the trajectories. Generally, similarity can come in degrees and in myriad types, not all of which

7 A partial order $\preceq$ on a set A is a binary relation on A that is reflexive ($\forall a, a \preceq a$), transitive ($\forall a, b, c$, if $a \preceq b$ and $b \preceq c$ then $a \preceq c$), and anti-symmetric ($\forall a, b$, if $a \preceq b$ and $b \preceq a$ then $a = b$).

8 Or that has more formal structure (e.g., category-theoretic) than is traced here (Halvorson, 2012, 2013, 2016; Halvorson and Tsementzis, 2017).

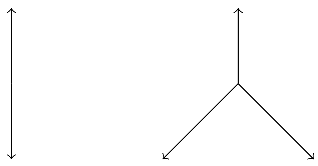

Fig. 1. A representation of two "path" models, the left one linear (L) and the right one with a fork (F).

are relevant for a given context, because not all make a difference to the questions investigated in that context. But if the context is unambiguous enough, one can apply the relevant notion of similarity with the following formal structure.

2.2 Similarity Structure

Let us begin with defining qualitative, binary similarity relations:

Definition 1. A *similarity relation* $\sim$ on a set X is a non-empty binary relation on X that is *quasi-reflexive*: for all $x, y \in X$ if $y \sim x$, then $y \sim y$ and $x \sim x$.

The relation $y \sim x$ is interpreted as "y is similar to x". Quasi-reflexivity requires that if any element is similar to another, then both the former and the latter are similar to themselves.

Now, similarity relations are commonly demanded (Carnap, 1967; Schreider, 1975; Mormann, 1996; Konikowska, 1997) to be not just quasi-reflexive, but *reflexive*—$x \sim x$ for every $x \in X$—and *symmetric*: if $y \sim x$ then $x \sim y$. While many useful similarity relations do satisfy these conditions, there are also some that do not. For example, consider a theory with many sorts of models, but exactly two of which are distinct paths, both infinitely long, with the first linear (L), and the second with a fork (F)—see Fig. 1. Say that one object, y, is *observationally similar* to another, x, denoted $y \sim_O x$, if they are both paths and any observer having traveled a finite distance in x has had experiences compatible with being in y.[9] Then $\sim_O = \{(L, L), (F, L), (F, F)\}$. Observational similarity is only quasi-reflexive because it simply does not apply to non-paths, and it is not symmetric because any observer having traveled any finite distance

9 This is a simplified analog of an independently interesting property described for relativistic spacetimes, *(weak) observational indistinguishability* (Malament, 1977; Manchak, 2009).

in L could well be on a branch of F, but observers in F who encounter the fork have experiences incompatible with being on L.

Of course, one could consider many other features in determining how these models are similar. As first argued explicitly by Goodman (1972), similarity judgments are contextual and often come in degrees or types.[10] Thus, considering the ways in which models are similar to one another can determine more than a single similarity relation.

Definition 2. A *similarity space* is an ordered pair $(X, \mathcal{S})$, where X is a set and $\mathcal{S}$ is a set of similarity relations on X, called a *similarity structure*.[11]

For such a similarity structure $\mathcal{S}$ containing $\sim_\alpha$, $y \sim_\alpha x$ is interpreted as "y is similar (in manner α) to x".

Nor do we need to restrict ourselves to a fixed similarity structure for the models of a theory, which should rather depend on a context of use: its similarity relations should be all and only those that make a difference to the range of investigative questions asked in that context. In the above example with paths, for instance, the context determined that only those differences which made a difference to the observations of any observer should be accounted for by the similarity relations. But a different context could well determine some symmetric or even reflexive similarity relations to be included in the relevant similarity structure.

In any case, once a set of models X of a theory has been equipped with a similarity structure $\mathcal{S}$ to form a similarity space $(X, \mathcal{S})$, one can define many interesting and useful relations from them.

Definition 3. The *domain at* $x \in X$ of a similarity relation $\sim$ on X is the set $D(x) = \{y \in X : y \sim x\}$.[12]

10 Many of the formal models of similarity influenced by this analysis has focused on cognitive judgments of similarity; see Decock and Douven (2011) for a review of and references to this literature.

11 Here, my terminology for a similarity structure follows Konikowska (1997) rather than Mormann (1996), who takes it to denote what I would call here a similarity space whose similarity structure consists of just a single similarity relation, or Schreider (1975), who does the same but calls them tolerance relations and tolerance spaces. Mathematically, a similarity structure is very close to being a base for a semi-quasi-uniform space (Pu and Pu, 1974), or what Császár (1974) calls a pseudo-uniform space.

12 These sets are called similarity neighborhoods in Mormann (1996) and tolerance classes in Schreider (1975), although, as alluded above, these authors assume that a similarity (tolerance) relation must be reflexive and symmetric.

The domain of $\sim$ at $x \in X$ is just the set of all the elements in X that are similar to x.

In the following, let $\mathcal{B}_x$ denote the set of domains at $x \in X$ of the similarity relations in $\mathcal{S}$ and $\wp X$ the power set of X.

Definition 4. The *closeness operator* $\mathrm{cl}_{\mathcal{S}} : \wp X \to \wp X$ of a similarity structure $(X, \mathcal{S})$ is defined by $\mathrm{cl}_{\mathcal{S}}(A) = \{x \in X : \forall D(x) \in \mathcal{B}_x, D(x) \cap A \neq \emptyset\}$.

The value of the closeness operator acting on a set $A \subseteq X$ is the collection of all models arbitrarily similar to any model in A, i.e., those similar to the models of the latter in all respects determined to be relevant by $\mathcal{S}$.

The closeness operator also has some important properties. First, note that it allows one to define important substructures of a similarity space:

Definition 5. The *common domain* of a similarity space $(X, \mathcal{S})$ is the set $\mathrm{cl}_{\mathcal{S}}(X)$.

The common domain of a similarity space is just the elements of the space the domain at which is non-empty for each relation in the space's similarity structure.

Proposition 1. *The closeness operator* $\mathrm{cl}_{\mathcal{S}}$ *of a similarity space* $(X, \mathcal{S})$ *is a preclosure operator on its common domain,*[13] *i.e., it satisfies the following conditions:*
1. $\mathrm{cl}_{\mathcal{S}}(\emptyset) = \emptyset$,
2. $A \subseteq \mathrm{cl}_{\mathcal{S}}(A)$ *for every* $A \subseteq \mathrm{cl}_{\mathcal{S}}(X)$, *and*
3. $\mathrm{cl}_{\mathcal{S}}(A \cup B) = \mathrm{cl}_{\mathcal{S}}(A) \cup \mathrm{cl}_{\mathcal{S}}(B)$ *for all* $A, B \subseteq \mathrm{cl}_{\mathcal{S}}(X)$.

The proof follows immediately from the relevant definitions. Clearly, if the interpretation of the closeness operator is that it yields all the models arbitrarily similar to the ones on which it acts, then it should indeed be the case that no models are arbitrarily similar to the empty collection of models; the models themselves are arbitrarily similar to themselves; and those that are arbitrarily similar to a model in either A or B are just those that are arbitrarily similar to a model in A or a model in B.

There is another important relation involving the concept of arbitrary similarity that holds not between sets and sets, but between sequences and points.

13 See Čech (1966, §14), who calls them simply closure operators; in modern usage that name is reserved for preclosure operators for which $\mathrm{cl}(\mathrm{cl}(A)) = \mathrm{cl}(A)$. The closeness operator of a topological similarity space (definition 14) satisfies this additional condition, which, together with the three in the proposition, are known as the Kuratowski closure axioms (Willard, 1970, p. 26).

Definition 6. A sequence $\{x_n\}_{n\in\mathbb{N}}$ of models of X *converges to* $x \in X$ *(with respect to* $\mathcal{S}$*)*, written $x_n \to x$, when for every $D(x) \in \mathcal{B}_x$, there is some $N \in \mathbb{N}$ such that $x_n \in D(x)$ if $n \geq N$.

In other words, a sequence of models converges to another when the former eventually become arbitrarily similar to the latter, i.e., the sequence is eventually contained in every $D(x) \in \mathcal{B}_x$.

In similarity spaces, convergence implies closeness, but not vice versa.

Proposition 2. *Let* $(X, \mathcal{S})$ *be a similarity space and* $x, x_n \in X$ *for each* $n \in \mathbb{N}$*. If* $x_n \to_{\mathcal{S}} x$*, then* $x \in \mathrm{cl}_{\mathcal{S}}(\{x_n\})$*. The reverse implication does not hold.*

Proof. The first part follows immediately from the definitions. For the second part, let $X = \mathbb{R}^{\mathbb{R}}$, the space of all real-valued functions of a single real variable. Further, for each $f \in X$, define $D^{x_1,\ldots,x_n}_{\epsilon_1,\ldots,\epsilon_n}(f) = \{g \in X : \forall k \in \{1,\ldots,n\}, |g(x_k) - f(x_k)| < \epsilon_k\}$ for any $\{x_1,\ldots,x_n\},\{\epsilon_1,\ldots,\epsilon_n\} \in \mathbb{R}^n$, and let $\mathcal{B}_f = \{D^{x_1,\ldots,x_n}_{\epsilon_1,\ldots,\epsilon_n}(f) : n \in \mathbb{N}\}$. If E is the set of $f \in X$ such that $f(x) = 0$ or 1 and $f(x) = 0$ only finitely often while $g(x) = 0$, then one can show (Willard, 1970, p. 71–2) that $g \in \mathrm{cl}_{\mathcal{S}}(E)$, yet no sequence in E converges to g. □

This result is important because it shows that, in general, the models arbitrarily similar to some collection of models are not necessarily those to which sequences from the latter converge.

Now consider another similarity space $(X', \mathcal{S}')$, letting $\mathcal{B}'_{x'}$ denote the set of domains at $x' \in X'$ of the similarity relations in $\mathcal{S}'$.

Definition 7. A function $f : X \to X'$ is *continuous (with respect to* $\mathcal{S}$ *and* $\mathcal{S}'$*) at* $x \in X$ when for every $D'(f(x)) \in \mathcal{B}'_{f(x)}$, there is some $D(x) \in \mathcal{B}(x)$ such that $f[D(x)] \subseteq D'(f(x))$. It is *continuous (simpliciter)* when it is continuous at each $x \in X$.

In other words, a function is continuous when similar elements in the function's range contain the image of sufficiently similar elements in the function's domain.

The continuous maps are the morphisms for similarity spaces, as the next proposition attests.

Proposition 3. *If* $(X, \mathcal{S}_X)$*,* $(Y, \mathcal{S}_Y)$*, and* $(Z, \mathcal{S}_Z)$ *are similarity spaces and* $f : X \to Y$ *and* $g : Y \to Z$ *are continuous, then* $g \circ f : X \to Z$ *is continuous.*

Proof. Analogous to the case of topological spaces, cf. Willard (1970, p. 45). □

Thus the similarity isomorphisms are just the bicontinuous functions.

2.3 Similarity and Topology

While one can investigate many features of theories by representing them in particular contexts with similarity spaces as given in the aforementioned definitions, a link with more well-established mathematical tools can be forged if these similarity spaces satisfy some very basic conditions.

Definition 8. A similarity space $(X, \mathcal{S})$ is *neighborly* when it satisfies the following conditions:
1. for every $x \in X$, there is some $\sim \in \mathcal{S}$ such that $x \sim x$; and
2. for every $\sim_\alpha, \sim_\beta \in \mathcal{S}$, there is some $\sim_\gamma \in \mathcal{S}$ such that $y \sim_\gamma x$ if and only if $y \sim_\alpha x$ and $y \sim_\beta x$.

The first condition guarantees that the structure $\mathcal{S}$ covers X, in the sense that there is some similarity relation which, for each $x \in X$, takes x to be similar to itself. The second condition can be understood in terms of the concept of a *restriction*. Formally, a restriction of a certain relation is just a relation that is a subset of that certain relation. So, a restriction of a similarity relation is one which is at least as strong or demanding as the latter. In these terms, the second condition requires that the family contain all the restrictions formed by finite intersections of its members: $\sim_\gamma = \sim_\alpha \cap \sim_\beta$. I.e., if it contains a relation of similarity in manner α, and one in manner β, then it contains a relation of similarity in manner α and β. Of these two, only the first involves any substantive constraint on the family of similarity relations, and a weak one at that, for the second can be satisfied merely by closing the set of similarity relations under finite intersections.

The domains of the similarity relations of a neighborly similarity space generate a canonical base for a neighborhood system, making that set into a pretopological space:[14]

Proposition 4. *The set of domains $\mathcal{B}_x$ for each $x \in X$ of a neighborly similarity space is a neighborhood base at x, i.e.,*
1. *if $D(x) \in \mathcal{B}_x$, then $x \in D(x)$; and*
2. *if $D_\alpha(x), D_\beta(x) \in \mathcal{B}_x$, then there is some $D_\gamma(x) \subseteq D_\alpha(x) \cap D_\beta(x)$.*

Definition 9. A *neighborhood system* at some $x \in X$ is a neighborhood base $\mathcal{B}_x$ for such that for any $U \in \mathcal{B}_x$, if $U \subseteq V \subseteq X$ then $V \in \mathcal{B}_x$.

14 See Čech (1966, Ch. III.14.B) for an exposition of pretopological spaces, called there closure spaces because a closure operator is taken as primitive (but is nonetheless definitionally equivalent to the neighborhood definition given here). Also, I omit the proofs of propositions 4 and 5, which follow immediately from the invoked definitions.

Definition 10. A *pretopology* on a set X is a set $\mathcal{T} \subseteq X \times \wp X$ consisting of a neighborhood system for each $x \in X$.

Proposition 5. *If $(X, \mathcal{S})$ is a neighborly similarity space, then the set $\bigsqcup_{x \in X}\{U \subseteq X : \exists D(x) \in \mathcal{S} : D(x) \subseteq U\}$ is a pretopology on X.*

The *induced pretopology* on a set (of models of a theory, say,) with a neighborly similarity space is that given in proposition 5, and a set equipped with a pretopology is called a pretopological space.

The definitions of several important concepts of pretopological spaces mirror those derived from similarity spaces. In the following, let a set X be equipped with a pretopology $\mathcal{T}$ of neighborhood systems $\mathcal{B}_x$.

Definition 11. The *closure operator* $\mathrm{cl}_{\mathcal{T}} : \wp X \to \wp X$ for the pretopological space $(X, \mathcal{T})$ is defined by $\mathrm{cl}_{\mathcal{T}}(A) = \{x \in X : \forall U \in \mathcal{B}_x, U \cap A \neq \emptyset\}$.

Definition 12. A sequence $\{x_n\}_{n \in \mathbb{N}}$ in X *converges to* $x \in X$ *(with respect to* $\mathcal{T}$*)*, denoted $x_n \to_{\mathcal{T}} x$, when for every $U \in \mathcal{B}_x$, there is some $N \in \mathbb{N}$ such that $x_n \in U$ if $n \geq N$.

Further, suppose that the pretopological space $(X', \mathcal{T}')$ has neighborhood systems $\mathcal{B}'_x$.

Definition 13. A function $f : X \to X'$ is *continuous (with respect to* $\mathcal{T}$ *and* $\mathcal{T}'$*) at* $x \in X$ when for every $U' \in \mathcal{B}'_{f(x)}$, there is some $U \in \mathcal{B}(x)$ such that $f[U] \subseteq U'$. It is *continuous (simpliciter)* when it is continuous at each $x \in X$.

As the astute reader will note, the definitions for these concepts in pretopological spaces are essentially the same as for their counterparts for similarity spaces; this justifies the notational equivocation between closeness and closure operators, and convergence (limits) and continuity in the two cases. The following proposition captures this equivalence between these similarity and pretopological concepts in the case of induced pretopologies.

Proposition 6. *Let $(X, \mathcal{S})$ and $(X', \mathcal{S}')$ be neighborly similarity spaces with induced pretopologies $\mathcal{T}$ and $\mathcal{T}'$, respectively. Then:*

1. *For any $A \subseteq X$, $\mathrm{cl}_{\mathcal{S}}(A) = \mathrm{cl}_{\mathcal{T}}(A)$.*
2. *For any sequence $\{x_n\}_{n \in \mathbb{N}}$ of X, $x_n \to_{\mathcal{S}} x$ if and only if $x_n \to_{\mathcal{T}} x$.*
3. *A function $f : X \to X'$ is continuous with respect to $\mathcal{S}$ and $\mathcal{S}'$ if and only if it is continuous with respect to $\mathcal{T}$ and $\mathcal{T}'$.*

Thus one can prove many statements about the similarity structure on a set by proving analogous statements about its canonical pretopology.

Many similarity structures naturally arising in the context of scientific theories satisfy a further condition:

Definition 14. A neighborly similarity space $(X, \mathcal{S})$ is *topological* when, for every $\sim_\alpha \in \mathcal{S}$, there is some $\sim_\beta \in \mathcal{S}$ such that if $y \sim_\beta x$ and $z \sim_\beta y$, then $z \sim_\alpha x$.

Being topological guarantees that for each similarity relation in the set, there is another at least as strong as it, which in particular contains the latter's transitive closure.[15] This condition thus formalizes the idea that for each similarity relation, one can always find another that is more restrictive—indeed, "twice" as much—having in particular all its domains contained in the domain of the first. It is the qualitative version of a sort of "$\epsilon/2$" principle, according to which, for each similarity relation, there is another whose domains are "half" as large.

Just as the similarity structure of a neighborly similarity space induces a pretopology, a that of a topological similarity space induces a topology.

Definition 15. A neighborhood base $\mathcal{B}_x$ is *topological* when, for any $D_\alpha(x) \in \mathcal{B}_x$, there is some $D_\beta(x) \in \mathcal{B}_x$ such that for all $y \in D_\beta(x)$, there is some $D_\gamma(y) \in \mathcal{B}_y$ such that $D_\gamma(y) \subseteq D_\alpha(x)$.

Proposition 7. *The set of domains $\mathcal{B}_x$ for each $x \in X$ of a topological similarity space $(X, \mathcal{S})$ is a topological neighborhood base at x, i.e.,*
1. *if $D(x) \in \mathcal{B}_x$, then $x \in D(x)$;*
2. *if $D_\alpha(x), D_\beta(x) \in \mathcal{B}_x$, then there is some $D_\gamma(x) \subseteq D_\alpha(x) \cap D_\beta(x)$; and*
3. *if $D_\alpha(x) \in \mathcal{B}_x$, then there is some $D_\beta(x) \in \mathcal{B}_x$ such that for all $y \in D_\beta(x)$, there is some $D_\gamma(y) \in \mathcal{B}_y$ such that $D_\gamma(y) \subseteq D_\alpha(x)$.*

Proof. The first two conditions follow immediately from the first two conditions for a neighborly set of similarity relations. For the third condition, pick any $D_\alpha(x) \in \mathcal{B}_x$, and note that the third condition for a basically neighborly set of similarity relations guarantees that there is some $\sim_\beta \in \mathcal{S}$ such that if $y \sim_\beta x$ and $z \sim_\beta y$, then $z \sim_\alpha x$. By logical exportation, for any $y \in D_\beta(x)$, if $z \in D_\beta(y)$ then $z \in D_\alpha(x)$. So the conclusion follows by letting $\gamma = \beta$. □

One can then define the induced topology on a set with a topological similarity structure in analogy with proposition 5. It is more traditional now, though, to understand a topology in terms of open sets rather than neighborhoods. According to this conception, a topology on X consists of a set of sets of X, its

15 Hence the other must be a restriction of the first, but this need not be strict, e.g., if $\sim_\alpha$ is already transitive.

open sets, that contains $\emptyset$ and X and is closed under arbitrary unions and finite intersections; see, e.g., Willard (1970) or Steen and Seebach (1978).

Proposition 8. *For any similarity structure on X whose domains form topological neighborhood bases $\mathcal{B}_x$ for each $x \in X$, $\{U \subseteq X : \forall x \in U, \exists D \in \mathcal{B}_x : D \subseteq U\}$ is a topology on X.*

Proof. See, e.g., Willard (1970, Theorem 4.5). □

The *induced topology* on a set with a topological similarity structure is that given in proposition 8.[16]

As the definitions (in terms of neighborhoods) of closure, convergence, and continuity are the same for topological spaces as they are for pretopological spaces, they cohere also cohere with those given by the similarity relations:

Proposition 9. *Let $(X, \mathcal{S})$ and $(X', \mathcal{S}')$ be topological similarity spaces with induced topologies $\mathcal{T}$ and $\mathcal{T}'$, respectively. Then:*
1. *For any $A \subseteq X$, $\mathrm{cl}_{\mathcal{S}}(A) = \mathrm{cl}_{\mathcal{T}}(A)$.*
2. *For any sequence $\{x_n\}_{n\in\mathbb{N}}$ of X, $x_n \to_{\mathcal{S}} x$ if and only if $x_n \to_{\mathcal{T}} x$.*
3. *A function $f : X \to X'$ is continuous with respect to $\mathcal{S}$ and $\mathcal{S}'$ if and only if it is continuous with respect to $\mathcal{T}$ and $\mathcal{T}'$.*

Thus, just as before, one can prove many statements about the similarity structure on a set by proving analogous statements about its induced topology.

One could continue to consider further conditions that a family of similarity relations might satisfy, but for current purposes the above will suffice. In focusing on similarity spaces and their induced (pre)topologies, therefore, I do not mean to suggest that these must capture the exact level of structure so as to be the "right" tools to analyze similarity relations amongst models of a theory. But, in the mathematized sciences, topological structure is natural and ubiquitous and one must start somewhere. Furthermore, the following applications will reveal how fruitful even this structure can be for analyzing some traditional questions in the philosophy of science.

16 A topological family of similarity relations also exactly satisfies the axioms for the base of a quasiuniformity on X (Steen and Seebach, 1978, p. 9, 37–38). Compared with the definition given by Steen and Seebach, I have reversed the role of the first and second places in the relation, but nothing substantial turns on this.

3 Intertheoretic Reduction

3.1 The Idea of Reduction

In order for our space agencies to send rockets with people and instruments into orbit, and complicated, delicate, and expensive space probes to other planets and the distant reaches of our solar system, they must calculate and predict with great accuracy these objects' trajectories through space. Yet in this task, they do not typically use Einstein's theory of general relativity, our currently best theory of space, time, and gravitation. They use instead the older and simpler Newtonian theory, and to great effect, despite the quite different structure for space-time which, on a realist construal, that theory ascribes to our world. Can one explain the success of Newton's theory in terms of Einstein's, at least for the ranges of phenomena for which it has been profitably applied?

The question of intertheoretic reduction between two theories, with which I am concerned here, is just this sort of question. It demands an explanation for the success of a theory, one we typically (but not necessarily) know to be empirically or explanatorily inadequate for some phenomena.[17] Thus it is a type of *synchronic* reduction, at least in the sense that it is a relation that holds between two theories rather than being a process or activity—*diachronic* reduction—that scientists or scientific communities undergo (van Riel and Van Gulick, 2016, §2.1). As a matter of terminology, the philosophical literature describes this question as concerning the reduction of *Newton's theory by Einstein's* (or, in other words, whether Einstein's theory *reduces* Newton's), emphasizing the explanation of the former by the latter, while the physics literature on the subject describes it as concerning the reduction of *Einstein's theory to Newton's*, emphasizing the circumstances under which the structure of the former is simplified to the latter. In the end nothing of substance really turns on this difference, but I shall adopt the philosophers' terminology since my concern here is the explanation of the success of Newton's theory by Einstein's.

Because this success is empirical, it will suffice to answer the explanatory question by showing that, in the circumstances in which it is successfully applied, Newton's theory gives predictions or empirical descriptions of phenomena

17 This is close to what Sklar (1967, p. 112) calls *explaining away* a theory, as how one might attempt to use quantum mechanics to explain "why Newtonian mechanics *seemed* to be correct; why it met with such apparent success for such a long period of time and under such thorough experimental scrutiny." However, I am here more concerned with empirical adequacy rather than "correctness," as such adequacy is not necessarily lost in a certain domain of application even if it is lost in another.

sufficiently similar to those of Einstein's: the success of Newtonian theory is the explanandum, its similarity (in those contexts of success) with general relativity the explanans. By adapting the apparatus of the previous section, one can attempt to capture this sort of similarity with the relevant similarity structure on the joint class of models of both theories, making this joint class into a similarity space. Of course, this apparatus would still be viable even if extra-empirical aspects of the theories were taken to be relevant for the explanatory question, or if one focused on features to be hypothetically, but not actually, explained. In this sense, it also provides a framework for understanding intertheoretic relationships of the same form—ones that focus on relevant ways in which these theories are similar or dissimilar, even if those ways do not concern actual empirical success (Nickles, 1973; Wimsatt, 1976; Batterman, 2002).

In what follows, I first (in section 3.2) describe the general framework for understanding intertheoretic reduction using similarity structure. Next, in section 3.3, I apply this framework to the aforementioned case of the reduction of Newton's theory of gravitation by Einstein's theory of general relativity, describing some partial results and open questions. Last, in section 3.4, I compare this framework for reduction with three others, that of Nagel (1961) and its descendants (Dizadji-Bahmani et al., 2010), the limit schema of Nickles (1973), and that of the structuralists (Schmidt, 2014), respectively, showing how, in a certain sense, it encompasses them.

3.2 Similarity Structures and Reduction

Consider a similarity space $(X, \mathcal{S})$ with closeness operator cl. In particular, X might be (or simply contain) the union of the models T and T' of two theories describing the same—or at least overlapping—ranges of phenomena. Often, T will be a newer, more empirically successful theory, while T' will be an older theory with a more restricted domain of success. In this case, the collection of similarity relations will be determined from all those properties of the models which make a difference to an explanatory question, e.g., about the empirical descriptions of phenomena that each model determines. Then the sorts of reductive relationship that holds between T and T' can be classified using the following terminology.

Definition 16. Let $(X, \mathcal{S})$ be a similarity space with closeness operator cl. If $T, T' \subseteq X$, then:

1. T *completely reduces* T' when $T' \subseteq \mathrm{cl}(T)$.
2. T *partially reduces* T' when $T' \cap \mathrm{cl}(T) \neq \emptyset$.

3. T *merely partially reduces* T' when T partially but not completely reduces T'.
4. T *fails to reduce* T' when T does not partially reduce T'.

If T completely reduces T', then the models of T' are arbitrarily similar to models of T *in the respects demanded by* $\mathcal{S}$. Thus, if $\mathcal{S}$ captures the similarity of empirical descriptions needed to explain the success of T', then the complete reduction of T' by T indicates that the potential success of *each* model of T'—each possible world or situation that the theory affords—can be explained in terms of models of T.

To take a toy example, suppose that the only relevant prediction that theory T' makes is that the value of a certain real parameter is 0. The theory T, on the other hand, only has models in which this parameter is at least $c/2 > 0$, but measurements of this common parameter at this time only have an accuracy of $\pm c$. Suppose that T is accepted for independent reasons, yet under certain circumstances T' seems to be sufficiently empirically adequate. Then one can model the relationship between the theories as follows. Consider a similarity space $(\{0\} \cup (c/2, \infty), \mathcal{S}_c)$, where $\mathcal{S}_c$ consists of all the similarity relations $\sim_\epsilon$ whose domains are $D_\epsilon(x) = \{y \in \mathbb{R} : |x - y| < \epsilon\}$ for $\epsilon \geq c$. Any model x has a smallest domain $D_c(x)$, so in particular $D_c(0) \cap (c/2, \infty) \neq \emptyset$.[18] It then follows from the definition of the closeness operator that $T' \subseteq \mathrm{cl}(T)$, i.e., T completely reduces T'. This shows that whenever the circumstances are such that T' is successful in describing the parameter of interest, there is a model of T that does so, too, because differences in values of the parameter of less than c are not relevant, due to the postulated limited measurement precision.

If, on the other hand, some theory T merely partially reduces some theory T', then some model of T' is *not* sufficiently similar to those of the newer, more successful theory T. This may not ultimately spell the failure of the explanation of the success of T', however. Let the *domain of successful application of* T' be the models $S' \subseteq T'$ that have found empirical success according to criteria used to generate $\mathcal{S}$ (Ludwig and Thurler, 2006; Ehlers, 1986). (This set is not in general determined by the theory itself or the similarity relations placed upon its models, but rather by the interface of theory and experiment.) If T completely reduces S', the success of the older theory would be explained even though T merely partially reduces T'. In such cases, a merely partial reduction can be evidence for delimiting the "physically reasonable" models of the older theory in

18 Similarity spaces with the sort of feature can be neighborly, but not topological, in general: the "$\epsilon/2$" principle, which being topological demands, prevents there from being a relation in the space's similarity structure with a smallest (non-trivial) domain.

light of the newer theory, as proponents of quantum gravity sometimes do with the quantum theory and general relativity (Sorkin, 1995).

If T merely partially reduces S', however, then we have a case of *Kuhn loss*: a successful model of the older theory which is not sufficiently similar to any model of the newer theory (Kuhn, 1996; Post, 1971). Accordingly, the success of the older theory would not be explained by the newer theory. Whether a reductive relationship holds between the models of two theories therefore can depend on the domain of application of the theory to be explained.

It also depends, of course, on the nature and specificity of the similarity relations between them. If these relations describe differing degrees of similarity in empirical descriptions, then they might well depend upon a contingent or historically particular degree of measurement precision and the types of measurements available or descriptions deemed relevant. Because of this, the status of the reductive relationship between different theories may change over time: two theories that were sufficiently similar in their shared domain of application may come apart as measurements become more precise or more types of measurement become possible. In this sense, although the concept of reduction investigated here is not of a process occurring in time (i.e., diachronic reduction), the relevant reduction relation to examine can depend on empirical facts or data known at a particular time.

3.3 An Example: Theories of Gravitation

As an example, from which I will continue to draw in the remainder, consider the case of Einstein's theory of general relativity and Newton's theory of gravitation.[19] In order to compare the two theories, it is helpful to construct a general framework for space-time models that encompasses both. I shall in particular adapt the *frame theory* (Ehlers, 1981, 1991, 1998), which takes a space-time to be an object $(M, t_{ab}, s^{ab}, \kappa, \nabla_a, T^{ab})$, where M is a smooth, paracompact real manifold of events, t_{ab} and s^{ab} are, respectively, the temporal and spatial metrics, $\kappa \in \mathbb{R}$ is the so-called causality constant, ∇_a is a torsion-free affine connection, and T^{ab} is the stress-energy tensor representing the presence of energy and momentum in the space-time. Given any curve $\gamma : I \to M$, the temporal metric determines the *duration* γ as $\int_I \sqrt{t_{ab}\xi^a\xi^b}\, ds$ if it is *timelike*, i.e., its tangent vector ξ^a satisfies $t_{ab}\xi^a\xi^b > 0$. Similarly, the spatial metric determines the *length* of γ as $\int_I \sqrt{s^{ab}\xi_a\xi_b}\, ds$ if it is *spacelike*, i.e., when ξ_a is covector satisfying

19 This subsection is largely based on Fletcher (2014, Ch. 3).

$s^{ab}\xi_a = \xi^b$ and $s^{ab}\xi_a\xi_b > 0$ on $\gamma[I]$. (While in general there will be many such covectors, the length of a spacelike curve is independent of the choice among them (Malament, 2012, p. 255).)

These objects satisfy two further conditions:

Causality $t_{ab}s^{bc} = -\kappa\delta^c_a$, where δ is the Kronecker delta tensor.
Compatibility $\nabla_a t_{bc} = 0$ and $\nabla_a s^{bc} = 0$.

The causality condition ensures that the timelike and spacelike directions at each event, as described by the tangent vectors there, are disjoint. The compatibility condition ensures that the standards of constancy determined by the affine connection match those of the temporal and spatial metrics: a quantity or field is changing with respect to the former whenever it is changing with respect to one of the latter.

The frame theory is not intended as a physical theory in its own right, but rather as a theoretical framework in which one can describe both relativistic and non-relativistic models of space-time and gravitation. One can see that the models of general relativity (M, g_{ab}), where M is the space-time manifold of event and g_{ab} is a Lorentz metric, are models of the frame theory as follows:

- the temporal metric is just $t_{ab} = g_{ab}$;
- the spatial metric is $s^{ab} = -\kappa g^{ab}$, where $\kappa = 1/c^2$ and c is the speed of light;
- ∇_a is the unique Levi-Civita connection compatible with g_{ab}; and
- the stress-energy tensor is constrained by the Ricci tensor R_{mn} associated with ∇_a according to a version of Einstein's equation,

$$T^{ab} = \frac{1}{8\pi G\kappa^2}\left(s^{am}s^{bn} - \frac{1}{2}s^{ab}s^{mn}\right)R_{mn},$$

where G is Newton's gravitational constant.[20]

To show that models of Newtonian gravitation are models of the frame theory requires a bit more work. First, I will exhibit how models of Newton-Cartan theory are models of the frame theory. Newton-Cartan theory is also known as “geometrized” Newtonian gravitation because, like general relativity, it represents the phenomenon of gravitation as a manifestation of curvature, but in the context of a non-relativistic space-time structure. Second, I will record how the models of standard Newtonian gravitation are, in a suitable sense, empirically equivalent with a subclass of those of Newton-Cartan theory.

20 For more on the formalism of general relativity, see, e.g., Hawking and Ellis (1973) or Malament (2012, Ch. 2).

Models of Newton-Cartan theory are of the form $(M, t_{ab}, s^{ab}, \nabla_a, T^{ab})$, where t_{ab} is a temporal metric of signature $(+, 0, 0, 0)$, s^{ab} is a spatial metric of signature $(0, +, +, +)$, ∇_a is a connection compatible with t_{ab} and s^{ab}, and the stress energy tensor T^{ab} is constrained by the so-called geometrized Poisson equation, $R_{ab} = 4\pi\rho t_{ab}$, where $\rho = t_{ab}T^{ab}$ is the mass density, and R_{ab} is the Ricci tensor associated with the connection. Additionally, the temporal and spatial metrics satisfy the condition $t_{ab}s^{bc} = 0$. Hence, these are models of the frame theory with $\kappa = 0$. This ensures that the structure of a Netwon-Cartan spacetime is non-relativistic, i.e., the manifold of events foliates into hypersurfaces of space at a time.[21]

Now, a more traditional formulation of Newtonian gravitation theory would take as its models structures of the form $(M, t_{ab}, s^{ab}, \nabla_a, T^{ab}, \phi)$, where the temporal and spatial metrics are as in the Newton-Cartan case, but the connection is taken to be flat, and the constraints on the stress-energy tensor are instead connected with a new field, the gravitational potential ϕ, according to Poisson's equation, $s^{ab}\nabla_a\nabla_b\phi = 4\pi\rho$. But there are systematic connections between the two theories. The one of most relevance for present purposes is expressed in the following proposition:

Proposition 10. *Let $(M, t_{ab}, s^{ab}, \nabla_a, T^{ab}, \phi)$ be a model of Newtonian gravitation. Then there is a unique connection ∇' such that:*

1. *$(M, t_{ab}, s^{ab}, \nabla'_a, T^{ab})$ is a model of Newton-Cartan theory;*
2. *for all timelike curves in M with tangent vector field ξ^a, $\xi^n\nabla'_n\xi^a = 0$ iff $\xi^n\nabla_n\xi^a = -s^{ab}\nabla_b\phi$; and*
3. *the Ricci curvature tensor R'_{ab} associated with ∇'_a satisfies the geometrized Poisson equation, $R'_{ab} = 4\pi\rho t_{ab}$.*

Every model of Newtonian gravitation theory is thus empirically equivalent to a model of Newton-Cartan theory, in the sense that all features of the former are the same, except all and only acceleration due to gravity has been replaced by geodesic motion in a space-time whose Ricci curvature satisfies the geometrized Poisson equation.[22] In general there will be many models of Newtonian grav-

21 For more on the formalism of Newton-Cartan theory, see, e.g., Malament (2012, Ch. 4).

22 In more detail, if one takes models of Newtonian gravitation whose gravitational potentials and connections are related by gauge transformations $\phi \mapsto \phi + \psi$ with $s^{ac}s^{bd}\nabla_b\nabla_d\phi = 0$ and $\nabla \mapsto (\nabla', t_{bc}s^{ad}\nabla_d\psi)$ to have the same representational capacities, then Newtonian gravitation is theoretically equivalent to a sector of Newton-Cartan theory, where this equivalence is understood as an equivalence of categories preserving empirical content (Weatherall, 2016).

itation associated in this way with a model of Newton-Cartan theory, but the point is that if models in the frame theory are related by similarity relations which depend only on empirical content, then exhibiting a reduction of general relativity to Newton-Cartan theory suffices to do so for Newtonian gravitation. Thus, the next step is to formulate the relevant similarity relations in terms of empirical similarity on the models of the frame theory.

I shall consider two cases. Both take as the relevant criteria of comparison pointwise similarity of any scalar fields constructible from the temporal and spatial metrics, the stress-energy tensor, their derivatives up to order k, and frame fields, the lattermost representing the idealized measuring instruments associated with observers. These fields represent possibly empirically measurable quantities in the spacetime. Both cases also restrict attention to when the spacetime models under comparison are defined on diffeomorphic (really, the same) manifolds of events. They differ only on which sorts of regions in which these fields should be compared. In the first case, these will only be compact (i.e., bounded) regions of spacetime. In the second case, these will be any regions of spacetime.

To formalize these ideas, consider any frame field $\{\overset{0}{e}{}^a, \overset{1}{e}{}^a, \overset{2}{e}{}^a, \overset{3}{e}{}^a\}$, letting h_{ab} be the Riemannian metric which is the inverse of $h^{ab} = \sum_{i=0}^{3} \overset{i}{e}{}^a \overset{i}{e}{}^b$. For any tensor field $F^{a_1 \ldots a_n}_{b_1 \ldots b_m}$ defined where h_{ab} is, one can define its h-norm field as

$$|F|_h = \left(h_{a_1 c_1} \cdots h_{a_n c_n} h^{b_1 d_1} \cdots h^{b_m d_m} F^{a_1 \ldots a_n}_{b_1 \ldots b_m} F^{c_1 \ldots c_n}_{d_1 \ldots d_m}\right)^{1/2}.$$

The h-norm field of $F^{a_1 \ldots a_n}_{b_1 \ldots b_m}$ is its Frobenius norm with respect to h—a field on M that describes the magnitude of $F^{a_1 \ldots a_n}_{b_1 \ldots b_m}$ by the square root of the sum of its squared components.

The h-norm of the *difference* between two tensor fields $F^{a_1 \ldots a_n}_{b_1 \ldots b_m}$ and $F'^{a_1 \ldots a_n}_{b_1 \ldots b_m}$, denoted $|F' - F|_h$, therefore measures how dissimilar the two are. So an observer (or observers) represented by a frame field determining a Riemannian metric h might judge two temporal metrics, say, to be similar when the h-norm of their difference does not exceed some value ϵ on a region S. And that observer might judge two spacetimes to be similar when not just their temporal metrics, but all their tensorial structure, is judged to be similar in this way.

In this vein, consider the following similarity relations, defined for each h determined by a frame field, $\epsilon > 0$, and compact region $C \subseteq M$:

$$(M, t'_{ab}, s'^{ab}, \nabla'_a, T'^{ab}) \sim_{h,\epsilon,S} (M, t_{ab}, s^{ab}, \nabla_a, T^{ab})$$
$$\text{iff}$$
$$\sup_C |t' - t|_h < \epsilon \;\&\; \sup_C |s' - s|_h < \epsilon \;\&\; \sup_C |\nabla' - \nabla|_h < \epsilon \;\&\; \sup_C |T' - T|_h < \epsilon,$$

When closed under intersection, the resulting similarity structure is topological for the models of the frame theory, and its induced topology is known as the (C^0 product) compact-open topology.[23]

Combining propositions 9 and 2 with the definition of reduction and the observations of Fletcher (2014, p. 65) yields the following proposition.

Proposition 11. *Suppose the models of the frame theory are equipped with a similarity structure, yielding a topological similarity space whose induced topology is the compact-open topology. Then the models of general relativity partially reduce the models of Newton-Cartan theory, hence the models of Newtonian gravitation.*

This proposition takes advantage of the fact that, for partial reduction, it suffices to show that there is a sequence of models x_n of general relativity that becomes arbitrarily similar to a model x of Newton-Cartan theory, i.e., the sequence converges to x in the topology induced by the similarity structure. This can be done with a sequence of Minkowski spacetime, which converge to Galilean spacetime. It is an open question whether the reduction is merely partial or in fact complete with respect to this or some other relevant similarity structure.

If the compact regions C in the above definition are then broadened to include *any* regions $S \subseteq M$, the result is still a topological similarity space but one whose induced topology, the open topology, is in general finer—has more open sets—than the compact-open topology. The nature of the relationship between general relativity and Newtonian gravitation is rather different with respect to this more discerning similarity structure.

Proposition 12. *Suppose the models of the frame theory are equipped with a similarity structure, yielding a topological similarity space whose induced topology is the open topology. Then the models of general relativity fail to reduce the models of Newton-Cartan theory, hence the models of Newtonian gravitation.*

Which of the two similarity structures is more apt to explaining the success of Newtonian gravitation? While the second is more comprehensive in its criteria of comparison between the theories, there is good reason to believe that this

23 I have made a few simplifications in the presentation of the compact-open topology. First, the ambiguity of the third conjunct can be resolved by noting that differences between affine connections can be assessing according to their (1,2)-tensorial connecting fields. Second, I have supposed that the derivatives of these tensor fields are not relevant for this context, but they can be easily included by including more conjunctions of bounded suprema of derivatives (taken with the connection compatible with h) of differences between these fields. See Fletcher (2014) for these details.

comprehensiveness extends far beyond the domain of successful application of the theory. Notably, it has not been so successful at cosmology. Rather, the success of Newton's theory has been confined to applications for bounded regions of spacetime, such as the observed history of a galaxy, various satellites within our solar system, and the falling of terrestrial objects. The first structure includes similarity relations which depend on these, but not on those which, surveying unbounded regions of spacetime, do not make a difference in explaining Newtonian gravitation's success. Thus this first structure is presently the more contextually appropriate of the two for the question of intertheoretic reduction.

3.4 Comparisons with Other Accounts of Intertheoretic Reduction

In this section I contrast briefly the present framework for intertheoretic reduction based on similarity structure with three others, those of Nagel (1961), Schaffner (1967), and their followers, and the structuralists (Schmidt, 2014). For the former, I will focus on the Generalized Nagel-Schaffner (GNS) model as articulated by Dizadji-Bahmani et al. (2010). In this model, a theory T is supplemented with auxiliary assumptions (such as boundary conditions) from which a more constrained theory T^* is deduced. Then one adds so-called "bridge principles" to T^* that unify its vocabulary with that of another theory T', yielding T'^*. Then, if T'^* makes at least as accurate predictions as T' yet is still "strongly analogous" to T', then T is said to reduce T'.

The similarity structure-based framework I am proposing here is aligned with the GNS model in some ways and opposed to it in others. Somewhat like how the GNS model uses bridge principles, reduction in terms of similarity structures demands that there be a common basis—say, of empirical descriptions—on which the two theories are compared. However, in the GNS model every term of the two theories must be bridged for there to be a complete reduction, while in my framework this is only necessary for the terms relevant (i.e., those that make a difference) to the explanatory context arising from the models considered. The function of strong analogy in the GNS model is comparable to that of the similarity structure in mine: both provide a contextual way of comparing two theories, although in the present framework the structure of similarity is both more concrete and richer than in the GNS model. Thus, in contrast to Dizadji-Bahmani et al. (2010, p. 409) and Butterfield (2011a, p. 939), who council giving up on the project of a precise characterization of analogy, similarity structure provides as much precision as can be made about the models of the theories themselves and which differences between them make a difference.

Auxiliary assumptions are often required for the GNS model to guarantee that T' (or T'^*) can be deduced, which must therefore always be consistent with the deducing one. But in the similarity structure-based account, the relationship between the theories under consideration is not one of deduction, even with aspects of similarity are taken into account, and accordingly the supplementation of the reducing theory T with auxiliary assumptions is no longer needed. Instead, one *exhibits a relationship* between the models of two theories based on a contextually determined similarity structure. In this sense, rather than adding to what is assumed, one is subtracting by considering a broader class of possibilities encompassing both theories. Indeed, the reducing theory need not be consistent with the reduced theory at all, as the case of the (at least partial) reduction of Newtonian gravitation by general relativity demonstrates.[24] In the (very!) special case in which the models of the reduced theory can be shown to be models of the reducing theory—in the sense of having isomorphic properties relevant to the context—the GNS framework itself can be construed (cf. proposition 1) as a special case of the similarity-structure account.

After the development of the Nagelian account was underway, others began to propose alternatives. Nickles (1973) contrasted this account with one that describes the reduction of one theory by another as the result of applying a "limit operation" (or some other operation O that need not be purely logical) on the latter to arrive at the former. Schematically, one might write $\lim T = T'$ or more generally $O(T) = T'$, with the operation perhaps depending on a "fundamental parameter" of T (Batterman, 2016, §2). What kinds of operations legitimately fall under this class, though? Indeed, one of the criticisms of this approach has been that what they are supposed to be or how they are to be delimited is quite vague (Schaffner, 1976; Rosaler, 2015). Consequently, various authors—e.g., Butterfield (2011b) and Rosaler (2015)—have attempted to make this precise—as a limit of models or of approximation of observed trajectories of a dynamical system, respectively. In particular, Rosaler focuses on the case of dynamical systems at different "levels" with a shared time parameter, using a so-called "bridge map" to compare the induced dynamics of one theory inside another, concluding that reduction occurs when the induced trajectory in state space is approximately the same as the trajectory given by the dynamics in that theory.[25]

24 Now, Dizadji-Bahmani et al. (2010) do remark that consistency only holds between T and T'^*, not necessarily with T', but the point is that there may be no particular "strongly analogous" theory consistent with T that is deduced at all.

25 While Rosaler (2015, p. 67) acknowledges that his theory has many affinities with Nickles's schema, he sees it as well to be in the spirit of the Nagel model. What's important

As with the GNS model, the framework based on similarity structures has both points of alignment and difference with these precise versions of Nickles's schema. According to propositions 1 and 2 and definition 16, the convergence of a sequence of models of one theory to a model of another implies that the former theory reduces the limit model. Hence, if this holds for all models of the other theory, that theory is reduced, too. In the case of Rosaler's account, if one assumes that the bridge maps preserve the empirical predictions (or whatever other features are contextually relevant) of the models of the theory they act upon, then his concept of reduction is also a case of reduction in the similarity structure-based approach much like that of the toy example on page 3.2.

The main significant differences with these are twofold. First, similarity structure-based approach is quite more general. As proposition 2 demonstrates, sequential convergence is sufficient but not necessary for reduction; there may be models of one theory that are arbitrarily similar to those of another in all the relevant ways, but which are not the limits of models of the latter. And unlike the theory of Rosaler (2015), the similarity structure-based approach is not restricted to dynamical systems with a shared time variable: it is general enough to apply to any models or theories whose models can be described with sufficient precision.

The second important difference is that—perhaps surprisingly—essentially none of authors who have engaged with variations of Nickles's schema have commented on the essential role that the choice of topological (or, more generally, similarity) structure plays in describing the relationship between theories as represented by their models. While Butterfield (2011b, p. 1075) seems to acknowledge that topological structure is the basis for the definition of the convergence of a sequence in the general case, he does not address how this structure is determined or its justification. Similarly, while Rosaler (2015) acknowledges that the degree of approximation needed for reduction of dynamical systems theories must be contextual, he assumes, without argument, that a Euclidean distance function on the systems' state space is the relevant way to determine that approximation. As this review emphasizes, though, and as I previously described in detail in the case of theories of gravitation (Fletcher, 2016), it is essential to show how the structure picked out captures the all and only the features relevant for answering the explanatory question that reduction asks. The framework of similarity structure makes this task central.

---

for my comparison here is not so much the classification of Rosaler's theory as its points of alignment and difference with the present similarity structure-based approach. See also Atmanspacher and beim Graben (2009) and references therein for a similar apparatus, but focusing more on defining emergence rather than reduction.

Finally, I shall compare the present framework with that of the structuralists. As with the others, there are points of similarity and difference. It is more difficult to characterize univocally what the structuralist position is on reduction because of the heterogeneity of the positions advocated by the authors described by that label (Schmidt, 2014). That said, it is fair to say that the structuralist position of reduction aims to characterize it in terms of showing how the models of one theory approximate those of another, with the notion of approximation formalized in terms of structure on those models, so-called "inaccuracy-sets." On this feature, there is alignment: both the structuralist and similarity structure-based approach use structure on the models of theories to compare them and give a definition of reduction relations. However, inaccuracy-sets are assumed to satisfy more formal conditions than similarity relations—in particular, uniform structures[26]—and are to encode some concept of empirical approximation (Mayr, 1981). Although this case is very important, similarity relations are not restricted to them. Also, analogously with the elaborations on Nickles's limit schema, the structuralist approach took the theory itself to require a single such uniform structure, while I allow for the similarity structure to vary by context.[27] In sum, the structuralist position also places structure on the models of a theory in service of answering questions about intertheoretic reduction, but this structure encodes only empirical descriptions, is non-contextual, and is stronger (assumes more) than the present approach.

4 Emergent Properties

4.1 The Concept of Emergence

O'Connor and Wong (2015) aptly remark: "Emergence is a notorious philosophical term of art," one with a variety of conflicting usages and definitions. The sort of emergence of concern here is that of one or more properties of a system or state of affairs, as described by one theory, from that provided by another

26 A base for a uniform structure is obtained by a topological family of similarity relations, each of which is symmetric.

27 Moulines (1980) allows different subsets of a single uniform structure to be applicable in different contexts, but this only lets the *degree* of similarity vary by context, rather than the type. Moulines (1976) also allows for two uniform structures, one on the possible models and another for the 'partial' possible models, which contain only non-theoretical terms, but these are required to be strictly compatible, so this is not really contextualist about similarity.

theory. Whether this is metaphysical or merely epistemological emergence will depend on whether one's (contextually appropriate) attitude toward the theories in question is more realist or anti-realist, but that question can be set aside for present purposes, as its answer does not substantively affect the formal features of the analysis. Further, just as the notion of reduction that I developed in §3 was synchronic in the sense of treating the relationships between theories rather than a dynamical or historical process, the notion of emergence here is also synchronic in the same way: it will describe the relationships of models of theories and novel properties thereof, not how properties of systems or states of affairs arise in time.

Informally, on the present account, a property of a model is emergent with respect to another class of models just when it is comparatively novel. This novelty can come in degrees, so I distinguish four different types of emergence partially ordered in strength. The weakest, *weak emergence*, requires only mere non-identity of the property—or the value of that property, if it is not a simple predicative property—of the model with any of the properties of the models in the comparison class. One might require not just mere non-identity, though, but a sort of unexpectedness or inexplicability. This can be formalized along at least two directions. The *strong emergence* of a property of a model requires that the property must also be not sufficiently similar to the properties of the models of the comparison class—it is unexpected because it is not even similar (in the relevant ways) to the properties available for consideration from the comparison class. This requires similarity structure on the values that a property can take on. The *non-reductive emergence* of a property of a model requires that the property must also be non-identical with the corresponding properties of the models arbitrarily similar to those in the comparison class—it is inexplicable in the sense that . This requires similarity structure on the joint collection of models. Finally, *radical emergence* is just the conjunction of strong and non-reductive emergence: the emergent property is not even sufficiently similar to the properties of the models arbitrarily similar to those in the comparison class. Accordingly, this requires similarity structure on both the space of values that a property can take on and the joint collection of models themselves. All of these concepts readily generalize from applying only to properties of individual models to the properties of sets of models.

Contrary to common usage, two of these concepts of emergence—the weak and strong varieties—will not only be *compatible* with reduction as developed in §3, but will often be a *consequence* of it. Indeed, after describing the more formal definitions of emergence using similarity structure in section 4.2, I formulate and prove some propositions to this effect.[28] This is because, in the present account, the one is not defined in terms of the failure of the other. Section 4.3 takes up the example of general relativity and Newtonian gravitation from section 3.3 again, arguing that absolute simultaneity is an emergent according to the different similarity structures previously considered. Finally, section 4.4 compares the present account of emergence in terms of similarity structure to two others that seem to me to be most related: that of Primas (1998), Bishop and Atmanspacher (2006), and Atmanspacher and beim Graben (2009), and that of Butterfield (2011a,b, 2014) and Bouatta and Butterfield (2012, 2015).

4.2 Similarity Structures and Emergent Properties

To begin, let X be a set of models represent some possible states of affairs, and consider a possible property that some of these models might represent. One can represent it with a certain map on X as follows.

Definition 17. A *valuation of a property* on a collection of models X is a (perhaps partial) map $\nu : X \nrightarrow V$, where V is called its space of *(property) values*. Further, if $x \in X$ is (not) in the domain of ν, then the property represented by ν is said to be *(not) definable for* x.

Philosophers will already be familiar with (classical) propositional properties, those which either obtain or not for a given model. For example, "being red" is a propositional property as applying to (some model of) everyday objects. Such properties have the Boolean domain $\{\top, \bot\}$ as their values. But properties as I have considered them here can also include quantities, such as "having (a certain) mass," whose space of values might be the real interval $(0, \infty)$, or even more complex structures.

Now suppose that X has been equipped with a similarity structure $\mathcal{S}$ with closeness operator $\mathrm{cl}_{\mathcal{S}}$ and a valuation $\nu : X \nrightarrow V$, whose space of property values V has also been equipped with a similarity structure $\mathcal{V}$ with closeness operator $\mathrm{cl}_{\mathcal{V}}$.

Definition 18. With respect to the models $A \subseteq X$, the values $\nu[B]$ of a collection of models $B \subseteq X$ are said to be:

28 I am not the first to suggest the compatibility of emergence with reduction, when the two are construed appropriately. See, for example, Wimsatt (1997, 2000), Primas (1998), Bishop and Atmanspacher (2006), and Butterfield (2011a,b), some of whose positions I consider in more detail in section 4.4.

1. *weakly emergent* when $\nu[B] \nsubseteq \nu[A]$, and *completely weakly emergent* when in addition $\nu[B] \cap \nu[A] = \emptyset$;
2. *strongly emergent* when $\nu[B] \nsubseteq \mathrm{cl}_{\mathcal{V}}(\nu[A])$, and *completely strongly emergent* when in addition $\nu[B] \cap \mathrm{cl}_{\mathcal{V}}(\nu[A]) = \emptyset$;
3. *non-reductively emergent* when $\nu[B] \nsubseteq \nu[\mathrm{cl}_{\mathcal{S}}(A)]$, and *completely non-reductively emergent* when in addition $\nu[B] \cap \nu[\mathrm{cl}_{\mathcal{S}}(A)] = \emptyset$; and
4. *radically emergent* when $\nu[B] \nsubseteq \mathrm{cl}_{\mathcal{V}}(\nu[\mathrm{cl}_{\mathcal{S}}(A)])$, and *completely radically emergent* when in addition $\nu[B] \cap \mathrm{cl}_{\mathcal{V}}(\nu[\mathrm{cl}_{\mathcal{S}}(A)]) = \emptyset$.

Any of these types of emergence is said to be *partial* when it is not complete.[29]

Logically, radical emergence entails the other three types of emergence, while weak emergence is entailed by the other three. In general, neither strong emergence nor non-reductive emergence entails the other. And in each of these cases, emergence can obtain either because the property valuations are definable for the models of A but different for the models of B, or they are not definable for the models of A.

In order for values of a property of models of B to be weakly emergent, no similarity structure on the space of models or property values is needed. The weakly emergent values of a property of models of B with respect to A are just those that are at least partially novel with respect to those of the models of A. Strong emergence obtains when values of a property of models of B are not only different from those of models of A, but also from all those arbitrarily similar to those values.[30] Thus similarity structure on the space of values of a property valuation is necessary for them to be strongly emergent. Non-reductive emergence, by contrast, requires only similarity structure on the models of X. It obtains when the values of a property of the models of B are novel with respect to those of the models of A and all the models arbitrarily similar to those in A. Finally, values of a property of models of B are radically emergent just when they are novel with respect to those even arbitrarily similar to the values of

29 Beyond these latter three, one can define infinitely many concepts of emergence based on iterations of the closeness operators: with respect to the models $A \subseteq X$, the values $\nu[B]$ of a collection of models $B \subseteq X$ can be said to be (n, m)-*emergent* when $\nu[B] \nsubseteq \mathrm{cl}^n_{\mathcal{V}}(\nu[\mathrm{cl}^m_{\mathcal{S}}(A)])$. However, if the similarity structures on the spaces of models and property values are topological, then their closeness operators are in fact topological closure operators, which are idempotent. In this case, the above four concepts are the only ones that are distinct.

30 Thus these precise characterizations of weak and strong emergence ought to mollify skeptical concerns that the criterion “novelty” is too vague without it being defined in terms of other criteria for emergence (Teller, 1992).

the models arbitrarily similar to those of A. Radical emergence thus requires similarity structures both on the property value space and the space of models.

There are several important features of the present graded account of emergence worth mentioning. First, it is relational, holding between the property values of some models and some other collection of models of an "appropriate comparison class" (Butterfield, 2011a, p. 922) from which the question of novelty or unexpectedness is posed. Changing this class may therefore change whether a property is emergent. This is as expected: if emergent properties are to be novel or unexpected, they can only be so with respect to such a class of expected properties.

Second, whether a set of values of a property are emergent in any sense stronger than weak depends on similarity structure, which is determined relative to a context of investigation. This shows how this sense of emergence can be sensitive to both the epistemological and pragmatic situation of an investigator, as the family of similarity relations determined (whether on the property value space or the space of models) depends on what makes a difference for the kinds of explanatory questions being asked.

Third, that certain values of a property are emergent does not imply that they are interesting (Teller, 1992). Whether a property is interesting, after all, generally depends on its implications for other sorts of inquiries and activities beyond its mere novelty or unexpectedness. One might nevertheless anticipate it to be somewhat likely that strongly emergent values of a property are interesting, provided the comparison class of models are informative about the values of the property as well. If the similarity structure on the space of property values is properly selected in a certain context of investigation, it will track differences that could make a difference to that context. Because strongly emergent values of a property are not even (closely) similar to those holding for the models of their comparison class, they are different in ways that could make a difference.

Fourth, both weak and strong emergence, but not non-reductive emergence, are compatible with complete reduction as described in the previous section. That is, it's possible for one theory to reduce another, yet for the models of the reduced theory to have emergent values of properties with respect to the models of the reducing theory. This can arise precisely because there may be models of the reduced theory, arbitrarily similar to those of the reducing theory, that nevertheless have values for a property that are not even similar to those of the reducing theory, as is the case with strong emergence.[31] This is not possible

31 One can induce a unique smallest similarity structure on a space of models from similarity structures on the spaces of relevant property values, analogously with the initial with non-reductive emergence, however; indeed, non-reductive emergence is a sufficient condition for the failure of total reduction. These observations are collected and strengthened slightly in the following proposition.

Proposition 13. *Let $(X, \mathcal{S})$ be a similarity space with closeness operator $\mathrm{cl}_{\mathcal{S}}$, and suppose that T reduces T' completely, with $T, T' \subseteq X$.*
1. *If $T' \subseteq T$, then $\nu[T']$ is not weakly emergent for any property valuation $\nu : X \nrightarrow V$.*
2. *If $T' \nsubseteq T$, then $\nu[T']$ is strongly emergent for some property valuation $\nu : X \nrightarrow V$ but non-reductively emergent for none.*

Proof. 1. By hypothesis that $T' \subseteq T$, $\nu[T'] \subseteq \nu[T]$ for any property valuation $\nu : X \nrightarrow V$.
2. Let $\chi_T : X \to \{0, 1\}$ be the characteristic function for T and $\mathcal{V}$ be a similarity structure on $\{0, 1\}$ consisting only of the relation $\{(0, 0), (1, 1)\}$; χ_T is a valuation of a property on X. By hypothesis, $\chi_T[T'] = \{0, 1\}$ while $\mathrm{cl}_{\mathcal{V}}(\chi_T[T]) = \{1\}$. Hence $\chi_T[T']$ is partially strongly emergent.
   To show that $\nu[T']$ is non-reductively emergent for no property valuation ν, note that by the assumption of complete reduction, $T' \subseteq \mathrm{cl}_{\mathcal{S}}[T]$, hence $\nu[T'] \subseteq \nu[\mathrm{cl}_{\mathcal{S}}(T)]$.

□

The compatibility of some types of emergence with reduction may sound like a radical conclusion if one assumes that emergence is close to being the negation of reduction, but it is in fact entirely natural here, where they are not: once one compares definitions 16 and 18, one sees that explaining the success of one theory from the perspective of another simply allows for former to have novel features.

4.3 An Example: Absolute Simultaneity

To understand these different grades of emergence and their properties, it is helpful to analyze some examples. In particular, I will take the case of the Newtonian and relativistic theories of gravitation discussed in section 3.3, and the

(weak) topology induced on a space from a set of maps that take the space as its domain (Willard, 1970, §8). In such cases, one can show that any strongly emergent properties in the models of a theory to which any theory reduces cannot be ones which make a difference in the similarity structure on the models. For example, if the similarity structure the encodes similarity of certain sorts of empirical descriptions, then no strongly emergent property can make a difference for those sorts of empirical descriptions.

property of having *a concept of objective simultaneity.* As is well known, one of the hallmark features of relativistic spacetimes is that simultaneity is in general relative to an observer, in the sense that different observers may determine different classes of events to be simultaneous with one another.[32] By contrast, a spacetime theory has a concept of objective simultaneity just when its assignments of simultaneity to different events depends only on the spacetime structure itself, rather than on, say, particular worldlines. Non-relativistic spacetimes have a concept of objective simultaneity determined directly from their temporal metric. So, in comparing the models of the two theories, we may assess whether objective simultaneity is emergent, and in what sense. Let T_R be the models of general relativity and T_N be the models of Newtonian gravitation, both of which are strict disjoint subsets of the models of the frame theory T_F described in the previous section. Further, let $\nu_s : T_F \to \{\top, \bot\}$ be the property valuation whose value is $\top$ when evaluated on a model with a concept of objective simultaneity, and is $\bot$ otherwise. Clearly, then, objective simultaneity is completely weakly emergent in the models of Newtonian gravitation with respect to the models of general relativity because $\nu_s[T_N] = \{\top\}$ and $\nu_s[T_R] = \{\bot\}$ are disjoint. Objective simultaneity is simply a novel property in Newtonian gravitation with respect to relativity theory.

Is objective simultaneity also strongly, non-reductively, or radically emergent? Answering these questions requires justifying a choice of similarity structure on the property value space $\{\top, \bot\}$, the space of models T_F, or both, respectively. To begin with the question of strong emergence, a quite natural similarity structure $\mathcal{V}$ for the property value space would consist just in the relation $\sim_T = \{(\top, \top), (\bot, \bot)\}$, which asserts that only (not) having a propositional property is similar to (not) having that property, i.e., “true” is similar to “true,” and “false” to “false.” (This singleton set of similarity relations is topological, and the canonical topology it generates is the discrete topology.) It follows that $\mathrm{cl}_{\mathcal{V}}(\{\bot\}) = \{\bot\}$, hence objective simultaneity is completely strongly emergent in the models of Newtonian gravitation with respect to the models of general relativity.

In section 4.3, two sorts of topological families of similarity relations were introduced for T_F, one whose induced topology is the compact-open topology, and the other whose induced topology is the open topology. As shown there in proposition 11, if the former is adopted for the models of the frame theory,

32 In Minkowski spacetime—the domain of special relativity—all observers always agree that timelike-related events are not simultaneous with one another, but may determine different sets of spacelike-related events to be simultaneous. In an arbitrary relativistic spacetime, this holds for events and observers sufficiently local to one another.

T_F, then in fact $T_N \cap \mathrm{cl}_{\mathcal{S}}(T_R) \neq \emptyset$, i.e., the models of general relativity partially reduce the models of Newtonian gravitation. Thus, since every model of Newtonian gravitation has a concept of objective simultaneity, $\nu_s[\mathrm{cl}_{\mathcal{S}}(T_R)] = \{\top, \bot\}$, hence $\nu_s[T_N] \subseteq \nu_s[\mathrm{cl}_{\mathcal{S}}(T_R)]$, meaning objective simultaneity is not non-reductively emergent, as one would expect from proposition 13. Similarity, since $\nu_s[\mathrm{cl}_{\mathcal{S}}(T_R)] \subseteq \mathrm{cl}_{\mathcal{V}}(\nu_s[\mathrm{cl}_{\mathcal{S}}(T_R)])$, it would also not be radically emergent. However, if the open topology is adopted for the models of general relativity, then $T_N \cap \mathrm{cl}_{\mathcal{S}}(T_R) = \emptyset$, i.e., the models of general relativity fail to reduce the models of Newtonian gravitation. Since $\nu_s[T_N \cap \mathrm{cl}_{\mathcal{S}}(T_R)] = \nu_s[T_N] \cap \nu_s[\mathrm{cl}_{\mathcal{S}}(T_R)]$, it follows that objective simultaneity would be completely non-reductively emergent. In fact, it would be radically emergent, as $\nu_s[\mathrm{cl}_{\mathcal{S}}(T_R)] = \{\bot\}$ would imply that $\mathrm{cl}_{\mathcal{V}}(\nu_s[\mathrm{cl}_{\mathcal{S}}(T_R)]) = \{\bot\}$.

Which of these is the "right" verdict? Is objective simultaneity radically emergent, or merely strongly emergent in the Newtonian theory? I have argued that, to this question, there is in general no context-free answer, one which depends only on the formal structure of the two theories. For instance, if the goal is a description of the properties of Newtonian spacetimes that are in various ways novel compared with those of relativistic spacetimes in light of the former's success in actual empirical descriptions, then, as I argued in section 3.3, there is good reason to believe that the compact-open topology is a better choice than the open topology. But if the goal is such a description in light of a "global view" of the features of spacetime, regardless of whether they are reflected in what is observable in the models, then the open topology may be a better.

4.4 Comparisons with Other Accounts of Emergent Properties

Various other authors have discussed senses of emergence that share features of the present account based on similarity structure. Here I discuss two: that of Primas (1998), Bishop and Atmanspacher (2006), and Atmanspacher and beim Graben (2009); and that of Butterfield (2011a,b, 2014) and Bouatta and Butterfield (2012, 2015). The former call their characterization *contextual emergence*, which "utilizes lower level features as necessary (but not sufficient) conditions for the description of higher-level features" (Atmanspacher and beim Graben, 2009). To exhibit cases of contextual emergence, these authors consider a theory in which systems of interest occupy a single state in a state space of a lower-level theory, which is equipped with a "fundamental topology" determining which trajectories or curves in the state space are continuous (and thus are candidates for representing change over time). One then considers which properties and differences between states make a difference to the context of investigation, coarse-graining or partitioning the state space into equivalence classes of states that, as far as this context is concerned, are indistinguishable: their elements have amongst themselves only differences that do not make a difference (e.g., are not measurable by certain context-dependent apparatus). The equivalence classes of the partition then form the elements of a new state space for the higher-level theory, with the associated quotient topology. The elements of this new higher-level state space may have properties that are not definable in the lower-level state space, even though those properties supervene on lower-level properties.[33] But in any case, this partitioning must be compatible with the fundamental topology in the sense that its quotient should commute with the dynamics at each level.[34]

The approach of contextual emergence shares with the present similarity structure-based approach a commitment to the importance and necessity of context for describing the relationships between different levels or theories, respectively. It also shares the use of topological structure in its analysis, but in somewhat different ways. Contextual emergence assumes and does not attempt to justify its "fundamental" topology on the lower-level state space, which is used only in ensuring that the coarse-graining commutes with the fundamental dynamics. By contrast, the present approach assumes neither of these: there need not be any "fundamental" similarity structure or requirements of commuting with dynamics. Indeed, the present approach does not assume that the models of the two theories to be related are related by any coarse-graining operation at all, or that the models of these theories must be continuous trajectories through state space. Rather, it builds (pre-)topological structure from families of contextually justified similarity relations, which also play an important role in defining three of the four principle types of emergence. However, while a more thorough comparison of the two approaches is still needed, it seems that that of contextual emergence can be subsumed by the that of similarity structure

33 These authors have in mind especially applications to the relationship between microscopic, complex descriptions of multipartite systems and macroscopic, more phenomenological descriptions of these systems as a whole, as is the relationship between mechanics and thermodynamics via statistical mechanics. Temperature is intended to be a contextually emergent property, for example. In these sorts of cases, the introduction of a context is accompanied by a probability distribution on the state space, so that emergence can arise not just from indistinguishable states, but also ignorance about what state is actual.
34 This sometimes requires restricting the lower-level theory's state space, which these authors call implementing a "stability" criterion, citing the KMS condition of quantum statistical mechanics as an example.

considered here. The models related by coarse-graining a state space may still be gathered together in a joint space and equipped with a similarity structure so that a coarse-grained state may be arbitrarily similar to many different fine-grained states, perhaps giving rise in the former to weakly or strongly emergent properties.

These notions of emergence are also related to those presented by Butterfield (2011a,b, 2014) and Bouatta and Butterfield (2012, 2015), who describe emergence as "behaviour that is novel and robust relative to some comparison class" (Butterfield, 2011b, p. 1066). Like them, my formalization of different types of emergence is precisely a sort of relevant comparative novelty. Unlike them, I consider the robustness of a property to be better understood an epistemological criteria about inferences using the models, not about what the models represent—see section 5. While I am sympathetic to demanding that an *interesting* novel (i.e., emergent) property be robust *sensu* Butterfield and Bouatta, I find it helpful to distinguish the representational issues from the inferential ones. (Indeed, one finds in their later expositions that robustness enters only nominally (Bouatta and Butterfield, 2015) or not at all (Butterfield, 2014).)

Now, despite describing emergence as requiring a sort of comparative novelty, Butterfield and Bouatta never give a formal definition of "novelty". The closest they seem to come to doing so is the following: "Here 'novel' means something like: 'not definable from the comparison class', and maybe 'showing features (maybe striking ones) absent from the comparison class' " (Butterfield, 2011a, p. 921). My own formalization of novelty is in accordance with this: a value of a property of a model is weakly emergent with respect to some other models when it does not obtain in the latter, and weak emergence is necessary for the other principle types of emergence I defined. Butterfield and Bouatta are also concerned to argue that such emergence is compatible with reduction, as I have done as well: one can have novelty properties in the models of a theory while still being able to explain that theory's success.

Another relevant point of comparison, and one about which they do say more, is how they see emergence interfacing with *limiting procedures*. Butterfield describes a "system"—a model in the sense I have been using the term—with an emergent property as follows:

> Often the system is a limit of a sequence of systems, typically as some parameter (in the theory of the systems) goes to infinity (or some other crucial value, often zero); and its properties and behaviour are novel and robust compared to those of systems described with a finite (respectively: non-zero) parameter. (Butterfield, 2011a, p. 921fn)

The “parameter,” usually a real-valued number, is here just a way of labeling the ordered elements of the sequence—though it is usually one of interpretive significance, this is not essential for making limits well-defined. In any case, this is a special case of the notion of weak or strong emergence defined here, for proposition 2 shows that the limit of a sequence of models of a theory, if it exists, is certainly arbitrarily similar to those models, yet there may be other models arbitrarily similar to those of the theory that are not the limits of any sequences.

A further way in which the present work builds on that of Butterfield and Bouatta—and shares with contextual emergence—is the analysis of the dependence of emergence on the selection and justification of the relevant family of similarity relations. Now, Butterfield (2011b, p. 1075) does mention that topological structure on properties themselves (what he calls “quantities”) is often needed to define limits, in contrast to the values of the properties, which, assuming that they are real numbers, have a canonical, natural topology. However, he does not describe the motivation, choice, or interpretation for such a topology. By contrast, I do not assume anything about the range of property values, e.g., that they be real numbers. Further, applying the analysis of section 2 shows how (pre)topological structure on the models of a theory or the values of a property arises not as a mere technical device on a space of property values, but from a motivated choice of similarity structure.

One final point of comparison is worth mentioning. Butterfield (2011b, p. 1069) discusses two senses of emergence for a property, strong and weak, the former applying to a novel property of a limit model and the latter applying “*before* we get to the limit. That is: in each example [of emergence], one can understand ‘novel and robust behaviour’ weakly enough that it *does* occur for finite N”—that is, for elements of the sequence. Do these types of emergence fit into the fourfold classification I have developed here? Yes, essentially. Butterfield’s strong emergence is a special case of weak or strong emergence as I have defined it. His weak emergence is also special case of weak or strong emergence, one in which the similarity structure encodes only finitely precise empirical measurements. For Butterfield, a property of a limit model is weakly emergent *in a particular modeling context* when that model sufficiently satisfies that context’s modeling criteria. Further descriptive precision does not make a difference to that context, even though it would in other modeling contexts. For example, one might fruitfully model the solar system as the interaction of various massive point particles under Newtonian gravitation in order, say, to send a space probe to Jupiter. In this context, a concept of absolute simultaneity could be said to be weakly emergent in Butterfield’s sense because it is known that in fact some general relativistic model would be more descriptively accurate in another context (in a way that would make a difference in that context), even though the Newtonian model suffices in this one. Thus weak emergence (in his sense) applies to a property of a model relative to two contexts of investigation: one in which the property counts as weakly or strongly emergent (in my sense) and one—typically one for which strictly more differences make a difference—in which it does not.

5 Modeling and Epistemology

5.1 The Principle of Stability

Scientists routinely build mathematical or computational models of a domain of phenomena, then infer or explain properties about the domain from properties of the models. The centrality of models to scientific practice indeed supports their status as the basic unit of analysis in the present work. But such models are almost always idealized (McMullin, 1985), misrepresenting (or omitting) features of their domains in ways that are not always apparent. However, when an idealization is relaxed so that the resulting model is more fully descriptively accurate, the properties inferred and explanations gleaned from the model may change. Yet much of the time this de-idealization is impractical—for otherwise the idealizations typically would not have been made in the first place! How, then, can inferences from idealized models be justified?

This is a problem considered not just in philosophy of science and epistemology, but within the sciences as well. For instance, the standard Friedmann–Lemaître–Robertson–Walker (FLRW) cosmological model contains a past singularity.[35] It is from this feature that cosmologists infer the finite age of the universe, the existence of a Big Bang. But this model also idealizes the content of the universe as being homogeneous and isotropic: at every instant of cosmological time, the distribution of matter is the same everywhere and in every direction. Thankfully, our experience hardly confirms such monotony. So how is this inference to the existence of a Big Bang therefore warranted?

It is difficult to provide conditions of sufficiency for warrant or justification. But the present framework of similarity structure affords the means to formulate a necessary condition. Namely, one can require that the inference would still have

35 For more on these models, see, e.g., Hawking and Ellis (1973, Ch. 5.3) or Malament (2012, Ch. 2.11).

been made if the idealization were relaxed or forced to hold only approximately—that is:

Principle of Stability (Informal) If an inference from a property of a model of a theory to a property of the world is to be justified, then the property of the model must obtain on all sufficiently similar models.[36]

The idea is that the justification should not depend on the arbitrarily precise selection of a single model. Once the models of a theory have been equipped with similarity structure, this principle can be interpreted as demanding that justified inference of a property of the world from a property of a model requires that the property of the model holds also of all the models in some domain of a similarity relation at that model.

Definition 19. Let $(X, \mathcal{S})$ be a similarity space and $\nu : X \twoheadrightarrow V$ a property valuation. ν is *stable at* $x \in X$ when there is some $D(x) \in \mathcal{B}_x$ such that $\nu[D(x)] = \{\nu(x)\}$. Otherwise, it is *unstable at* x

Principle of Stability (Formal) For any property valuation ν on a similarity space, if an inference from $\nu(x) = v$ to the corresponding property of the world is justified, then ν is stable at x.

If the family of models is topological, then this is equivalent to the existence of a topological neighborhood of the model, all of whose elements have the property. This is just the concept of the stability of a property at a point of a topological space. So, when similarity is encoded in topological structure, the principle of stability takes the (topological) stability of a property of a model to be a necessary condition for one to infer the corresponding property about the world.

In light of this connection with similarity and topology, Fletcher (2016) has suggests that physicists' talk of the "physical significance" of a property of a model just refers to the warrant to infer that property about the physical phenomena the models describes. Thus when Hawking (1971, p. 395) asserts that "the only properties of space-time that are physically significant are those that are stable in some appropriate topology", he means to assert a *necessary* condition for inference from a (space-time) model to properties of the world to be justified—a version of the principle of stability.[37] An articulation of this prin-

36 This is analogous to what Jones (2006, §5.3) calls "Earman's Principle," which applies specifically to the stability inferences of effects from idealized systems under de-idealization—see Earman (2004, p. 191).

37 See also Geroch (1971, p. 70) and Hawking and Ellis (1973, p. 197).

ciple can already be found as early as that of Duhem (1954, Part II, Ch. III), who argued that inferences and explanations from mathematical models of phenomena (“mathematical deductions”) must satisfy an additional criterion if they are not to be “otiose” and “condemned to eternal sterility.” In order to be “useful” in real physical applications, a given inference from a mathematical model must also be demonstrated to be approximately correct when the (idealized) assumptions of the model are only approximately true.

In the remainder of this section I will not try to defend the principle of stability or describe any more of its history or classification as an epistemological principle (even though those are worthy pursuits—see section 7.4). Rather, I will confine attention to exploring a couple of the principle’s implications: the pervasiveness of unstable properties and the dependence of stability on the relevant family of similarity relations.

How common should we expect unstable properties to be? There’s a certain sense in which they are pervasive, as captured by the following proposition.

Proposition 14. *Consider a similarity space* $(X, \mathcal{S})$*. For any* $x \in X$*, if there is some similarity relation in* $\mathcal{S}$ *with a domain at* x *that non-empty and no similarity relation in* $\mathcal{S}$ *with a domain at* x *that is the singleton* $\{x\}$*, then the property valuation* $\chi_x : X \to \{\top, \bot\}$*, defined by*

$$\chi_x(y) = \begin{cases} \top & \text{if } y = x, \\ \bot & \text{if } y \neq x, \end{cases} \tag{1}$$

is unstable at x*.*

Proof. If the conditions of the hypothesis are satisfied for $x \in X$, then every domain at x of a similarity relation on X contains both x and some $y \neq x$. Since $\chi_x(y) = \bot$ when $y \neq x$ while $\chi_x(y) = \top$ when $y = x$, there is no non-empty domain at x on which the values of χ_x are only $\top$. □

χ_x is a property valuation for “being x.” If one uses a model x to represent some phenomenon, then the property represented by χ_x is just that x truly does represent the phenomenon. The above proposition shows that, if the context of investigation determines that the only relevant ways in which models are similar always makes one model similar to a distinct one, then “being represented by a single particular model” is an unstable property. The principle of stability then entails that one is never justified in inferring that just one particular model represents the phenomena. Thus, quite pervasively, there are always some properties of a model that one is not justified in inferring of the world, even though one may be justified in using that model.[38]

In section 4.4, I described how Butterfield (2011a, p. 921) demanded that emergent properties must be "robust," in the sense of being "the same for various choices of, or assumptions about, the comparison class." If this sense of robustness is understood as a sort of *stability* of the property under perturbations of the model used, then it is just an instance of stability in the sense of this section. If this were right, then, as I suggested in section 4.4, this notion of robustness would not be a necessary feature of emergent properties per se, but would rather be, in the light of the principle of stability, a necessary condition for one to infer such an emergent property of some phenomenon from a model. Such robustness would therefore be at most a requirement for an emergent property to be interesting or useful, not to exist.

5.2 An Example: Chronology and Causality in General Relativity

Whether a property is robust or stable in this sense, though, depends crucially on which other models are a part of the theory and on the family of similarity relations placed on those models. For example, consider again relativity theory. Some of its models exhibit closed timelike or causal curves, possible world-histories of particles that loop back on themselves. These are widely considered to be a sort of time travel, so spacetimes with closed timelike or causal curves are said to violate chronology or causality, respectively, while those without are respectively called chronological or causal. So, one can ask whether a spacetime has any of these features stably, so that, according to the principle of stability, they might be inferred of a world modeled by them. One central result about this is the following (Hawking and Ellis, 1973, p. 198, Proposition 6.4.9):

Proposition 15. *A relativistic spacetime is stably causal with respect to the open topology if and only if it admits of a global time function.*

38 Clearly, therefore, there is also a connection with reduction and emergence. For example, suppose a similarity space consisting just of the models of two theories has symmetric similarity relations. If one theory completely reduces the other while there being strongly emergent properties in the latter (with respect to the former), then those emergent properties are unstable in the models in which they obtain. Further, sharper propositions along these lines might be proven.

A global time function is a continuous function of spacetime events that is strictly increasing along each future-directed causal curve. There are many spacetimes that support a global time function: the Minkowski and FLRW spacetimes are just two well-known examples.[39] Thus, with respect to a similarity structure that induces the open topology—one which considers global features of spacetime to be relevant—all spacetimes sufficiently similar to these are also causal, hence they pass the necessary condition for inferring causality of a world or state of affairs they model.

However, if one selects as the relevant similarity structure one whose induced topology on spacetimes is the compact-open topology, the situation is totally different (Fletcher, 2016, p. 376, Corollary 1):

Proposition 16. *No relativistic spacetime is stably causal in the compact-open topology.*

Such a family of similarity relations determines spacetimes to be sufficiently similar based on their similar empirical descriptions of compact regions of spacetime, which are the best candidates for the regions we can actually measure. The above proposition thus shows that, if the relevant way in which spacetimes must be similar is to be based not on global similarity but similarity restricted to what is measurable, then the principle of stability would prevent one from ever inferring of the world that it is free from time travel. The situation is the opposite for spacetimes allowing for time travel (Fletcher, 2016, p. 377, Proposition 5):

Proposition 17. *Every relativistic spacetime which violates chronology does so stably in the compact-open topology.*

The necessary conditions for inferring the existence of time travel, as prescribed by the principle of stability, are always met according to this way of determining spacetime similarity.

Which of these is the "correct" result, if any? As I have argued here and elsewhere (Fletcher, 2014, 2016), the choice of similarity structure on a collection of models (and any topology it may induce thereon) should be based on contextually dependent factors: all of and no more than whatever makes a difference to the inquiry or question at hand should be included. While this methodological contextualism may seem pragmatically agreeable, what is perhaps surprising about it is that, through the principle of stability, it puts contextual constraints

39 Note that a spacetime admitting of a global time function in no way implies that there is a preferred such function, which would define a concept of objective simultaneity on that spacetime.

on what inferences from models may be warranted or justified, without changing anything about the models themselves.

6 Counterfactuals and Laws in Science

6.1 Ordering Semantics for Counterfactuals from Similarity Structure

Scientists use models not just using classical logical deduction, to predict phenomena or describe the world, but also to reason counterfactually about what would have happened, what would be happening, and what would happen. As philosophers (Lewis, 1981, 1973b, §2.3) and linguists (Kratzer, 1981) have given semantics for counterfactual conditionals in terms of comparative similarity (or possibility) as a sort of preordering relation on worlds,[40] one might wonder whether the present framework of similarity structure can be adapted to these semantics, even though they have been interested in everyday language, while we are here concerned with the models of a scientific theory. The answer, as I show below, is affirmative: any similarity structure determines a certain partial preorder, which is precisely the structure on worlds used in the ordering semantics for the above analyses of counterfactuals. In what follows, I first describe how the determination occurs, then draw out some implications of this analysis for scientific reasoning and laws of nature in section 6.2.

Let $(X, \mathcal{S})$ denote a similarity space.

Definition 20. For any $x, y, z \in X$, *x is at least as similar to z as y is (with respect to $\mathcal{S}$)*, written $x \lesssim_z y$, when for all $\sim \in \mathcal{S}$, if $y \sim z$ then $x \sim z$.

This three-place relation determines the meaning of "greater comparative similarity" through strictly more extensive satisfaction of the similarity relations in $\mathcal{S}$. In other words, the model x is at least as similar to the model z as the model y when every qualitative manner in which y is similar to z is also one in which x is similar to z.

The following proposition follows immediately from the definitions.

Proposition 18. *For each $z \in X$, $\lesssim_z$ satisfies the following conditions:*

40 Strictly speaking, originally Lewis (1973b) assumed a total preorder (satisfying certain other properties), according to which each pair of worlds is comparable under the relation, but seemingly came to relax that assumption later (Lewis, 1981).

1. *(Conditional Reflexivity) For any* $x \in X$*, if there is some* $y \in X$ *such that either* $x \lesssim_z y$ *or* $y \lesssim_z x$*, then* $x \lesssim_z x$*.*
2. *(Transitivity) For any* $w, x, y \in X$*, if* $w \lesssim_z x$ *and* $x \lesssim_z y$ *then* $w \lesssim_z y$*.*
3. *(Conditional Weak Centering) If there are some* $x, y \in X$ *such that* $x \lesssim_z y$*, then for all* $x \in X$*,* $z \lesssim_z x$*.*

The first two properties—conditional reflexivity and transitivity—make $\lesssim_z$ a partial preorder.[41] Conditional weak centering formalizes the idea that any model is at least as similar to itself as any other, if there are any at all to which it is similar. It is a consequence of the fact that similarity relations have been defined as being right quasi-reflexive: if any element is similar (in some manner) to another, then the latter is similar to itself. If $\mathcal{S}$ is neighborly, though, then each $\lesssim_z$ is (unconditionally) weakly centered: $x \in X$, $z \lesssim_z x$.

Such structures can provide interpretations of counterfactual conditionals. Let $A, C : X \to \{\top, \bot\}$ be (Boolean) property valuations of models of X—each either holds or not at a given model. Models at which A holds—those $x \in X$ for which $A(x) = \top$—are called A-models. If we symbolize the counterfactual conditional as "if A were the case, then, C would be the case" as 'A > C', then following Lewis (1981), we could write its truth conditions as follows in terms of concepts derived from similarity structure:

> 'A > C' is true at a model z if and only if, for any A-model $x \in \bigcup_\alpha D_\alpha(z)$, there is some A-model y such that $y \lesssim_z x$ and C holds at any A-model w such that $w \lesssim_z y$.

In other words, 'A > C' is true at a model z just when C holds of all A-models sufficiently similar to z. Now, it is not essential that one use precisely Lewis's 1981 semantics for counterfactuals: any account whose truth conditions can be logically constructed from the structure of a partial preorder on models can be accommodated here.[42] Indeed, my goal here is not to argue for a particular semantics, but to show how any based on partial preorders can be modeled with a family of similarity relations.

One difference with Lewis's (and others') accounts concerns the nature of the objects related by the partial preorder. Traditionally, these have been possible worlds, which are thought to consist in or determine the truth values of all propositions, perhaps requiring deep metaphysical commitments. By contrast,

41 By contrast, a total preorder $\lesssim$ on a set A satisfies (unconditional) reflexivity: $a \lesssim a$ for all $a \in A$.

42 Swanson (2011) has criticized Lewis's semantics and offered a replacement in terms of cutsets definable from the partial preorder. I find his argument plausible, but do not introduce his semantics only to minimize the introduction of new technical concepts.

the models of the present similarity structure-based approach merely represent the descriptions of phenomena by some scientific theory or other. Unlike worlds, these are generally intended just to represent phenomena in the domain of applicability of the theory, not to be complete descriptions of the universe. In this sense, they are more like states of affairs than possible worlds, but unlike states of affairs, which are typically understood as non-exhaustive collections of propositions or parts (possibly with abstractions) of worlds, models themselves need not be directly a part of the world of phenomena but merely *represent* such a part.

Accordingly, another difference is that one's ontological commitment in using the semantics is to whatever one has committed to by using the models, which covaries with the strength of one's realist attitude towards the models. This is entirely compatible with views from constructive empiricism through selective realism to strong realism. In other words, it is not metaphysical claims about the status of possible worlds (or states of affairs) but the commitments to and practice with scientific models and theories that ground the semantics.

Connected with this, a third differences lies in the fact that counterfactuals considered here can only be evaluated relative to some theory or model about the phenomena they describe. In this sense, the present goal is a semantics for scientific counterfactuals—e.g., "according to theory T, if A were the case, then B would be the case"—rather than for natural language in toto. It is (in part) this more restricted goal that give the present framework its power.

6.2 Similarity Structure and Laws of Nature

While similarity structure yields a familiar technical apparatus to give meaning to counterfactuals, it also allows for novel treatments of two aspects of laws of nature. First, note that the above semantics requires no reference at all to laws of nature as such; they are strictly superfluous, for the models and similarity relations on them suffice to provide counterfactual semantics. Thus, it suggests a novel version of antirealism about such laws (van Fraassen, 1989; Giere, 1999) that does not necessarily extend to counterfactuals—or at least a strong form of deflationism that minimizes the special role of lawhood. Indeed, the usual main complaint about antirealism about laws is that it has trouble grounding nomic claims and inferences, especially the use of counterfactuals in science (Carroll, 2012, §5). However, in the account sketched above in section 6.1, all one needs for this purpose are models and contextually determined similarity relations on those models.

While my account, as a formal semantics grounded in scientific models, is *compatible* with antirealism about laws, it does not *presuppose* such antirealism. For instance, it is also compatible with a *systems* account of laws, in which the laws are best axiomatizations of true deductive systems for the world—where "best" is understood as some balance of simplicity and deductive strength (Lewis, 1973b, Ch. 3.3). Systems accounts can also use models and similarity relations to determine the meaning of scientific counterfactuals while additionally asserting that certain *descriptions* of those models are privileged as natural laws.

There is a useful metaphor that illustrates this idea while also explaining, under the antirealist interpretation, why seemingly substantive talk about laws is so common in science. Imagine the models of a theory as the minuscule parts of an intricately designed statue. Just as it is difficult for a sculptor to order a block of marble of precisely the shape of her finished statue, or even communicate what that shape is to a quarry, it is often difficult to describe the models of theory directly. The sculptor begins by ordering a block of material of the right sort which is expansive enough to contain the form of the statue, and her application of skill to create the intended statue consists in successive blows that remove material from the block. There is no unique sequence of chiselings that the sculptor must enact in order to create the statue; any sequence which yields the intended statue suffices, though some sequences are more elegant, reproducible, and measured than others. Analogously, the theoretician interested in describing or representing a theory with a class of models begins by selecting a sufficiently common or well-understood class of mathematical structures—such as vector spaces or probability spaces—that is also expansive enough to contain the intended models. Her application of skill consists in the selection and description of more or less easily describable restrictions on that structure to produce the models of a theory. For instance, she might require a probabilistic theory of belief over propositions of a first-order language to be regular, i.e., assign positive probability to any sentence which is not logically false. These restrictions might be said to be a system of natural laws, as they provide, within the context of the theory, necessary connections, correlations, restrictions, and determinations among phenomena. But like with the sculptor's statue, what is of ultimate substance for the theoretician is the finished product, the collection of models representing the theory, not the particular successive choices of restrictions on the class of mathematical structures.

Now, just as the various sequences of chiselings a sculptor can make to yield an intended statue can be clear or obscure, simple or complicated, the particular "laws" the theoretician formulates can be so as well. Being clear and simple facilitates communication and eases the costs of applying a theory, e.g., for making scientifically informed counterfactual judgments. These are pragmatically virtuous features, and explain why laws and features thereof feature importantly in much scientific discourse. But this virtue does not necessarily accrue because of any metaphysical significance to the choice of how to describe a collection of models, just as the choice of chiselings away of excess material need not be essential to the significance of the sculptor's statue. Antirealists about natural law may indeed deny such significance at all, while proponents of systems accounts might endow this significance to particular descriptions meeting sufficient standard of simplicity and strength. In either case, though, the pervasive centrality of laws in much of science can be understood as arising from pragmatic concerns. One can decide for oneself how metaphysically significant these cuts and cleavages are.

The second novel feature of the present account is that it provides equally well a semantics for counterlegals—conditional statements that suppose conditions contrary to natural law—as it does for counterfactuals themselves. Because, as just discussed, laws of nature play mostly a pragmatic role in the descriptions of classes of models, the consideration of different theories embodying different such laws requires no real emendation or extension. Models of different theories may well be considered within the same space on which similarity structure is placed, as was already done in the consideration of the account of intertheoretic reduction in section 3. Thus the antecedent of a counterlegal conditional can be evaluated in exactly the same way as a counterfactual conditional, as the property designated by certain (non-actual) laws obtaining merely picks out a collection of models in the space, just as that designated by certain (non-actual) facts does so in the semantics for the conditional. The only restriction is that the truth of these counterlegals, much like with counterfactuals, can only be evaluated relative to some more or less explicit joint collections of models. Hence, the semantics is not for the counterlegals of everyday language, but for scientific counterlegals: "according to the theories $\{T_\alpha\}$, if laws L were the case, then C would be the case".

7 Envoi

As the previous sections detail, many central issues concerning scientific theories can be treated with flexibility and precision using similarity structure and the (pre)topological structure it induces on the models of those theories. In fact, another general advantage of this approach and program is that nothing in the analysis essentially depends on identifying particular classes of models with particular theories. As I mentioned in section 2.2, many philosophers of science have argued that in some domains of science, models, rather than theories, are the important unit of analysis (Bailer-Jones, 2002; Craver, 2002), and the present approach is entirely compatible with this. The role of a theory here is only to demarcate, name, or label a certain class of models, and it is largely immaterial whether these are deemed to constitute or represent theories.

Furthermore, unlike the ambitions of some formal approaches to questions in philosophy of science, the present approach cannot demonstrate anything of much substance of a scientific theory or class of models without input from some (real or imagined) scientific practice. This is because a theory or class of models does not come equipped with a similarity structure intrinsically, but only gains that structure through the specification of an investigative context. Whatever matters in the application of a particular scientific theory or model bears on how those models are relevantly similar. Requiring users of the formalism to pay close attention to the details of what matters in a scientific context, rather than abstracting those details away, is an advantage because it makes the formalism applicable to real science in which context is crucial.

Another advantage is that the formalism does not presuppose any particular position on the scientific realism/antirealism debate. Whether certain models can be similar to one another in some respect does not depend on whether the models are of the structure of reality, merely of phenomena, or some admixture of the two. It does not even depend on this structure being of the same type for all models, allowing for various forms of selective scientific realism. Now, as I discussed in the section 6.2, it *does* suggest a more thin or nominal conception of natural laws, but advocates of more metaphysically substantial conceptions of laws are quite free to ignore this—nothing in the formalism demands it, although more would need to be said about how universals or antireductionist accounts of laws interface with it.

The approach does have a few disadvantages. In my treatment so far, the models have been abstract, mathematical, or otherwise formal. Because similarity relations are simply right quasi-reflexive binary relations, it is possible in principle for them to hold between more informal abstract or even concrete objects, but these tend to have vague or ambiguous properties to which the successful application of the approach may be more difficult. Whether this is possible requires further investigation which I have not yet pursued. Thus, the need for formality may not be an intrinsic disadvantage, one that arises only because of the approach's present state of development. In any case, that the approach is more powerful as the models are made more formally precise provides incentives for sharpening the descriptions and properties of the models in which one is interested and being more explicit about the theory (if any) one is using, an independent virtue in itself.

Beyond these general advantages and possible limitations, work on each of the topics treated above in terms of similarity structure can be extended further. They also suggest many further open research questions varying across its technical, philosophical, and historical aspects. The following subsections will elaborate on some of these, suggesting just how much more there is to come.

7.1 Similarity and Topology

The framework of families of similarity relations presents both technical and conceptual avenues for future research. Technically, sets equipped with similarity structure are a (moderately distant) generalization of uniform spaces—really, bases for semi-quasi-uniform spaces (Pu and Pu, 1974; Császár, 1974)—and should admit of a more systematic development along similar lines. For example, while I have indicated ways in which similarity structure on a set induces (pre)topological structure on that set, more systematic functorial relations between these sorts of structures and others—uniform, metric, etc.—needs to be explored. More development is also needed on the relationships between different similarity structures and the ways in which requiring mappings between two spaces to be continuous with respect to the structure on one of the spaces can induce similarity structure on the other. Such a development would be needed to show how, for instance, similarity on the value spaces of properties of a collection of models induces similarity on the models themselves. And just as with topological spaces, one can also investigate separation, countability, compactness, and connectedness properties, and invariance with respect to the isomorphisms of the models.

Many of these technical questions have corresponding conceptual aspects. For example, I have stressed that all and only the sorts of differences between models that make a difference to the context of investigation, whether explanatory or otherwise, ought to be taken into account when constructing the similarity structure for that context. Further detailed explication of this idea should show how enough information about this context fixes the relevant similarity structure, and how adding or removing potential differences that could make a difference would add or remove similarity relations from the structure, respectively. I suspect that the best way to do this is in terms of the properties of the models, similarity between which induces similarity of models as suggested above.

Finally, it is an open question whether the whole framework will require modification or adaption to the case of probabilistic theories and models. On the one hand, the present framework does not make any explicit assumptions about the models under consideration being *non*-probabilistic, so there is no corresponding explicit reason to expect modification to be needed. On the other hand, probability theory introduces new convergence and similarity concepts—e.g., convergence almost surely, in probability, and in distribution—whose subsumption under the present ones is not entirely clear.[43] More work needs to be done to assess the relationships between these concepts and those introduced here before the present framework can be applied to many important cases involving probabilistic scientific theories, such as statistical and quantum mechanics.

7.2 Intertheoretic Reduction

There is a wide avenue for research pursuing applications of the definition of reduction given in section 3.2 to specific cases of interest, which would involve determining the most relevant similarity structures for the models of a pair of theories of interest, the resulting character of the reductive relationship between them—complete, merely partial, or neither—and their domains of successful application. This last feature is important because it has implications for the assessment of a putative reduction as an explanation. As described in section 3.3, Fletcher (2014) has already made progress on these questions for the case of general relativity and Newtonian gravitation, but even there, the question of whether the reduction is complete or merely partial is not yet answered. Further concrete examples abound, including the relationships between the following: theories of statistical mechanics and thermodynamics; quantum and classical mechanics; field and particle theories, whether quantum or not, relativistic or not, etc.; continuum mechanics, rigid body mechanics, and particle mechanics; classical and molecular genetics; and theories of macro- and microeconomics.

Beyond wider applications, the similarity structure-based account of reduction also suggests avenues for research concerning two other venerable topics in philosophy of science: theories of explanation, and conceptual continuity across scientific theories. I have suggested in section 3 that reduction, as described in the above section, can count as an explanatory relation of one collection of models by another relative to a context of features encoded in the similarity relations

43 See, e.g., Billingsley (1995) for more on the concepts and mathematics of probability theory.

on them. However plausible it may be, I have not offered a systematic theory to undergird this claim. Now, perhaps a systematic theory of such explanations is not needed, but whether this is so, it would still be illuminating to see how the this type of approximative or similarity-based notion connects—or does not connect—with other well-known accounts of explanation.

Regarding conceptual continuity, the exhibition of a reduction relation between the models representing two theories shows *some* continuity between the theories, at least concerning the properties that make a difference for the similarity relations. Conversely, the extent to which reduction *fails* reveals discontinuity. By adjusting the similarity structure used to be larger or smaller, one can test whether the addition or subtraction of specific properties to be tracked by the similarity relations makes a difference for the reduction relation. This suggests that there may be a *maximal* similarity structure that tracks exactly those properties whose corresponding concepts are continuous in the theories—a family that cannot include another similarity relation without entailing a downgrading of the reduction (e.g., from complete to merely partial). Can a more precise definition be given? Given such a definition, do such maximal structures exist? If they do, what do they entail about the conceptual continuity between particular theories of interest?[44]

7.3 Emergent Properties

Just as with the case of reduction, there is a wide avenue for research pursuing applications of the different types of emergence defined in section 4.2 to specific cases of interest, such as those arising from quantum entanglement or holistic properties of complex composite systems. This would as well involve determining the most relevant similarity structure on the models of a pair of theories of interest and the character of the emergence that some properties of the one have—weak, strong, non-reductive, or radical, and among these either partial or complete. As proposition 13 suggests, it will be easy to find cases of emergent properties of some sort or other, but I hypothesize that many cases will not

44 If this approach can be made more precise, it would also raise the question of to what extent it is representation-dependent, i.e., dependent on the models that one uses to represent a theory. If there is such dependence, then the approach would more properly reveal information about continuity between (classes of) models, rather than theories themselves.

be interesting![45] What, though, makes cases of emergence like supporting a concept of objective simultaneity (for Newtonian gravitation with respect to general relativity) interesting, but not having a mass of 1 kg? I have suggested that it is the emergent property's implications beyond being merely novel that make it so, but this suggestion certainly needs elaboration, if not emendation. Perhaps the properties emergent in the models of one theory because they are not even definable in the models of the other deserve special attention here.

Even more so than reduction, there are many different theories of emergence that philosophers have articulated. I have described some of the relationships between the notion described here and those of robust novelty (Butterfield, 2011a,b, 2014; Bouatta and Butterfield, 2012, 2015) and contextual emergence (Primas, 1998; Bishop and Atmanspacher, 2006; Atmanspacher and beim Graben, 2009), but the connections with others that are sensitive to the scientific usage of the term (e.g., Gillett (2016) and Humphreys (2016b)) deserve to be explored.

7.4 Modeling and Epistemology

The principle of stability, as formulated using the apparatus of similarity relations, raises many interesting questions about its historical status, its broader role in science, what relationship it bears with analogous principles concerning scientific modeling and epistemology, the epistemological status of the principle itself, and its connection with several traditional topics in epistemology. In the first place, while I have located a version of the principle in the writing of Duhem, but was Duhem the first to consider or formulate it? What, all things considered, was Duhem's considered view on the subject? Into which sciences has the principle penetrated, and to what extent?

There are similarly named principles and concepts in the sciences, comparison with which would also be useful. For example, in the theory of dynamical systems, there is a concern with the (structural) stability of a system, its qualitative invariance under certain types of perturbations. Stability in this sense seems just to be a particular example of stability in the sense described by the principle, with the qualitative features in question being the ones tracked by the relevant similarity relations. In that field there has been a stability dogma

45 Here I appeal to a version of what Humphreys (2016a) calls the *rarity heuristic*: a satisfactory account of emergence can't make it too common. More specifically, while on the present account emergence will likely be common, *interesting* emergent properties ought to be more uncommon.

about the relevance and interpretation of stable and unstable features, which has seemingly played a role similar to that of intended for the principle of stability (Schmidt, 2011). Further work is needed to assess the relations between the concepts and history of stability in the principle and in dynamical systems theory, and with other similar concepts used in science, such as robustness.

Further comparisons between the principle and similar principles in traditional epistemology are also needed. For instance, many philosophers have proposed that a condition of *safety* be necessary for someone's belief to count as knowledge (Ichikawa and Steup, 2017, §5.3): in all nearby worlds in which that person holds the belief, the belief is true. If one replaces "nearby worlds" with "sufficiently similar models," "belief" with "inference," and "truth" with "a property obtaining," then this condition seems like the justified-knowledge analog of the principle of stability, which applies to justified inference. As there is much literature discussing versions of this principle, perhaps connections with it could be forged to provide mutual insight.

Modal conditions on knowledge are also often discussed in the context of the question of epistemic closure: whether, given one's store of knowledge, one is justified in adding to that store any claim it logically entails. It has been suggested, for example, than an analysis of knowledge which takes safety as a necessary condition for justification can yield a failure of epistemic closure (Luper, 2016, §2.1). By comparison, I described in the section on modeling and epistemology how, generically, the principle of stability entails that there will be properties of models justifiably used for scientific inference that are not justified in being inferred of their target—for Duhem (1954), for example, this included the (in)stability of the solar system! This failure of transmission of justification, which seems like a close analog of a failure of epistemic closure, deserves further study and scrutiny.

Another worthy investigation concerns whether the principle of stability has any connection with Fitch's paradox of knowability, a logical result which seems to state that if every truth is possibly knowable, then all truths are in fact known (Brogaard and Salerno, 2013). Since the consequent is universally agreed to be false, this entails that there must be some unknowable truth. The sorts of properties of models to which the principle of stability prevents inference could well be true of the state of affairs they represent, so that, if justification is necessary for knowledge, a proposition expressing that such a property obtains might be a concrete example of an unknowable truth.[46]

46 Such an instantiation may require some realist commitments to the theory being used, but this too should be taken up in future investigation.

Finally, the epistemological status of the principle deserves further investigation: should the principle be generally accepted? Restrictedly? What are the grounds for doing so? In fact, it is difficult to find instances in which an author questions, furnishes an argument for, or provides a reference or citation for the principle. One possible explanation for this is that the principle is *constitutive* of a certain type of epistemic activity involving *imprecision* between model and phenomena (or world), *continuity* between various models and various state of phenomena, and the need for *justification* of inference from model to phenomena. Would that these of these conditions constitute modeling activity provide a contingent transcendental argument for this epistemic principle (cf. Chang (2008))?

7.5 Counterfactuals and Laws in Science

As with previous concepts, more applications within various sciences of the account of counterfactual semantics need to be explored. This is of course also a test of the particular semantical proposals for using the partial preorders generated by similarity structures. Attempting its wider application, e.g., in the analysis of time symmetry, measurement, and non-locality in quantum theory, may also show that these formal tools can be of use in analyzing problems involving modal concepts specific to the particular sciences. The approach considered here is plausibly capable of this flexibility because the similarity structure one selects for a class of models depends on contextually relevant features of those models—they arise bottom-up from the context of investigation, rather than being imposed from the top down.

A related direction of research also integral to understanding science is the analysis of causation. Lewis (1973a) proposed such an analysis in terms of counterfactuals, but his focus on the overall similarity of possible worlds too often lets features irrelevant to a particular scientific context make a difference in the truth value of counterfactuals (Menzies, 2014, §3). Meanwhile, the formalism which has had the most scientific success, based on structural equations modeling, such as that of Pearl (2009) and elaborated by Woodward (2003), uses the concept of primitive interventions which are given meaning through scientific context. However, I conjecture that the oft-claimed irreducibility of these latter formalisms to Lewis's is based on the particular features of Lewis's account of similarity, not on the essential ideas of the similarity approach itself. Indeed, I believe that a counterfactual foundation for interventions can be provided which is also powerfully responsive to the questions posed within scientific theorizing and modeling.

Finally, the new, minimal view of laws of nature suggested by the approach to counterfactuals also deserves further exploration and elaboration, both in its affinities for antirealism about laws and for the systems approach. I have suggested that it solves many of the problems associated with the antirealist position, for example; to what extent is this the case? And to what extent does it do so for the systems approaches? What is its relationship—e.g., compatibility or incompatibility?—with other views of law of nature? These questions are currently open; their answers are likely to provide and develop even more seeds of fruitful future research.

Acknowledgment: Thanks to Patricia Palacios, Elanor Taylor, and Marek Kuś for comments on earlier drafts, and to Neil Dewar's seminar on "Reduction and Emergence" at the MCMP in Summer, 2017 for helping me catch embarrassing mistakes—naturally, all future embarrassment is my responsibility. Parts of the research leading to this work have been supported by a National Science Foundation Graduate Research Fellowship and a Marie Curie International Incoming Fellowship (PIIF-GA-2013-628533) within the Seventh European Community Framework Programme.

References

Atmanspacher, H. and beim Graben, P. (2009). Contextual emergence. *Scholarpedia*, 4(3):7997. Revision #122848.

Bailer-Jones, D. M. (2002). Models, metaphors and analogies. In Machamer, P. and Silberstein, M., editors, *The Blackwell Guide to the Philosophy of Science*, pages 108–127. Blackwell Publishers Ltd.

Balzer, W., Moulines, C. U., and Sneed, J. D. (1987). *An Architectonic for Science*, volume 186 of *Synthese Library*. D Reidel, Dordrecht.

Bartha, P. (2016). Analogy and analogical reasoning. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2016 edition. https://plato.stanford.edu/archives/win2016/entries/reasoning-analogy/.

Batterman, R. (2016). Intertheory relations in physics. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2016 edition. https://plato.stanford.edu/archives/fall2016/entries/physics-interrelate/.

Batterman, R. W. (2002). *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence*. Oxford University Press, New York.

Billingsley, P. (1995). *Probability and Measure*. Wiley, New York, 3rd edition.

Bishop, R. C. and Atmanspacher, H. (2006). Contextual emergence in the description of properties. *Foundations of Physics*, 36(12):1753–1777.

Bouatta, N. and Butterfield, J. (2012). Emergence and reduction combined in phase transitions. In Kouneiher, J., Barbachoux, C., Masson, T., and Vey, D., editors, *Frontiers* *of Fundamental Physics: The Eleventh International Symposium*, volume 1446 of *AIP Conference Proceedings*, pages 383–403.

Bouatta, N. and Butterfield, J. (2015). On emergence in gauge theories at the 't Hooft limit. *European Journal for Philosophy of Science*, 5(1):55–87.

Brogaard, B. and Salerno, J. (2013). Fitch's paradox of knowability. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2013 edition. https://plato.stanford.edu/archives/win2013/entries/fitch-paradox/.

Butterfield, J. (2011a). Emergence, reduction and supervenience: A varied landscape. *Foundations of Physics*, 41(6):920–959.

Butterfield, J. (2011b). Less is different: Emergence and reduction reconciled. *Foundations of Physics*, 41(6):1065–1135.

Butterfield, J. (2014). Reduction, emergence, and renormalization. *Journal of Philosophy*, CXI(1):5–49.

Carnap, R. (1967). *The Logical Structure of the World: Pseudoproblems in Philosophy*. University of California Press, Berkeley, 2nd edition. Rolf A. George, translator.

Carroll, J. W. (2012). Laws of nature. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2012 edition. http://plato.stanford.edu/archives/spr2012/entries/laws-of-nature/.

Čech, E. (1966). *Topological Spaces*. Interscience, London, revised edition. Edited by Zdeněk Frolík and Miroslav Katětov.

Chang, H. (2008). Contingent transcendental arguments for metaphysical principles. *Royal Institute of Philosophy Supplement*, 63:113–133.

Craver, C. F. (2002). Structures of scientific theories. In Machamer, P. and Silberstein, M., editors, *The Blackwell Guide to the Philosophy of Science*, pages 55–79. Blackwell Publishers Ltd.

Császár, Á. (1974). Invariant structures and transformation groups. In Császár, Á., editor, *Topics in Topology*, volume 8 of *Colloquia Mathematica Societas János Bolyai*, pages 165–172. North-Holland.

Decock, L. and Douven, I. (2011). Similarity after Goodman. *Review of Philosophy and Psychology*, 2(1):61–75.

Dizadji-Bahmani, F., Frigg, R., and Hartmann, S. (2010). Who's afraid of Nagelian reduction? *Erkenntnis*, 73(3):393–412.

Duhem, P. (1914/1954). *The Aim and Structure of Physical Theory*. Princeton University Press, Princeton, 2nd edition. Philip P. Wiener, translator.

Earman, J. (2004). Curie's Principle and spontaneous symmetry breaking. *International Studies in the Philosophy of Science*, 18(2–3):173–198.

Ehlers, J. (1981). Über den Newtonschen grenzwert der Einsteinschen gravitationstheorie. In Nitsch, J., Pfarr, J., and Stachow, E.-W., editors, *Grundlagenprobleme der modernen Physik*, pages 65–84. Bibliographisches Institut, Mannheim.

Ehlers, J. (1986). On limit relations between, and approximative explanations of, physical theories. In Marcus, R. B., Dorn, G. J., and Weingartner, P., editors, *Logic, Methodology and Philosophy of Science VII*, volume 8 of *Proceedings of the International Congress of Logic Methodology and Philosophy of Science*, pages 387–403. Elsevier Science Publishers B.V.

Ehlers, J. (1991). The Newtonian limit of general relativity. In Ferrarese, G., editor, *Classical Mechanics and Relativity: Relationship and Consistency*, pages 95–106. Bibliopolis, Naples.

Ehlers, J. (1998). The Newtonian limit of general relativity. In Richter, A. K., editor, *Understanding Physics*, pages 1–13. Copernicus Gesellschaft e.V., Katlenburg-Lindau, Germany.

Fletcher, S. C. (2012). What counts as a Newtonian system? The view from Norton's dome. *European Journal for Philosophy of Science*, 2(3):275–297.

Fletcher, S. C. (2014). *Similarity and Spacetime: Studies in Intertheoretic Reduction and Physical Significance*. PhD thesis, University of California, Irvine.

Fletcher, S. C. (2016). Similarity, topology, and physical significance in relativity theory. *British Journal for the Philosophy of Science*, 67(2):365–389.

Frigg, R. and Hartmann, S. (2012). Models in science. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2012 edition. http://plato.stanford.edu/archives/fall2012/entries/models-science/.

Frigg, R. and Nguyen, J. (2016). Scientific representation. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2016 edition. https://plato.stanford.edu/archives/win2016/entries/scientific-representation/.

Geroch, R. (1971). General relativity in the large. *General Relativity and Gravitation*, 2(1):61–74.

Giere, R. N. (1988). *Explaining Science: A Cognitive Approach*. University of Chicago Press, Chicago.

Giere, R. N. (1999). *Science without Laws*. University of Chicago Press, Chicago.

Gillett, C. (2016). *Reduction and Emergence in Science and Philosophy*. Cambridge University Press, Cambridge.

Goodman, N. (1972). *Problems and Projects*. Bobbs-Merrill, Indianapolis.

Halvorson, H. (2012). What scientific theories could not be. *Philosophy of Science*, 79(2):183–206.

Halvorson, H. (2013). The semantic view, if plausible, is syntactic. *Philosophy of Science*, 80(3):475–478.

Halvorson, H. (2016). Scientific theories. In Humphreys, P., editor, *The Oxford Handbook of Philosophy of Science*, pages 585–608. Oxford University Press, New York.

Halvorson, H. and Tsementzis, D. (2017). Categories of scientific theories. In Landry, E., editor, *Categories for the Working Philosopher*, pages 402–429. Oxford University Press, New York.

Hawking, S. W. (1971). Stable and generic properties in general relativity. *General Relativity and Gravitation*, 1(4):393–400.

Hawking, S. W. and Ellis, G. F. R. (1973). *The Large Scale Structure of Space-Time*. Cambridge University Press, Cambridge.

Hesse, M. B. (1966). *Models and Analogies in Science*. University of Notre Dame Press, Notre Dame.

Humphreys, P. (2016a). Emergence. In Humphreys, P., editor, *The Oxford Handbook of Philosophy of Science*, pages 759–778. Oxford University Press, New York.

Humphreys, P. (2016b). *Emergence*. Oxford University Press, New York.

Ichikawa, J. J. and Steup, M. (2017). The analysis of knowledge. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford

University, spring 2017 edition. https://plato.stanford.edu/archives/spr2017/entries/knowledge-analysis/.

Jones, N. J. (2006). *Ineliminable idealizations, phase transitions, and irreversibility*. PhD thesis, The Ohio State University.

Konikowska, B. (1997). A logic for reasoning about relative similarity. *Studia Logica*, 58(1):185–226.

Kratzer, A. (1981). The notional category of modality. In Eikmeyer, H. J. and Rieser, H., editors, *Words, Worlds, and Contexts: New Approaches in Word Semantics*, pages 38–74. de Gruyter.

Kuhn, T. (1996). *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago, 3rd edition.

Lewis, D. (1973a). Causation. *Journal of Philosophy*, 70(17):556–567.

Lewis, D. (1973b). *Counterfactuals*. Blackwell, Oxford.

Lewis, D. (1981). Ordering semantics and premise semantics for counterfactuals. *Journal of Philosophical Logic*, 10(2):217–234.

Ludwig, G. and Thurler, G. (2006). *A New Foundation of Physical Theories*. Springer, Berlin.

Luper, S. (2016). Epistemic closure. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2016 edition. https://plato.stanford.edu/archives/spr2016/entries/closure-epistemic/.

Malament, D. B. (1977). Observationally indistinguishable space-times. In Earman, J., Glymour, C., and Statchel, J., editors, *Foundations of space-time theories*, volume 8 of *Minnesota studies in the philosophy of science*, pages 61–80. University of Minnesota Press.

Malament, D. B. (2008). Norton's slippery slope. *Philosophy of Science*, 75(5):799–816.

Malament, D. B. (2012). *Topics in the Foundations of General Relativity and Newtonian Gravitation Theory*. Chicago Lectures in Physics. University of Chicago Press, Chicago.

Manchak, J. B. (2009). Can we know the global structure of spacetime? *Studies in History and Philosophy of Modern Physics*, 40(1):53–56.

Mayr, D. (1981). Investigations of the concept of reduction II: Approximative reduction of theories with inaccuracy-sets: Uniform structures on theories, their completion and embedding. *Erkenntnis*, 16(1):109–129.

McMullin, E. (1985). Galilean idealization. *Studies in the History and Philosophy of Science*, 16(3):247–273.

Menzies, P. (2014). Counterfactual theories of causation. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2014 edition. https://plato.stanford.edu/archives/spr2014/entries/causation-counterfactual/.

Mormann, T. (1996). Similarity and continuous quality distributions. *The Monist*, 79(1):76–88.

Mormann, T. (2013). Topology as an issue for history of philosophy of science. In Andersen, H., Dieks, D., Gonzalez, W. J., Uebel, T., and Wheeler, G., editors, *New Challenges to Philosophy of Science*, pages 423–434. Springer Netherlands, Dordrecht.

Moulines, C. U. (1976). Approximate application of empirical theories: A general explication. *Erkenntnis*, 10(2):201–227.

Moulines, C. U. (1980). Intertheoretic approximation: The Kepler-Newton case. *Synthese*, 45(3):387–412.

Nagel, E. (1961). *The Structure of Science: Problems in the Logic of Explanation*. Harcourt, New York.

Nickles, T. (1973). Two concepts of intertheoretic reduction. *The Journal of Philosophy*, 70(7):181–201.

O'Connor, T. and Wong, H. Y. (2015). Emergent properties. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2015 edition. https://plato.stanford.edu/archives/sum2015/entries/properties-emergent/.

Pearl, J. (2009). *Causality: Modeling, Reasoning, and Inference*. Cambridge University Press, Cambridge, 2nd edition.

Post, H. (1971). Correspondence, invariance and heuristics: In praise of conservative induction. *Studies in History and Philosophy of Science*, 2(3):213–255.

Primas, H. (1998). Emergence in exact natural sciences. *Acta Polytechnica Scandinavica*, 91:83–98.

Pu, H. W. and Pu, H. H. (1974). Semi-quasi-uniform spaces. *Portugaliae Mathematica*, 33(3):177–184.

Rosaler, J. (2015). Local reduction in physics. *Studies in History and Philosophy of Modern Physics*, 50:54–69.

Sarkar, S. (2016). Ecology. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2016 edition. https://plato.stanford.edu/archives/win2016/entries/ecology/.

Schaffner, K. F. (1967). Approaches to reduction. *Philosophy of Science*, 34(2):137–147.

Schaffner, K. F. (1976). Reductionism in biology: Prospects and problems. In Cohen, R. S., Hooker, C. A., Michalos, A. C., and Van Evra, J. W., editors, *PSA 1974: Proceedings of the 1974 Biennial Meeting Philosophy of Science Association*, pages 613–632, Dordrecht. Springer Netherlands.

Scheibe, E. (2001). *Between Rationalism and Empiricism*. Springer, New York. Edited by Brigitte Falkenburg.

Schmidt, H.-J. (2014). Structuralism in physics. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2014 edition. http://plato.stanford.edu/archives/win2014/entries/physics-structuralism/.

Schmidt, J. C. (2011). Challenged by instability and complexity: On the methodological discussion of mathematical models in nonlinear sciences and complexity theory. In Hooker, C., editor, *Philosophy of Complex Systems*, volume 10 of *Handbook of the Philosophy of Science*, pages 223–254. Elsevier, Oxford.

Schreider, J. A. (1971/1975). *Equality, Resemblance, and Order*. Mir Publishers, Moscow, rev. edition. Martin Greendlinger, translator.

Sklar, L. (1967). Types of inter-theoretic reductionn. *British Journal for the Philosophy of Science*, 18(2):109–124.

Sorkin, R. D. (1995). A specimen of theory construction from quantum gravity. In Leplin, J., editor, *The Creation of Ideas in Physics: Studies for a Methodology of Theory Construction*, pages 167–179. Springer Netherlands, Dordrecht.

Steen, L. A. and Seebach, Jr., J. A. (1978). *Counterexamples in Topology*. Springer, New York, 2nd edition.

Swanson, E. (2011). On the treatment of incomparability in ordering semantics and premise semantics. *Journal of Philosophical Logic*, 40(6):693–713.

Teller, P. (1992). A contemporary look at emergence. In Beckermann, A., Flohr, H., and Kim, J., editors, *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism*, Grundlagen der Kommunikation und Kognition / Foundations of Communication and Cognition, pages 139–154. de Gruyter.

Thompson, P. (2007). Formalisations of evolutionary biology. In Matthen, M. and Stephens, C., editors, *Philosophy of Biology*, Handbook of the Philosophy of Science, pages 485–523. North-Holland, Amsterdam.

van Fraassen, B. (1989). *Laws and Symmetry*. Clarendon, Oxford.

van Riel, R. and Van Gulick, R. (2016). Scientific reduction. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2016 edition. https://plato.stanford.edu/archives/win2016/entries/scientific-reduction/.

Weatherall, J. O. (2016). Are Newtonian gravitation and geometrized Newtonian gravitation theoretically equivalent? *Erkenntnis*, 81(5):1073–1091.

Weisberg, M. (2013). *Simulation and Similarity: Using Models to Understand the World*. Oxford University Press, Oxford.

Willard, S. (1970). *General Topology*. Addison-Wesley, Reading, MA.

Wilson, M. (2006). *Wandering Significance*. Oxford University Press, Oxford.

Wilson, M. (2009). Determinism and the mystery of the missing physics. *British Journal for the Philosophy of Science*, 60(1):173–193.

Wimsatt, W. C. (1976). Reductive explanation: A functional account. In Cohen, R. S., Hooker, C. A., Michalos, A. C., and Van Evra, J. W., editors, *PSA 1974: Proceedings of the 1974 Biennial Meeting Philosophy of Science Association*, pages 671–710, Dordrecht. Springer Netherlands.

Wimsatt, W. C. (1997). Aggregativity: Reductive heuristics for finding emergence. *Philosophy of Science*, 64(Supplement):S372–S384.

Wimsatt, W. C. (2000). Emergence as non-aggregativity and the biases of reductionism(s). *Foundations of Science*, 5(3):269–297.

Winther, R. G. (2015). The structure of scientific theories. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2015 edition. http://plato.stanford.edu/archives/fall2015/entries/structure-scientific-theories/.

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