Symmetries and Representation

Forthcoming in *Philosophy Compass*

Geoffrey Hall, Sebastián Murgueitio Ramírez*

February 27, 2024

Abstract

It is often said in physics that if two models of a theory are related by a symmetry, then the two models provide (or *could* provide) two different representations of the very same situation, alike the case of two maps of different color for the very same city. It is also said that the situations represented by two models of a theory are indiscernible in some ways when the models in question are related by a symmetry of the theory, just like the situation in the interior of the cabin of a train when the train is at rest in the station is empirically indiscernible from the situation in the interior when the train is moving uniformly (in classical mechanics, these two situations are represented by two models related by a boost). In recent years, philosophers of physics have focused a lot of attention in developing various principles that aim to elucidate these and similar remarks on symmetries, models, physical equivalence, and representation that are widespread in physics practice. The goal of the current article is to provide a critical review of these principles, and suggest a new framework for thinking about these kinds of questions. One important upshot of the paper is that questions of indiscernibility, and questions of the representational capacity of models, must be distinguished from one another.

Keywords: Symmetry, theory, interpretation, models, equivalence, representation

Contents

*We want to thank Caspar Jacobs and two anonymous referees for their insightful feedback. Sebastián also wants to thank Neil Dewar for a helpful conversation on this topic.

Introduction

In January of 2001, a workshop dedicated to the philosophy of physics was held at the University of Oxford. It was titled "Symmetries in Physics: New Reflections" and was organized by Katherine Brading and Elena Castellani. The success of the workshop encouraged the organizers to create a collection of papers on the philosophy of symmetries titled "Symmetries in Physics: Philosophical Reflections" [2003]. In the words of Brading and Castellani: "It became clear from the success of the workshop, the enthusiasm and sense of shared work-in-progress, that the time is right for a collection of papers in philosophy of physics on the subject of symmetry" [2003, p. ix]. In part because of the influence of the papers in the collection (many of which were written by scholars still actively working on the area), research on the philosophy of symmetries has grown significantly in the two decades that have passed since then, to the point that it seems fair to say that the philosophy of symmetries has become an established sub-discipline in the philosophy of physics, distinct from (but obviously related to) the philosophy of quantum mechanics and the philosophy of spacetime. Although the field is now rather rich and varied, there are two questions that still guide a substantial amount of the conversation: what is the relationship between symmetries and representation? And what is the relationship between symmetries and measurement? The goal of the present paper is to offer a critical overview of some of the main positions around the first question, and suggest new ways in which the debate can be further developed (for a discussion of the second question, see the accompanying article to this series [REDACTED]).

Section 1 introduces the framework we will employ for talking about models, theories, representation, and symmetries. Section 2 introduces a hierarchy of indiscernibility principles that might be thought to govern the relation between the symmetries of a theory and the possible situations the theory represents. Section 3 introduces two principles concerning the relationship between symmetry and representation that are sometimes conflated in the literature. We then argue for the importance of distinguishing these principles from the indiscernibility principles introduced in section 2, and make more precise the idea that a model is apt at representing a situation. Finally, section 4 assesses the prospects of one of the most discussed indiscernibility principles in the literature according to which situations represented by symmetry-related models are empirically equivalent.

1 Preliminary notions

1.1 Theories and models

At a high level of abstraction, a symmetry of a theory is an invertible transformation between models of the theory that preserves the laws.[1] Consider, as a case study, the theory of classical mechanics, applied to a block sliding on the floor without friction. We can think of a model of the block in this context as a map $x : T \to S$ from a (mathematical) space T representing time to a (mathematical) space S representing possible states the block could be in at each time. Not just any map is a permissible model. At the core of the theory, one finds Newton's laws[2] together with special force laws that depend on the specific case: the force law for a spring is different from the force law for gravitation, and different from the force law for the friction produced by a surface, and

[1]There are different views about what symmetries are because there are different takes on whether symmetries must preserve certain restrictions in addition to preserving the laws. If the restrictions are purely formal, then one says that the definition of symmetry is *formal*. If one says that certain physical properties must be preserved, then one says that the definition is *ontic*. And if one demands that things such as observations are preserved, then one says that the definition is *epistemic*. See Baker [2022] , Read and Møller-Nielsen [2020] and Dasgupta [2016] for recent discussions about this debate.

[2]One can also formulate the theory through variational principles (e.g., Hamilton's principle), and there are interesting questions about the circumstances under which various formulations (e.g., Lagrangian vs Hamiltonian) are equivalent, and about whether, even if equivalent, some are more perspicuous than others. See North [2021] for an insightful discussion of some of these questions.

so on.[3] In applications, these laws can be encoded in certain equations $L(\cdot)$ that can be thought as constraining the space of possible models $x(t)$, in the sense that only those maps or functions $x(t)$ that satisfy the laws $L(x)$ are models, or *solutions*, of the theory. So in this case, if $x(t)$ represents the position of the block as a function of time, it will be a model or *solution* of Newtonian mechanics (applied to the block) only if it satisfies $\frac{d^2}{dt^2}x(t) = 0$, the law-equation stating that the acceleration of an object in the absence of external forces is zero.[4]

A theory such as classical mechanics can have different law-equations for different kinds of systems, such as a law-equation for gravity and a law-equation for springs, but also can have different law-equations for the *same* type of systems that differ over the number of objects in the system. The law-equation for a planet around the Sun, for example, is different from the law-equation of five planets around the Sun even though both differential equations are, of course, about gravitating objects.[5] For this reason, when we talk of a theory in terms of a solution space of models, we mean something rather more fine-grained than our intuitive sense of theory. For instance, we should distinguish between the classical theory of n-independent massive particles from the classical theory of $n+1$-independent particles because the space of solutions is, strictly speaking, different.

The above highlights a three-fold process for setting up a theory.[6] First, we encode the laws of the theory in terms of a system $L(\cdot)$ of equations. Second, we specify a space of *kinematically possible* models, by specifying the space of structures that the law-equations will constrain. For instance, in the above case of the block, we could think of the space of kinematically possible models as consisting of any function $x : T \to S$ from a structure T representing time and a structure S representing the position.[7] Both T and S may themselves have some structure (for instance they

[3] See Cartwright [1999, ch. 2] for a discussion regarding the relationship between Newton's second law and the special force laws.

[4] For instance, $x(t) = \cos(t)$ will not be a model of the theory for the block in this context, as it does not solve that law-equation. In contrast, $x(t) = vt$, where $v = \frac{d}{dt}x$ represents the speed, is a model (using the terminology that will be explained below, both functions would be *kinematically possible models* for the block but only the second one will count as a *dynamically possible model*).

[5] Both equations are derived from the equation for Newton's second law and the force function representing gravity. Wallace [2022a] uses "sector" to distinguish between law-equations dealing with different numbers of objects.

[6] We do not mean to imply that this is how scientists actually develop and set up theories, but it is a helpful way of thinking of the structures of rather mathematized theories of the sorts discussed in the philosophy of physics. Below we will complicate this picture a bit when we introduce the notion of an admissible interpretation of a theory.

[7] Typically, specifying the kinematically possible models of the law-equations requires a background structure on which differential equations can be defined. Such a structure is called a "jet bundle," and it allows us to specify not only the values of x at $t \in T$ but also the values of its ith-order derivatives. Only some among the kinematically possible models $x : T \to S$ defined in the jet bundle will solve the relevant law-equations of the theory (the ones that do solve them define a manifold in that space associated with the corresponding differential equations). For a quick introduction, see the appendix of Belot [2013]. For a thorough treatment, see Olver [2000].

need enough structure to interpret the law-equation $L(\cdot)$.) Third, we carve out the *dynamically possible* models from the space of kinematically possible models by specifying which models are solutions to, or satisfy, the law-equations. To go back to the block example, both the functions $x(t) = \cos(t)$ and $x(t) = t$ can be taken as kinematically possible models for the block, representing possible trajectories of the block as a function of time, but only the second function counts as a dynamically possible model because it is the one that satisfies $\frac{d^2}{dt^2}x(t) = 0$, the law-equation for Newton's second law in the absence of a force (recall that the block is moving in a friction-less surface).[8] *Symmetries then are invertible transformations that map solutions of the law-equations to solutions (i.e., map dynamically possible models to dynamically possible models) while preserving the underlying structure inherent to the space of kinematically possible models.*[9]

Models of a theory in this sense are mathematical objects (e.g., solutions to differential equations). However, within the context of physics and other sciences, they are used also as representational vehicles. A specific curve $x(t)$ in $\mathbb{R}^3$ satisfying the law-equations of Newtonian gravitation can represent a trajectory that a massive body follows. In such cases facts about the specific model $x(t)$ and the background space $\mathbb{R}^3$ license certain beliefs and assertions about some target situation in a given context, such as the Earth as it moves around the Sun. What are the representational properties of models?

1.2 Models and states of affairs

Here is a simple picture. For each model m of a theory T there are some states of affairs (or situations) that are the *admissible interpretations* of m, in the sense that m can be used to represent any of these situations.[10] The notion of an admissible interpretation requires some clarification. The term is often used for vague languages, for instance, when 'bald' is said to have many admissible interpretations. Usually, this means there are many different specific but similar properties that

[8] It is worth noting that there could be additional distinctions among models, such as that between the dynamical models that satisfy the boundary conditions or the initial conditions of a problem, and those that do not. See Wolf and Read [2023] and Murgueitio Ramírez [2021, Ch. 4] for recent work exploring the importance of boundary conditions when thinking of the models of a theory.

[9] Hence, an invertible transformation that does not preserve the underlying structure is not a symmetry, despite what Dasgupta [2016] seems to suggest with some examples that change the number of objects.

[10] We are using 'state of affairs' here as a bit of a placeholder. It should be thought to include temporally extended states of affairs, like the complete history of a given material object, possible situations that constitute subsystems of some larger system, as well as "complete" states of affairs, like entire possible worlds. When context makes clear what type of states of affairs a given model is representing (for instance some particular object in motion, a possible situation, a possible world, a world history), we will use a more specific language.

'bald' could be interpreted as picking out consistent with the norms of English and the intentions of the speaker.[11] Our use is more general. A state of affairs being an admissible interpretation of a model requires, at a minimum, that using that model to represent that state of affairs in a given context does not violate the norms and conventions in the community that uses this model. But the situations that can be represented by a given model may form a heterogeneous bunch: we do not build in at the outset any kind of thesis concerning indiscernibility of admissible interpretations.[12] These norms might involve certain sorts of quasi-compositional constraints on how one obtains admissible interpretations. For instance, mere stipulations like "m represents the state of affairs s" will not do, at least not normally. Instead, there are usually certain variables in the model that are constrained to be interpreted in certain ways, which in turn constrain which states of affairs the model can be used to represent.[13]

One might worry that the the notion of an admissible interpretation for models would require us to have some prior language sufficient to represent the states of affairs models can be used to represent. For instance, Dewar raises the worry that:

> Presumably, the interpretation [of the models] itself must be specified in some appropriate language. So we then must confront the question: why is that language in any better shape than the language with which we began? What fixes the interpretation of (the linguistic statement of) the interpretation? [Dewar, 2022, p. 63]

But the supposition that the models of a given theory have admissible interpretations does not require any such prior language. It is often only through understanding the details of the model and through actually using it to model real systems, that one is able to have beliefs concerning the states of affairs the model represents.[14] The question of in virtue of what a given state of

[11] See Keefe [2000] for further details on supervaluationism

[12] Another theory that could have perhaps worked equally well here is an indexical theory along the lines of Kaplan [1989] applied to the interpretation of models. On this theory, models get associated with a *character*, which is function contexts of use to states of affairs. Intuitively it maps each context to the specific states of affairs it is used to represent. The plurality of admissible interpretations associated with a model are then those states of affairs such that, for some context, the character of the model maps that context to one of those states of affairs.

[13] Because of these constraints, when we say that a model *can* represent a given state of affairs, we do not mean merely that it is possible that the model represents that state of affairs. The relevant notion of "can" here is being used in the same way that it is used when stating the rules of a game. Just as in chess, certain moves are admissible, in the sense that they are permitted by the rules of a game, in the context of interpreting a scientific theory, certain states of affairs are admissible as representational contents of certain models.

[14] See Giere [2010] for an approach to how scientists *use* various scientific models to represent physical systems. See also Cartwright [1999, ch. 2].

affairs is an admissible interpretation of the model is a hard question that we will not attempt to answer in full here. That being said, it is important to point out that there is a traditional approach to foundational theories of meaning that holds that the representational properties of public representational vehicles, like words in natural language or models of a given theory, must be explicable in terms of the mental representations of the agents who use those representational vehicles.[15] Arguably, theories like this are not apt in the context of models of a scientific theory since one's ability to mentally represent the situations the model can be used to describe depends on one's ability to use and understand the model in the first place. But there are alternative metasemantic frameworks more amenable to explaining the representational properties of models. For instance, in a broadly Lewisian framework, the correct assignment of admissible interpretations to models might be determined by a ranking of possible assignments by how well they fit use and the world.[16] This sort of view opens up an intriguing possibility, suggested by [Dewar, 2022, p. 63], that the symmetries of a theory play some role in determining which assignment better fits use. After all, moving between symmetry-related models constitutes an important part of the use of scientific models. Part of the goal of this paper will be to both spell out what various notions of equivalence might be relevant here and to distinguish them from one another. But ultimately, this kind of proposal is only plausible if assigning symmetry-related models the same, or similar, representational profiles really does better fit use.

Two points of clarification: first, the space of solutions to a given differential equation can be used to model different types of things in different contexts. One can use the same differential equation to model the vibrations of atoms in a crystal lattice and to model the behavior of stocks [Ahn et al., 2018]. Thus which interpretations of models are admissible depends on which theory we are talking about.[17] The second point we want to address is that of idealization in science. Models

[15] See Speaks [2010] for an overview of mentalistic foundational theories of meaning.

[16] The general framework is laid out across several papers, including Lewis [1975], Lewis [1983] and Lewis [1984]. A lot of subsequent discussion has focused on "reference magnetism," the controversial posit according to which some features of the world are easier to represent than others. But there is also a lot to be learned from the basic thought that metasemantics proceeds by determining which abstract assignment of representational content best fits use and the world, even if one is skeptical of the idea that some parts of the world are more eligible to be represented than other.

[17] For our purposes, it will suffice to think of a theory as being specified by a collection Δ of differential equations, a collection $\mathcal{M}$ of models (all the kinematic models, which include both solutions (dynamical models) and non-solutions to the differential equations), and a function f from models to the set of admissible interpretations. Which such function f from models to admissible interpretations is correct will depend on which theory we are talking about and the norms that the users of that theory employ.

of a given theory often involve heavy idealization.[18] For example, one might model the motion of a block attached to a spring in a lab by using a solution to the equation of a classical harmonic oscillator. But no real spring is precisely described by solutions to this equation since no spring is made of perfectly elastic material. There are, however, cases where a solution imprecisely describes the behavior of the real spring, and so provides a model of it. The question, then, is whether situations with real, non-idealized springs, are admissible interpretations of this model. One option is to recognize a distinction between those states of affairs that the model can "strictly speaking" be used to represent, and those that the model can be used to represent in some looser sense.[19] Another option is to only permit the idealized states of affairs as admissible interpretations, but to recognize that those states of affairs can themselves be used to model situations with real springs.[20] We will mostly stay neutral on this question in what follows except to flag points where it matters.

The main question investigated in this paper is 'What is the relationship between symmetries and representation?' Broadly speaking, we can distinguish two approaches to answering this question implicit in the literature. On one approach, symmetries and representation are jointly constrained by objective indiscernibility relations on the targets of representation: symmetry-related models are required to represent indiscernible situations, where a situation being indiscernible from another one is an objective, non-conventional matter. A second approach focuses solely on the relation between symmetries and the representational capacities of models. In the following sections, we lay out various ways of developing these two approaches and the relationships between them.

2 Symmetries and Indiscernibility Principles

An important proposal for elucidating the connection between symmetries and representation focuses on connecting symmetry to indiscernibility. In particular, any one of the following hierarchy of principles linking symmetry to indiscernibility might be thought to be an important constraint on model representation:

[18]For a recent and detailed discussion of idealization in science, see Potochnik [2020].

[19]This gives rise to two interpretation functions f_{precise} and $f_{\mathsf{imprecise}}$: one maps each model to those states of affairs it can precisely represent, and the other maps each model to those states of affairs it precisely represents *and* those that it imprecisely represents.

[20]This latter sort of theory posits an ambiguity in the use of "model" in the literature: there are mathematical models, and these provide precise descriptions of idealized states of affairs; these idealized states of affairs then constitute scientific models of the actual, messy, states of affairs in the concrete world.

Absolute Indiscernibility For any theory T and models m and m' of T representing situations s and s', if m and m' are related by a symmetry, then s and s' are absolutely indiscernible.

Qualitative Indiscernibility For any theory T and models m and m' of T representing situations s and s', if m and m' are related by a symmetry, s and s' are qualitatively indiscernible.

Physical Indiscernibility For any theory T and models m and m' of T representing situations s and s', if m and m' are related by a symmetry, s an s' are physically indiscernible.

Empirical Indiscernibility For any theory T and models m and m' of T representing situations s and s', if m and m' are related by a symmetry, then s and s' are empirically indiscernible.

Lawlike Indiscernibility For any theory T and models m and m' of T representing situations s and s', if m and m' are related by a symmetry, then s and s' are indiscernible with respect to the laws.

Here two situations being Fly indiscernible means something like agreeing on all of the F-facts, or agreeing on the pattern of instantiation of F type properties.[21] For example, one can think of Galileo's famous ship thought experiment [2022] as an illustration of Empirical Indiscernibility in the following way: if two models m and m' of classical mechanics are related by a constant velocity transformation, then the situation s associated with the objects in the *interior* of the ship's cabin when the ship is at rest with respect to the shore is empirically indiscernible from the situation s' associated with those same objects when the ship is moving uniformly.[22]

Indiscernibility principles are most often discussed in the context of spacetime theories which are taken to represent a set of possible worlds.[23] The symmetries are smooth permutations of

[21] Principles like these are not always explicitly discussed in these terms in the literature but are often implicit in debates about theoretical equivalence. The list is meant to be representative of some of the more theoretically interesting notions of indiscernibility appealed to but not exhaustive. For example, one might suggest that a theory comes equipped with resources that determine which properties are empirically detectable according to it. If that is the case, then one might want to distinguish between situations that are empirically indiscernible according to a theory, and situations that are empirically indiscernible in a more absolute manner (regardless of the theory). We read Read and Møller-Nielsen [2020] as developing an argument along these lines.

[22] Notice that the proper treatment of idealization matters when evaluating indiscernibility principles. Every model is symmetry related to itself, and so if a given model can represent both a situation with an idealized spring and a situation with a non-idealized spring, it is unlikely that any indiscernibility principle can hold, for situations with idealized springs are discriminable in all sorts of ways from situations with ordinary springs.

[23] The information that such models provide about worlds includes facts about the geometry of spacetime, which paths in spacetime objects of various types follow, which paths they *can* follow, as well as facts about the distribution of various fields over spacetime.

points that preserve certain features of the geometry of spacetime, such as relative distance.[24] Now suppose one had a symmetry that intuitively represents a shift of all matter three meters to the right.[25] Then, there is some intuitive appeal to principles like Empirical Indiscernibility: after all, agents and measurement devices in these worlds would behave just the same in the initial world and the shifted world.[26] But once one has granted Empirical Indiscernibility, there are broadly theoretical grounds for "leveling up." For suppose one provides an interpretation of a theory on which the situations represented by symmetry-related models are empirically equivalent but not physically equivalent. Then by the lights of this interpretation, the theory posits physical facts that play no role in the empirical predictions of the theory, which many regard as a theoretical cost.[27] [28]

In our view, even if some such principles are plausible for spacetime theories, they don't generalize to symmetries more broadly (as we will see in section 4, even Empirical Indiscernibility seems to fail). That being said, we do think that Lawlike Indiscernibility together with an indiscernibility principle on which the type of indiscernibility is permitted to change from theory to theory may be defensible and important. This principle says the following:

Salient Indiscernibility For any theory T and models m and m' of T representing situations

[24] Though things are more complicated in the case of general relativity where there isn't any kind of fixed metric on spacetime but rather many different possible metrics that reflect different curvature depending on the distribution of matter.

[25] This is a standard kind of case that has been debated since at least the Leibniz-Clarke correspondence [Clarke, 1717].

[26] See Dewar [2022] for an extended argument that infers that symmetry-related models are empirically equivalent from the fact that the dynamics of measuring devices are preserved under symmetries. For a critical discussion, see the accompanying paper to this series [REDACTED] and the references there.

[27] Similar sorts of defeasible "leveling up" arguments could be given for Qualitative Indiscernibility (and perhaps even Absolute Indiscernibility). For instance, one might motivate physicalism with respect to our spacetime theory so that physical indiscernibility implied qualitative indiscernibility on the grounds that positing qualitative differences without physical differences is a theoretical cost. The argument here though would have to contend with the very large literature on whether qualitative phenomenal properties of subjects supervene on physical properties, however. In turn, one could try to motivate Absolute Indiscernibility from Qualitative Indiscernibility by motivating anti-haeceitism, according to which there are no merely non-qualitative differences between worlds (see Baker [2010] for an argument along these lines, and Ismael and van Fraassen [2003] for a different approach towards defending Qualitative Indiscernibility). Here too one could appeal to the theoretical cost of positing too much structure. In short, then, symmetries of spacetime theories might motivate some of these stronger indiscernibility principles, and it is fair to say that a substantial amount of the recent philosophical literature on symmetries have focused on the spacetime case

[28] But even in the spacetime case there are subtle issues, as Belot [2017] has recently argued. Philosophers have also used gauge symmetries in theories like electromagnetism to motivate indiscernibility principles. One question that has been studied in the recent philosophy literature is whether, contrary to what has been assumed for many years, there are cases in which gauge-related models do not represent the same states of affairs. The consensus used to be that gauge-related models always represent the same states of affairs, but the community is a bit more divided now. See Brading and Brown [2004], Greaves and Wallace [2014], Teh [2015], Friederich [2015], Murgueitio Ramírez [2023] and Gomes [2019] for discussions.

s and s', there are some salient properties $F_1, \ldots, F_n$ such that m and m' are related by a symmetry only if s and s' are indiscernible with respect to the pattern of instantiation of $F_1, \ldots, F_n$.

The notion of "salience" depends on the theory.[29] Salient properties could be properties represented by the underlying structure of the solution space. So for a simple example, take two models of classical mechanics m and m' representing the world history of two particles in Newtonian space. Then, one salient property preserved by a symmetry going from m to m' is the relative distance between two particles in the situation described. In our view, Salient Indiscernibility and Lawlike Indiscernibility may function as something like conventional constraints on representation: given two symmetry-related models, there is a semantic constraint on admissible interpretations requiring that any admissible interpretation of the one be a situation that agrees on both the laws and the pattern of instantiation of the salient properties with any admissible interpretation of the other (like relative distance in the case of Newtonian gravitation[30]). But this leaves open a good number of questions: for instance, whether situations that agree on laws and the pattern of distribution of the salient properties are empirically or physically equivalent. It also leaves open another question, to which we now turn: are symmetry-related models associated with the same set of admissible interpretations?

3 Representational capacities

3.1 Two different principles

The second approach to clarifying the relationship between symmetries and representation focuses on the representational capacities of models. Many authors working on symmetries have thought that symmetries establish something like a relation of "synonymy" between models of a theory:

> If two models of a theory are related by a symmetry, those models represent the same possibility [Baker, 2022, p. 1784].

> Two states of affairs related by a symmetry transformation are really just the same state of affairs differently described [Greaves and Wallace, 2014, p. 60].

[29] We read Read and Møller-Nielsen [2020] as defending a version of this principle.

[30] However, as we discussed in the last section, some symmetries of the Kepler problem do change the eccentricity of planetary orbits, and so the distance.

A related, but importantly different, principle is articulated in the following quotations:

> Two solutions of a classical theory's equation of motion are related by a symmetry if and only if ... they are equally well- or ill-suited to represent any particular physical situation. [Belot, 2013, p. 1].

> [...] if a particular mathematical model may be used to represent a given physical situation, then any isomorphic model may be used to represent that situation equally well. [Weatherall, 2018, p. 332].

> Two models of a physical theory are symmetry-related iff they can represent the same possible physical situations [Luc, 2022, p. 72].

The thesis that symmetry-related models do represent the same thing is different from the thesis that symmetry-related models *could* represent the same thing. Thus we have the following two theses[31]:

Sym1 Any two symmetry-related models of a physical theory are equally apt to represent the same things under the correct interpretation of the theory.[32]

Sym2 Any two symmetry-related models of a physical theory do represent the same thing under the correct interpretation of the theory.

It is important to distinguish between **Sym1** and **Sym2** since there are cases in which distinct symmetry-related models can be used within a single context to represent distinct but symmetric states of affairs. The choice of which model to use to represent which state of affairs in such cases is often arbitrary.[33] For example, one can use the model $x_r(t) = vt + c$ to represent the position of a red car moving with constant velocity v that starts the motion next to a tree, and the model $x_b(t) = vt + c + d$ to represent the position that car would have had if it has started moving some

[31] Note that the second set of quotations are actually bi-conditionals, suggesting that perhaps representational properties of models could themselves play some role in constraining which transformations are symmetries, as opposed to symmetries constraining the space of admissible interpretations of a theory. There is a literature here on whether symmetries are formal, ontic or epistemic (e..g., see Read and Møller-Nielsen [2020]).

[32] In our view, one can assess the status of these principles without providing a fully worked out theory of what interpretations of *theories* are. However, a sketch might be the following: an interpretation of a theory is the assignment of admissible interpretations to models of a theory. Models m and m' are equally apt to represent the same situations given the correct interpretation of theory if and only if the interpretation that best matches use $+X$ assigns the same situations as admissible to either, where X is whatever else goes into the correct foundational theory of meaning in addition to use.

[33] For more on the importance of distinguishing **Sym1** and **Sym2**, or at least theses like them, see Luc [2022], Murgueitio Ramírez [2021, ch. 4], Fletcher [2020] and Roberts [2020].

meters ahead from the same tree. The two models are related by a symmetry of Newton's second law for zero force, namely, by a translation.[34] Hence, we have two symmetry-related models of a theory that represent different things under the correct interpretation of the theory, and so this serves as a simple counterexample to **Sym2**.[35] Notice, however, that the prior example is consistent with **Sym1**, for we could have chosen to flip the models, representing the car when starting next to the tree using $x_b(t) = vt + c + d$ and when starting some meters ahead using $x_r(t) = vt + c$ (where d now represents a negative displacement). Hence, these two symmetry-related models are apt to represent the same things, even though, within a given context, they do fail to actually represent the same thing.[36]

Now, it is often assumed that principles like **Sym1** and **Sym2** ought to be understood as restricted to global or universe symmetries, that is, symmetry transformations of the whole material content of a world as it happens in the Leibniz-Clarke correspondence (see Clarke [1717]). In such cases, the prior example of the car and the tree does not quite work, for there we were considering symmetry transformations–spatial shifts–of the car alone. Once **Sym1** and **Sym2** are restricted to global symmetries of models interpreted as representing the whole material content of the world, there aren't going to be uncontroversial cases in which **Sym1** is true and **Sym2** is false because there are not going to be uncontroversial cases in which **Sym2** is false at all. Many philosophers accept the principle according to which there are no "shifted" nomologically possible worlds. Our goal here, however, is merely to point out a logical distinction between **Sym1** and **Sym2**. And while it may be *true* that there are no distinct nomologically possible worlds related by a symmetry, this doesn't strike us as a logical truth, but rather a substantive metaphysical hypothesis. In the next section, we will discuss the relationship between shifted worlds and **Sym2** in more depth. The point we want to make here is just that accepting **Sym1** while rejecting **Sym2** is at least a coherent position to take, and so it is important to distinguish these principles.

[34] In particular, $x \mapsto x + d$ is a symmetry of $\frac{d^2}{dt^2}x = 0$, which is Newton's second law for zero force.

[35] See also Fletcher [2020, Sect. 3] for a discussion of other ways in which **Sym2** fails.

[36] The first thing to note is that **Sym1** and **Sym2** are independent of one another: **Sym1** does not entail **Sym2** nor vice versa. To see that **Sym2** does not entail **Sym1**, just note that while two models might diverge in terms of their aptness for representing some situation given the correct interpretation of the theory, it doesn't follow that they *do* represent anything at all. For instance, perhaps models only represent in a content, somewhat similarly to the way that proper indexicals like 'here' and 'I' don't represent anything *simplicter* but only represent something when placed in a suitable context.

3.2 Sym1, Sym2, and Indiscernibility

Up until this point, we have considered two different approaches aiming at clarifying the connection between symmetries and representation, one focused on indiscernibility principles and one focused on the representational capacities of models. Now we will investigate the connection between these approaches. In particular, how do the requirements imposed by **Sym1** or **Sym2** relate to the hierarchy of indiscernibility principles discussed earlier? Consider the case of Newtonian gravitation for point-like masses. A model, roughly, is a smooth association of indices t representing time with a list of points $x(t)$ representing the positions of particles in space at the time represented by t. Given a model like this, we can consider a symmetry-related model $x'(t)$ by translating the position of each particle at each time by some fixed distance (this is a shifted model). Now if we imagine that the model is representing an entire possible world, then the effect of this translation leaves all of the relative distances between all material objects in that world fixed, and it leaves the pattern of dynamical behavior of all material objects fixed.[37] Many philosophers have thought such cases are paradigmatic examples of worlds satisfying some strong indiscernibility principle: the translation from $x(t)$ to $x'(t)$ is simply a re-description of the same world we started with.[38] Our question is: does **Sym2** enforce any sort of strong indiscernibility principle here?

One might think that it does. For one could say that **Sym2** ensures that a $x(t)$ and $x'(t)$ in fact represent the same world, to think otherwise is to confuse difference in how $x(t)$ and $x'(t)$ represent with difference in what $x(t)$ and $x'(t)$ represent. This could be so, but it's worth pointing out that this goes beyond what **Sym2** says. To see this consider the following theory of how representation works. Suppose that **Sym1** is true and that, moreover, in a given context, each model represents all of the worlds it can be used to represent. Finally, suppose that $x(t)$ represents a world w and $x'(t)$ represents a shifted world w' that is discernible from w. Then by **Sym1**, w' is an admissible interpretation of $x(t)$ and w is an admissible interpretation of $x'(t)$. And so if models always plurally represent all of their admissible interpretations, it follows that $x(t)$ represents a world if and only

[37] Since the gravitational force acting on a particle depends only on the relative distances between particles, such translations preserve solution-hood.

[38] However, arguably, this translations does not leave absolute position fixed. One model represents particle i as being located at the position in space represented by a certain point p at the time represented by t and another in the position represented by $p + y$. In a Newtonian world, these would be intuitively different: one could point to some region of space three meters from them and say $p + y$ represents *that* and p represents *this* demonstrating their own location. See Bryan and Read [2022] for a recent discussion on indexicals in the context of symmetries.

if $x'(t)$ does. Thus **Sym2** is secured; but we've taken no stand on the question of discernibility of the relevant worlds.[39] And so in particular it doesn't take a stand on whether w and w' being admissible interpretations of symmetry-related models enforces any kind of indiscernibility in w and w'.

This shows that given certain possible representational conventions, **Sym2** comes apart from indiscernibility. But are these conventions plausibly our own? Couldn't one stipulate that models represent only one thing? It is of course open to one to say this, though it is very important to recognize that mere act of stipulating that something is so does not guarantee it to be so: the stipulations one makes elsewhere might constrain what can be consistently stipulated here, and moreover, the conventions that are present in the community might bar those stipulations from going into effect. In other words, *not just any set of stipulations on a representational relation is guaranteed to be satisfied*. There are trivial cases of this (no representational relation satisfies the stipulation "x does and does not represent y"). But there are also subtler cases. For instance no representational relation satisfies the stipulation "each natural number represents a unique real number and every real number is represented by some natural number". In the present case, what is important is what symmetries themselves can be used to represent in a given community. For instance, suppose symmetries can be given an "active" reading, on which they represent the result of uniformly redistributing matter in a certain way (e.g., three meters to the right). This seems like a possible convention. But once this convention is in place, it simply isn't in one's power to stipulate, holding conventions fixed, that models related by a symmetry both represent the same things, and that each model represents one and only one world. The world must cooperate in order for any such stipulation to succeed. If there are discernible worlds in which all particles are shifted, then given that symmetries themselves *can* be read "actively" – as corresponding to a real shift in the world – no stipulation would succeed in getting symmetry related models to all pick out one and only one of those shifted worlds.[40]

[39] See Jacobs [2021] for a response along these lines against "Leibniz Equivalence," a principle similar to **Sym2** that says that symmetry-related models (invariably) represent the same state of affairs.

[40] This is relevant regarding the "motivationalist" vs "interpretationalist" debate about symmetries (for a detailed overview of this debate, see Luc [2023]). In particular, what we say here suggests that an interpretationalist cannot simply declare that symmetry-related models are equivalent (e.g., see Dewar [2019]) because even if they try to stipulate that, the stipulation might simply fail to achieve anything because of background norms and stipulations elsewhere. For more on the debate about whether principles in the vicinity of **Sym1** and **Sym2** could be used to infer interesting conclusions about shifted worlds see, in particular, Teitel [2021] and Jacobs [2023].

3.3 What is the scope of Sym1?

Consider, again, the example of the car. We already mentioned that $x_r(t) = vt + c$ and $x_b(t) = vt + c + d$, which are related by a symmetry of the equation for Newton's second law (for the case of zero force), are equally capable of representing the car's motion in a case where it starts moving next to a tree, and in a case where it starts moving some meters ahead of the tree. But what does it mean to say that a certain model of a theory is apt at representing a particular situation?[41]

As noted earlier, $x_r(t) = vt + d$ and $x_b(t) = vt + d + c$ are both apt at representing the car's motion. But why is $x_w(t) = 0.5at^2 + vt + c$, which is not symmetry-related to the other two models, not apt to do the same? After all, $x_w(t)$ seems apt to represent the car provided that one understands this model as representing the motion from the perspective of a non-inertial frame that has acceleration a with respect to the road. But if one says this, then it seems that non-dynamical models are equally apt at representing the same things as dynamical models, and so **Sym1** seems to become much less interesting. To address this worry, note the rather obvious point that part of the reason (if not the whole reason) $x_r(t)$ and $x_b(t)$ are apt at representing the motion of the red car is that they solve the relevant law-equation characterizing the car's behavior. In contrast, $x_w(t)$ is not a solution of that equation, and so it cannot capture the correct behavior. More generally, a minimum requirement for a model m to be apt at representing a system is for it to represent the system as obeying the laws it actually obeys.[42] And if m satisfies this minimum requirement, so does any other model m' related to m by a symmetry of the law-equation, after all, the symmetries of an equation necessarily map (relate) solutions of the equation to solutions (the same cannot be said about a model m'' that is related by a non-symmetry transformation to m).

Having said this, it is true that $x_w(t)$ would be a good way of representing the motion of the car from the perspective of an accelerating frame. However, if one does adopt the latter perspective, one must then invoke different law-equations, for instance, equations that relate this apparent acceleration of the car to inertial or "fictional" forces, and in this new context, $x_r(t)$ no longer is apt (it is not a solution of the "new" law-equations). Thus, when **Sym1** talks about two models being

[41] It is worth pointing out that the symmetry literature has focused most of their attention on the question of whether two models are "equally capable" or "equally apt" to represent the same situation, but have not discussed in detail the question of what it means to say that a particular model is capable of representing a particular situation.

[42] Up to certain approximations that depend on the context, e.g., at low speeds, we do not need to consider relativistic laws.

equally apt at representing the same thing, this really means equally apt given certain conventions about how to use the theory in a particular context (e.g. in the context of an inertial frame), it does not mean that they are equally apt given any conventions whatsoever (the latter would indeed make **Sym1** much less interesting).

Say, then, that the aptness of a model to represent a situation requires that the model satisfies the law-equations that characterize the laws of the situation. Is that sufficient? We do not intend to answer this question here, but we do want to make two remarks. First, the following requirement is implausible: for a model to be apt at representing an object in a situation, the model must have enough resources, say enough variables, to represent each and every physical property of that object in that situation.[43] Very few (if any) of the models used in science satisfy such a strong requirement, and presumably, this is why the models are useful, to begin with. But arguably, we do want the model to capture what we take to be the salient properties of a certain situation. An important part of being able to capture these properties requires being able to capture the laws that connect them. This is why the model needs to solve the law-equations associated with the situation. But the laws are, presumably, not all that matters in concrete applications. There are, in addition, the so-called "initial conditions" that give us additional information about the history of an object, such as what the particular state of a system at a certain moment in time is (the "boundary conditions" are also important, as these encapsulate how does the system behave around its spatial boundaries).[44]

In general, whether or not two models can capture the same initial conditions and the same boundary conditions depends on the case. For example, say that the car is next to a tree exactly at noon (this is an initial condition for the car), and say that we model the car using $x_r(t)$. Let's stipulate, for example, that $t = 0$ represents noon. Hence, $x_r(0) = c$ will represent the location of the car as it passes by the tree, and so this model is capable of representing this initial condition. But given these same stipulations, it seems clear that $c + d$ is *not* apt at representing this same initial condition, and so, given these stipulations, $x_b(t)$ is not apt to represent the initial condition

[43] In some cases philosophers talk of models of physical theories representing possible worlds. By this, they do not mean that the model explicitly represents each and every fact that obtains in that world. Instead, the idea is that the model represents something like a supervenience base for the physical facts in that world. So while the model may not represent some physical fact f that obtains in w, it will represent some set of facts X such that worlds that agree on the facts in X and the laws agree on f.

[44] It is worth noting that the term "initial condition" is a bit vague, as sometimes it is used to refer to physical properties of an object in a situation (e.g., the car is by the tree at noon), and sometimes to refer to the specific values of a curve or model at a Cauchy-surface (e.g., an initial condition might be $x = c$ when $t = 0$).

of the car. One can, of course, use $x_b(t)$ to represent the car and this initial condition, but then one has to stipulate that $c + d$, not c, represents the car's location at noon. So each model is apt at representing the car and the relevant initial condition but only if one uses different conventions for each model. In general, however, there is no guarantee that if a certain model can capture both the initial conditions and the boundary conditions given certain (accepted) conventions, any model symmetry-related to the first one can also capture the same things with other accepted conventions. After all, finding the class of models that can satisfy the boundary conditions for a given situation is a non-trivial task, one that goes beyond finding the symmetries of the equations.[45]

Second, and related, even models that on the surface seem rather different from one another can be apt at representing the same things, including the same initial conditions. For example, a model such as $x(t) = 0$ (zero for all times) can be apt at representing the car's motion, provided that one assumes an inertial frame that is co-moving with the car. Similarly, and contrary to some things Belot [2013] suggests, even $x(t) = 0$, which is symmetry-related to other solutions of the spring equation, can be apt at representing the oscillations of a given spring. However, to do so, one needs to assume a frame that is oscillating together with the spring, in a way analogous to a frame that is co-moving with the car. Whether or not that kind of frame is natural is up for debate. In any case, we want to stress that the existence of models that represent the correct laws and on the surface look very different from one another does not, on its own, pose a problem for **Sym1** precisely because the considerable flexibility in the kinds of acceptable conventions adopted by physicists when using them. Having said this, we do think that some examples like these, of models that seem to be very different from one another, do raise a challenge to some of the indiscernibility principles discussed earlier. We study one such case in the next and last section.

4 Symmetries and Empirical Indiscernibility

Consider again two models of Newtonian gravitation, m and m', that are related by a translation d. Say that m represents w, and m' represents a shifted world w' in which all the objects are ten

[45] See Bluman and Kumei [1990, ch. 4] for a thorough treatment of symmetries of differential equations with boundary conditions, and see Murgueitio Ramírez [2021, 124-127] for a philosophical discussion applying some of these resources to the problems raised by Belot [2013]. In light of these considerations, one might propose a modification of **Sym1** that says something like "models related by boundary-preserving symmetry-transformations are capable of representing the same things." One might be able to address some of the challenges raised by Belot in [2013] and [2017] by considering a modified version of **Sym1** along these lines.

meters to the right compared to their (absolute) locations in m. As presented, these models seem to represent distinct but empirically indiscernible states of affairs. For a long time, it was assumed in the philosophy literature that this is not an isolated example but that, quite generally, for any physical theory, symmetry-related models always represent empirically indiscernible situations.[46] However, in the last ten years or so, scholars started to point out the existence of symmetry transformations that seem to represent empirically very distinct states of affairs (e.g., see Belot [2013]). To give a concrete example, discussed in detail by Murgueitio Ramírez [2024], there are symmetry transformations of the law-equation for an ideal spring that seem to represent empirically very distinct states of that spring.[47] Consider, as an illustration, an astronaut flying outside the ISS who uses both hands to compress a little spring (one hand at each end). When they are far enough from the ISS, the astronaut suddenly releases the spring into outer space. What would the spring do? It will start to oscillate, as it tries to restore its equilibrium position. In contrast, if the astronaut had released the spring without compressing it or stretching it first, then the spring would have remained in equilibrium (it would not have oscillated). One can represent the case in which the spring starts to oscillate with a model like $x(t) = A\cos(t)$, and the case in which it does not with a model like $x(t) = 0$. Crucially, these two models *are* related by a symmetry of the equation for an ideal spring (see Murgueitio Ramírez [2024]). And so Empirical Indiscernibility seems to fail in such a case.

When faced with counterexamples, one way to respond is by offering a restriction of the proposed principle that captures the paradigmatic cases while avoiding the counterexample. There are at least two important proposals along these lines in the literature when addressing cases such as the spring just mentioned. The first one is to restrict the class of models Empirical Indiscernibility applies to to models that (a) represent whole universes, as opposed to subsystems within the universe, and (b) to global symmetries between those models (e.g., see [Luc, 2022, p. 25]). In this reading, all that Empirical Indiscernibility says is that nomologically possible worlds represented by symmetry-related models are empirically indiscernible. In order to present a challenge to this restricted version of Empirical Indiscernibility, the above example needs modification. Consider the following two possible worlds. In the first, the only material object that exists is a spring that is always

[46] For some examples of philosophers defending this or similar theses, see Roberts [2008], Healey [2009], Dewar [2015] , and Dasgupta [2016].

[47] For technical details, see Wulfman and Wybourne [1976, p. 516].

oscillating in just the way that it oscillates in the above case after the astronaut releases it from an initial state of compression. In the second, the only material object that exists is a spring that has always been at rest in equilibrium, with no oscillation whatsoever. One can then use model $x(t) = A\cos(t)$ to represent the behavior of the oscillating spring in the first world and $x(t) = 0$ to represent the non-oscillating spring in the second world. As before, these models are related by symmetry. But intuitively, the worlds that they represent are empirically discernible by virtue of containing springs that instantiate empirically discernible states (i.e., states of oscillation and states of equilibrium).

The second line of response is to instead restrict Empirical Indiscernibility to those symmetries of the target system that can be *extended* to include symmetries of the measurement device as well. Thus, if the symmetry in question is not well-defined for whatever object has been explicitly or implicitly used to measure states of the original target system, then we are no longer within the scope of Empirical Indiscernibility according to this proposal (see Wallace [2022b] and also Luc [2022]). In the present case, one might require the symmetries that relate the oscillatory spring and the immobile spring to count as symmetries of a ruler used to measure the difference in the length of the spring. As these transformations are not symmetries of a (rigid) ruler, one can then say that this case is out of the scope of Empirical Indiscernibility. Authors who have strong instrumentalist inclinations might be sympathetic to this kind of response, but those who think that it makes perfect sense to say that two or more states are empirically distinct without having to appeal to some measurement device that measures the difference between the states might think that restricting Empirical Indiscernibility in this manner is artificial.[48]

[48]Similar considerations apply to other symmetries that have been discussed in the literature, like those in the Kepler problem that relate the orbits of a planet around a much more massive object like the Sun to other orbits with different eccentricity (see Belot [2013], Murgueitio Ramírez [2021, ch. 4] and Wallace [2022a] for a philosophical discussion, and see Prince and Eliezer [1981] for a physics discussion.) Even if the universe only consisted of a planet around a star, and even if there was no ruler or detector capable of measuring anything, a plausible case can be made for the claim that a perfectly circular orbit, represented with a model m, and an orbit with high eccentricity, represented with a model m', would correspond to empirically distinct situations. Hence, the fact that one can find a symmetry of the theory (i.e., a symmetry of the equation for the Kepler problem) that relates m and m' is a potential problem for Empirical Indiscernibility. In future work, it would be worth exploring if there are other ways of resisting these kinds of cases (different from the two strategies discussed above), perhaps by requiring the symmetries to preserve certain boundary conditions (or even initial conditions) along the lines hinted at in the prior section.

5 Conclusion

This paper considered several different principles connecting the symmetries of a physical theory and representation that intend to capture the ways that some models are used in some physical theories (e.g., **Sym2** might capture how different gauges can be used when describing the same electromagnetic phenomena). However, it is unclear whether most of the principles are sufficiently general to serve as constraints on the interpretation of physical theories more broadly. In the majority of cases, the use to which models are put seems to require nothing more than Lawlike Indiscernibility. In particular, when using models to represent that various states of affairs are so, one pragmatically important constraint is that our models can agree on certain facts about the *type* of behavior of objects even when used from different perspectives, like in the case of different inertial frames. Thus, whether symmetries are interpreted actively or passively, symmetry-related models *ought* to agree on the laws of the situations they are representing. On some occasions, one might also require the models to agree on other facts, such as the initial and boundary conditions (or on the values of some other quantities), in which case one might also appeal to Salient Indiscernibility. But these two principles radically underdetermine the space of possible situations a given model is apt to represent. Hence, most of the principles discussed in the literature and examined in detail in this paper seem to occupy a small part of a larger and rather unexplored picture connecting symmetry and representation.

References

K. Ahn, M. Y. Choi, B. Dai, S. Sohn, and B. Yang. Modeling stock return distributions with a quantum harmonic oscillator. *Europhysics Letters*, 120(3):38003, January 2018. doi: 10.1209/0295-5075/120/38003.

David John Baker. Symmetry and the Metaphysics of Physics. *Philosophy Compass*, 5(12):1157–1166, 2010. ISSN 1747-9991. doi: 10.1111/j.1747-9991.2010.00361.x.

David John Baker. What Are Symmetries? *Ergo: An Open Access Journal of Philosophy*, 9(n/a), 2022. doi: 10.3998/ergo.3592.

Gordon Belot. Symmetry and Equivalence. In Robert Batterman, editor, *The Oxford Handbook of Philosophy of Physics*, pages 318–339. Oxford University Press, 2013.

Gordon Belot. Fifty Million Elvis Fans Can't be Wrong. *Noûs*, pages 946–981, 2017. ISSN 1468-0068. doi: 10.1111/nous.12200.

G.W. Bluman and S. Kumei. *Symmetries and Differential Equations*. Applied Mathematical Sciences. Springer New York, 1990. ISBN 978-1-4757-4309-8.

Katherine Brading and Harvey R. Brown. Are Gauge Symmetry Transformations Observable? *The British Journal for the Philosophy of Science*, 55(4):645–665, 2004. ISSN 0007-0882. doi: 10.1093/bjps/55.4.645.

Katherine Brading and Elena Castellani, editors. *Symmetries in Physics: Philosophical Reflections*. Cambridge University Press, 2003. ISBN 978-0-521-82137-7. doi: 10.1017/CBO9780511535369.

Cheng Bryan and James Read. Shifts and reference. In Antonio Vassallo, editor, *The Foundations of Spacetime Physics: Philosophical Perspectives*. Routledge, 2022. ISBN 1-03-210720-0.

N. Cartwright. *The Dappled World: A Study of the Boundaries of Science*. Cambridge University Press, 1999. ISBN 978-0-521-64411-2.

S. Clarke. *A Collection of Papers, Which Passed between the Late Learned Mr. Leibnitz and Dr. Clarke in the Years 1715 and 1716 : Relating to the Principles of Natural Philosophy and Religion*. 1717.

Shamik Dasgupta. Symmetry as an Epistemic Notion (Twice Over). *The British Journal for the Philosophy of Science*, 67(3):837–878, 2016. ISSN 0007-0882. doi: 10.1093/bjps/axu049.

Neil Dewar. Symmetries and the philosophy of language. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 52(Part B):317–327, 2015. ISSN 1355-2198. doi: 10.1016/j.shpsb.2015.09.004.

Neil Dewar. Sophistication about symmetries. *The British Journal for the Philosophy of Science*, pages 485–521, 2019. ISSN 0007-0882. doi: 10.1093/bjps/axx021.

Neil Dewar. *Structure and Equivalence*. Cambridge University Press, 2022. ISBN 1-108-91046-7.

Samuel C Fletcher. On Representational Capacities, with an Application to General Relativity. *Foundations of Physics*, 50(4):228–249, 2020. ISSN 0015-9018. doi: 10.1007/s10701-018-0208-6.

Simon Friederich. Symmetry, Empirical Equivalence, and Identity. *The British Journal for the Philosophy of Science*, 66(3):537–559, 2015. ISSN 0007-0882. doi: 10.1093/bjps/axt046.

Galileo Galilei. *Dialogue on the Two Greatest World Systems.* Oxford University Press, 2022. ISBN 978-0-19-884013-8.

R.N. Giere. *Explaining Science: A Cognitive Approach.* Science and Its Conceptual Foundations Series. University of Chicago Press, 2010. ISBN 978-0-226-29203-8.

Henrique Gomes. Gauging the boundary in field-space. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 67:89–110, 2019. ISSN 1355-2198. doi: 10.1016/j.shpsb.2019.04.002.

Hilary Greaves and David Wallace. Empirical Consequences of Symmetries. *The British Journal for the Philosophy of Science*, 65(1):59–89, 2014. ISSN 0007-0882. doi: 10.1093/bjps/axt005.

Richard Healey. Perfect Symmetries. *The British Journal for the Philosophy of Science*, 60(4): 697–720, 2009. ISSN 0007-0882. doi: 10.1093/bjps/axp033.

Jenann Ismael and Bas C. van Fraassen. Symmetry as a Guide to Superfluous Theoretical Structure. In Katherine Brading and Elena Castellani, editors, *Symmetries in Physics: Philosophical Reflections*, pages 371–392. Cambridge University Press, 2003. ISBN 978-0-521-82137-7.

Caspar Jacobs. Invariance or equivalence: A tale of two principles. *Synthese*, 199(3):9337–9357, December 2021. ISSN 1573-0964. doi: 10.1007/s11229-021-03205-5.

Caspar Jacobs. Are Models Our Tools Not Our Masters? *Synthese*, 202(4):1–21, 2023. doi: 10.1007/s11229-023-04352-7.

David Kaplan. Demonstratives: An Essay on the Semantics, Logic, Metaphysics and Epistemology of Demonstratives and Other Indexicals. In Joseph Almog, John Perry, and Howard Wettstein, editors, *Themes From Kaplan*, pages 481–563. Oxford University Press, 1989. ISBN 978-0-19-534534-6.

R. Keefe. *Theories of Vagueness.* Cambridge Studies in Philosophy. Cambridge University Press, 2000. ISBN 978-0-521-65067-0.

David Lewis. New Work for a Theory of Universals. *Australasian Journal of Philosophy*, 61(4): 343–377, 1983. doi: 10.1080/00048408312341131.

David Lewis. Putnam's Paradox. *Australasian Journal of Philosophy*, 62(3):221–236, 1984. doi: 10.1080/00048408412340013.

David K. Lewis. Languages and Language. In Keith Gunderson, editor, *Language, Mind, and Knowledge*, volume VII of Minnesota Studies in the Philosophy of Science, pages 3–35. University of Minnesota Press, 1975. ISBN 0-8166-5779-3.

Joanna Luc. Arguments from scientific practice in the debate about the physical equivalence of symmetry-related models. *Synthese*, 200(2):1–29, 2022. doi: 10.1007/s11229-022-03618-w.

Joanna Luc. Motivationalism vs. interpretationalism about symmetries: Some options overlooked in the debate about the relationship between symmetries and physical equivalence. *European Journal for Philosophy of Science*, 13, August 2023. doi: 10.1007/s13194-023-00539-4.

Sebastián Murgueitio Ramírez. *On the Empirical Significance of Physical Symmetries.* PhD thesis, University of Notre Dame, 2021.

Sebastián Murgueitio Ramírez. A Puzzle Concerning Local Symmetries and Their Empirical Significance. *British Journal for the Philosophy of Science*, 2023.

Sebastián Murgueitio Ramírez. On Symmetries and Springs. *Philosophy of Science*, pages 1–37, January 2024. ISSN 0031-8248, 1539-767X. doi: 10.1017/psa.2023.170.

Jill North. *Physics, Structure, and Reality.* Oxford University Press, Oxford, 2021. ISBN 0-19-264647-8.

P.J. Olver. *Applications of Lie Groups to Differential Equations.* Applications of Lie Groups to Differential Equations. Springer New York, 2000. ISBN 978-0-387-95000-6.

Angela Potochnik. *Idealization and the Aims of Science.* University of Chicago Press, September 2020. ISBN 978-0-226-75944-9.

G. E. Prince and C. J. Eliezer. On the Lie symmetries of the classical Kepler problem. *Journal of Physics A: Mathematical and General*, 14(3):587–596, March 1981. doi: 10.1088/0305-4470/14/3/009.

James Read and Thomas Møller-Nielsen. Redundant epistemic symmetries. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 70:88–97, 2020. ISSN 1355-2198. doi: 10.1016/j.shpsb.2020.03.002.

Bryan W. Roberts. Regarding 'Leibniz Equivalence'. *Foundations of Physics*, 50(4):250–269, 2020. doi: 10.1007/s10701-020-00325-9.

John Roberts. A Puzzle About Laws, Symmetries and Measurability. *British Journal for the Philosophy of Science*, 59(2):143–168, 2008. doi: 10.1093/bjps/axn009.

Jeff Speaks. Theories of Meaning (Stanford Encyclopedia of Philosophy). In Edward N. Zalta, editor, *Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2021 edition, 2010.

Nicholas J. Teh. Galileo's Gauge: Understanding the Empirical Significance of Gauge Symmetry. *Philosophy of Science*, 83(1):93–118, 2015. ISSN 0031-8248. doi: 10.1086/684196.

Trevor Teitel. What Theoretical Equivalence Could Not Be. *Philosophical Studies*, 178(12):4119–4149, 2021. doi: 10.1007/s11098-021-01639-8.

David Wallace. Isolated Systems and Their Symmetries, Part I: General Framework and Particle-Mechanics Examples. *Studies in History and Philosophy of Science Part A*, 92(C):239–248, 2022a. doi: 10.1016/j.shpsa.2022.01.015.

David Wallace. Observability, redundancy and modality for dynamical symmetry transformations. In James Read and Nicholas Teh, editors, *The Philosophy and Physics of Noether's Theorem*, pages 322–353. Cambridge University Press, Cambridge, 2022b. ISBN 978-1-108-66544-5.

James Owen Weatherall. Regarding the 'Hole Argument'. *British Journal for the Philosophy of Science*, 69(2):329–350, 2018. doi: 10.1093/bjps/axw012.

William J. Wolf and James Read. Respecting boundaries: Theoretical equivalence and structure beyond dynamics. *European Journal for Philosophy of Science*, 13(4):47, September 2023. ISSN 1879-4920. doi: 10.1007/s13194-023-00545-6.

C. E. Wulfman and B. G. Wybourne. The Lie group of Newton's and Lagrange's equations for the harmonic oscillator. *Journal of Physics A: Mathematical and General*, 9(4):507, April 1976. ISSN 0305-4470. doi: 10.1088/0305-4470/9/4/007.