Mechanistic Abstraction

Worth Boone

Gualtiero Piccinini
University of Missouri-St. Louis, piccininig@umsl.edu

Follow this and additional works at: https://irl.umsl.edu/philosophy-faculty

Part of the Philosophy of Science Commons

Recommended Citation
DOI: https://doi.org/https://doi.org/10.1086/687855
Available at: https://irl.umsl.edu/philosophy-faculty/8

This Article is brought to you for free and open access by the Philosophy at IRL @ UMSL. It has been accepted for inclusion in Philosophy Faculty Works by an authorized administrator of IRL @ UMSL. For more information, please contact marvinh@umsl.edu.
We provide an explicit taxonomy of legitimate kinds of abstraction within constitutive explanation. We argue that abstraction is an inherent aspect of adequate mechanistic explanation. Mechanistic explanations—even ideally complete ones—typically involve many kinds of abstraction and therefore do not require maximal detail. Some kinds of abstraction play the ontic role of identifying the specific complex components, subsets of causal powers, and organizational relations that produce a suitably general phenomenon. Therefore, abstract constitutive explanations are both legitimate and mechanistic.

1. Introduction. A constitutive explanation explains a phenomenon (capacity, function, activity, process) in terms of some suitably organized subphenomena (subcapacities, subfunctions, subactivities, subprocesses; e.g., Fodor 1968; Cummins 1983; Craver 2007; Bechtel and Richardson 2010). Some constitutive explanations are mechanistic: they attribute the explanatory subphenomena to structural components (as opposed to black boxes; see Piccinini and Craver 2011). Some constitutive explanations are more abstract than others. That is, some constitutive explanations include fewer (causally relevant) details than others about their explanandum, their explanantia, or both. Question: is abstraction within constitutive explanation compatible with mechanistic explanation, or does it lead to nonmechanistic forms of explanation?

To answer this question, let’s distinguish between two kinds of roles that abstraction could play within explanation: epistemic roles and ontic roles. This distinction maps onto epistemic and ontic conceptions (or modes) of explanation.
According to the epistemic conception, explanations are *representations* that meet certain epistemic requirements. Events are explained by subsuming them under appropriate representations. For example, we explain why someone developed lung cancer by pointing out that they smoked and that people who smoke have a certain probability of developing lung cancer. Thus, we explain the lung cancer by subsuming this particular instance under a statistical generalization. The epistemic conception easily accommodates abstract explanation. We can represent the world more or less abstractly, so we can give more or less abstract explanations.

According to the ontic conception, by contrast, the adequacy conditions for explanations are facts in the world. To explain a phenomenon is to identify the explanatory facts themselves. In our example, there are biological facts that explain why someone develops lung cancer from smoking (e.g., carcinogens accumulate and are metabolized into forms that covalently bond with DNA to form DNA adducts). These facts provide the adequacy conditions of our explanatory descriptions: there is still a sense in which representations explain, but only insofar as they denote the explanatory facts.

The ontic conception of mechanistic explanation has been interpreted as implying that the more (relevant) details an explanatory description includes, and hence the less abstract it is, the better it explains. After all, the world is maximally concrete. The more relevant details we include, the more completely we denote the relevant facts or events. Therefore, the less abstract a model is, the better it approximates an ideally complete mechanistic explanation. Call this the *requirement of maximal detail*.2

The requirement of maximal detail implies that abstract constitutive explanations—since they omit relevant details—are partial, inferior, or inadequate. Critics have balked at these implications and defended abstract explanations (Weiskopf 2011; Barberis 2013; Levy and Bechtel 2013; Barrett 2014; Chirimuuta 2014; Levy 2014). As these critics see things, proponents of the ontic conception are committed to the requirement of maximal detail and must therefore reject abstract constitutive explanations. These critics conclude that either explanation is solely epistemic (not ontic) or abstract explanations are not mechanistic precisely because they fail the requirement of maximal detail (or both).

This debate bears on at least five important questions. First, there is the proper characterization of constitutive explanation itself—whether mechanistic explanation requires maximal detail, whether abstract constitutive ex-

---

1. On causal relevance, see Craver (2007, chaps. 3 and 4).
2. Some statements by some mechanist philosophers (e.g., Craver 2007, 114; Kaplan 2011, 347) are sometimes interpreted as implying the requirement of maximal detail, although the same philosophers also defend the use of abstraction within explanation. For a parallel debate in neuroscience, see Markram (2006) and Eliasmith and Trujillo (2014).
Planations are legitimate, and whether abstract constitutive explanations are mechanistic or are, instead, distinct and autonomous from mechanistic explanations.

A second question is whether explanation is epistemic or ontic. If the requirement of maximal detail holds and abstract explanations are legitimate, this appears to challenge the ontic conception. Some critics respond by rejecting the ontic conception. Another response is to assign abstraction merely epistemic roles: abstraction is a feature of our limitations in representing the world (see Craver 2014). We defend a third alternative: abstraction plays ontic as well as epistemic roles.

A third related question is whether and when multiple realizability occurs. If abstract constitutive explanations are legitimate, abstract explanations may denote higher-level phenomena that are multiply realizable because they are invariant over changes in the details they abstract from. If such abstract explanations are illegitimate or incomplete at best, then multiple realizability may be ruled out. The status of multiple realizability has important consequences for the metaphysics of mind: if mental states are multiply realizable, mental states are realized by but not identical to brain states; if mental states are not multiply realizable, they may well be identical to brain states (Polger and Shapiro 2016).

A fourth related question is the proper characterization of computational explanation. Computational explanation is paradigmatically abstract. If abstraction leads to inadequate explanations, then computational explanations are inadequate. If abstraction does not lead to inadequate explanations but does lead outside mechanistic explanation, then computational explanation may be adequate but not mechanistic (see Shagrir 2010; Haimovici 2013). If abstraction in explanation is both legitimate and compatible with mechanistic explanation, however, computational explanation may well be both mechanistic and legitimate (see Piccinini 2015).

A fifth related question is whether higher-level explanations, which abstract away from relevant lower-level details, are eliminable, are reducible to, are irreducible to, or can be integrated with lower-level explanations. If higher-level explanations are distinct and autonomous from lower-level explanations, they are irreducible to lower-level explanations. If abstract explanations are illegitimate, they should be eliminated. If they are legitimate and mechanistic and there is no multiple realizability, they may be reducible to lower-level explanations. If they are legitimate and there is multiple realizability, they may be integrated with lower-level explanations by forming multi-level mechanistic explanations.

In what follows, we address these questions by providing an explicit taxonomy of legitimate kinds of abstraction within constitutive explanation. We reject the requirement of maximal detail and argue that abstraction is an inherent aspect of adequate mechanistic explanation. Mechanistic explana-
tions—even ideally complete ones—typically involve many kinds of abstraction and do not require maximal detail. Some kinds of abstraction play the ontic role of identifying the specific complex components, subsets of causal powers, and organizational relations that produce a suitably general phenomenon. Therefore, abstract constitutive explanations are both legitimate and mechanistic.

Before we proceed, it will be helpful to lay out our basic assumptions. Complex mechanisms come in levels of functional organization. Mechanisms (level 0) compose larger mechanisms (level +1) and are composed by components, which are in turn (smaller) mechanisms (level −1). Mechanisms have properties, which are sets of causal powers. The properties of mechanisms (level 0), when such mechanisms and their properties are suitably organized, compose the properties of the larger mechanism that contains them (level +1). By the same token, the properties of a mechanism (level 0) are composed by the properties of its suitably organized components (level −1). The relation between the properties of a mechanism (level 0) and the properties of its organized components (level −1) is the proper subset relation: the causal powers possessed by a mechanism taken as a whole are a proper subset of the causal powers possessed by its organized components (Piccinini and Maley 2014).

At each level of organization, a mechanistic model articulates how the relevant properties of the relevant components, suitably organized, produce the phenomenon (see Craver 2007; Bechtel and Richardson 2010; Glennan, forthcoming). Ontically, a phenomenon is mechanistically explained by the mechanism that produces it and its relevant properties. An ideally complete mechanistic model is an articulation of the mechanism that produces a phenomenon and its relevant properties at all levels of organization, including an articulation of the compositional relations between the whole mechanism, its components and their properties and organization, their subcomponents and properties and organization, and so on (see Craver 2006, 360).

In section 2, we list several distinct epistemic roles that abstraction plays within the mechanistic framework. These epistemic roles are relatively uncontroversial and consistent with a range of views of the nature of mechanistic explanation. By contrast, in section 3 we outline several distinct ontic roles that abstraction plays within the mechanistic framework. We argue that even ideally complete mechanistic models involve abstraction.

2. Epistemic Roles of Mechanistic Abstraction. Mechanistic abstraction plays a number of epistemic roles. These roles are dictated by our epistemic interests and limitations rather than any ontic norms of mechanistic explanation. Such abstract explanations are thus incomplete in the relevant sense but may nonetheless be indispensable to science. Thus, incomplete abstract explanations may be explanatorily adequate in many contexts. The adequacy of such explanations may shift, however, as scientific knowledge and methods evolve.
Two types of abstraction play epistemic roles: abstraction of one or more levels of organization among others, and abstraction of one or more aspects of one level among others. These types of abstraction may be necessary because we are interested in some specific aspect of the phenomenon at the expense of others, because including more details makes the explanation intractable, or simply because we are ignorant of the details we omit.

Epistemically motivated abstractions are often mixed with simplifications and idealizations. Simplifications and idealizations introduce features into a description or model that distort or misrepresent the target system. Although simplification and idealization are not the main topic of this paper, it is important to realize that epistemically motivated abstractions, simplifications, and idealizations are part and parcel of typical mechanistic models (see Milkowski 2016).

One obvious role for abstraction in mechanistic explanation is that many details of a system may as yet be unknown. In such cases, mechanistic explanations take the form of mechanism sketches. In mechanism sketches, black boxes or filler terms may be included to highlight the incompleteness of a mechanistic model and motivate future research.

Even when certain details are known, there are other nontrivial epistemic reasons to omit them from a mechanistic model. In the first place, mechanistic explanations are typically offered at one or a few levels of organization among others. Including all levels of a mechanism in an explanation may be unfeasible for several reasons: not enough is known about some levels (as we just said), too much information would be needed to describe all the levels and it would be unfeasible to include it all, or much of the additional information is irrelevant to the levels of interest. Even when multiple levels of mechanistic organization are included in a model, typically the included levels remain only a subset of all the levels of organization of the entire multilevel mechanism. Other levels are omitted. Thus, multilevel mechanistic models—the ones actually developed in practice, as opposed to ideally complete ones—typically employ this form of abstraction.

A second epistemic role for abstraction comes from considerations of solubility and tractability in mathematical and computational models. Mechanistic explanations are often presented in the form of (interpreted) mathematical or computational models—for example, the Hodgkin–Huxley model of the action potential. Typically, such models become intractable or difficult to manipulate if they include too many details about their target systems. Instead, mathematical and computational models are typically constructed not only by abstracting away from many details of the target system but also by replacing those details with simplifications and idealizations that distort or misrepresent the target system. Nevertheless, these models retain crucial epistemic roles (e.g., for prediction and discovery) and often retain explanatory value to the extent that they denote some components, causal powers, or organizational relations of target systems that are relevant to the explanandum phenomenon.
A third epistemic role for abstraction is that we are often interested in one component or property of a mechanism at the expense of other components or properties. The result is another type of mechanism sketch, or partial (elliptical) mechanistic explanation. Consider what it takes to explain why, under special conditions, a mechanism functions differently than it normally does. Explaining a deviation from normal functioning typically requires simply demonstrating some aberrant feature of the mechanism, while omitting the details relevant to normal function (van Eck and Weber 2014).

Interim conclusion: our epistemic interests and limitations often demand abstractions, namely, abstraction from unknown details, abstraction from some levels of a mechanism in favor of one or more levels, abstractions in the service of mathematical and computational tractability, and abstraction from some aspects of one level in favor of other aspects of that level.

3. Ontic Roles of Mechanistic Abstraction. The epistemic roles of abstraction are consistent with a range of views on the nature of mechanistic explanation, including the requirement of maximal detail. Contrary to the requirement of maximal detail, however, even ideally complete mechanistic models require abstraction.

Within an ideally complete mechanistic explanation, abstraction serves two interrelated roles: (1) to explain a phenomenon at a particular degree of generality, and (2) to identify all and only the components, properties, and organizational relations that constitute the mechanism at different levels. To fulfill these roles, two types of abstraction must be performed: (i) abstraction to sufficiently general types of components, properties, and organizational relations; and (ii) abstraction from lower levels of organization to higher levels of organization.

To a first approximation, we construe types as classes of components, property instances, or relation instances whose members resemble one another. We leave open whether property types are tropes that resemble one another, universals, or classes of particulars. As shorthand, we will say that a type of property is a set of causal powers, a type of component is a class of components that share relevant properties, a type of organizational relation is an organizational relation between types of components and their properties, and a regularity is a manifestation of a type of mechanism (component types, property types, and organizational relation types).

At face value, the requirement of maximal detail entails that all explanations are explanations of token events; obtaining any level of generality requires at least some omission of the particular details of token events (see Ross 2015). But explananda can themselves be regularities. For instance, the link between smoking and lung cancer (explanandum) is itself a regularity in the world, which we explain by situating it within other causal regularities—namely, the processes by which carcinogens are metabolized into forms that covalently bond with DNA to form DNA adducts. On the ontic
conception, we thus explain such regularities in the same way we explain token events—both are perfectly legitimate explananda; they simply differ in their scope or generality.3

Explanandum phenomena can be more specific or more general in a variety of ways. We may wish to explain how a particular rat navigates a particular maze at a particular time (a particular event), how a particular rat navigates in general, rat navigation more generally, rodent navigation, mammalian navigation, and so forth. Neural recordings from a particular rat engaged in a particular behavior at a particular time may be part of the evidence for discovering explanations of all of the above explananda, typically in combination with other sources of evidence. Depending on which phenomenon we aim to explain and how general it is, we need to invoke more specific or more general types of components, properties, and organizational relations.

Generalizing beyond token events requires isolating target phenomena that are stable across variations in those events. For example, memory researchers may implant electrodes in the hippocampi of specific rats and gather recordings while those rats navigate a specific maze at a specific time. But those researchers are typically not interested in merely explaining how the particular rats in the test population navigated those particular mazes at those particular times; rather, they more often seek to generalize their results to conclusions about the functions of rat hippocampi in general, or rodent hippocampi more generally, or even mammalian hippocampi.

If we are explaining a particular navigation event, we are free to appeal to idiosyncratic features of that event. If, by contrast, we are explaining rat navigation in general, we have to appeal to features shared by the relevant mechanisms present in normal rats. In such cases, we target a mechanism type, as opposed to a mechanism token (see Machamer, Darden, and Craver 2000, 15). This requires abstracting away from the idiosyncratic details of particular rats and particular navigation events and specifying the relevant subset of causal powers, common to all tokens of the relevant types of component, which, when appropriately organized, produce the phenomenon at the specified degree of generality—here, rat navigation.

Explaining still more general phenomena like mammalian navigation requires abstracting away from the specifics of rat physiology to find relevant similarities across the class of organisms that feature in the explanandum. This may require identifying an even more restricted subset of causal powers common to the more general type of mechanism that is shared by dif-

3. Abstraction and generality are distinct notions. Abstraction does not always increase generality, as can be seen in the examples discussed in sec. 2. But generality requires abstraction: for any explanation to have scope beyond token events, some idiosyncratic details of those token events must be omitted.
ferent sorts of mammals. These abstractions succeed to the extent that they track stable features of those progressively more general types of system. Those stable features delimit different, and progressively more general, types of mechanisms. Identifying and explaining those different mechanism types requires omitting the idiosyncratic details of less abstract types of mechanism in order to reach a description that is general enough to denote the relevant features that the less abstract types have in common.

Beyond a certain degree of generality, explananda may no longer share enough relevant features to constitute a unified phenomenon that can properly be a target of explanation. Alternatively, the systems that are responsible for a phenomenon may no longer share relevant features, or the relevant features that are shared may no longer productively explain the supposed target phenomenon. In other words, there may be cases in which no relevant subset of causal powers can be identified; in such cases there is no relevant mechanism type to be identified and no unified phenomenon to be explained.

As stated above, an ideally complete mechanistic explanation of a phenomenon includes all and only the details that produce the explanandum at all levels of organization. This requires both omitting details that are irrelevant to producing the phenomenon and including all the relevant features at all relevant levels of organization. Moving from one level to a higher level requires another kind of abstraction. The details of lower-level mechanism types, which implement the level under consideration, are omitted. As a result, the causal powers being described at the higher level are only a subset of the causal powers of the lower levels.

Consider the rat hippocampus and its components. The causal powers of the rat hippocampus taken as a whole are only a subset of the causal powers of the neurons and other components that make up the rat hippocampus, even when the rat hippocampus’s neurons and other components are taken to be organized into the rat hippocampus itself. For instance, the rat hippocampus’s neurons have the causal power to send neurotransmitters to one another under various conditions, whereas the rat hippocampus taken as a whole lacks that power. Instead, the hippocampus as a whole has the causal power of, for instance, tracking the location of a rat in its environment. This function of the hippocampus may well be stable over massive amounts of variation in the properties of the individual neurons that compose the system—and so in a relevant sense does not constitutively depend on those individual details of the system.

When we go up a level in our mechanistic explanations (e.g., from neurons to hippocampi), we abstract away from some of the causal powers of the components in order to single out the specific causal powers that produce the phenomenon at the higher level. For instance, we single out the causal powers that are specific to hippocampi as wholes, as opposed to their
organized components. In other words, we identify the aspects of the world that explain the phenomenon at the higher level.

This kind of ubiquitous abstraction tracks one form of multiple realizability between levels of mechanistic organization, because the same set of causal powers, which operates at one level, may be embedded in different supersets at lower levels (Piccinini and Maley 2014). For example, the causal power that all functioning corkscrews share (lifting corks out of bottles) may be a subset of different supersets (lifting corks out of bottles by screwing into the cork and directly pulling it out, by leveraging it out using an arm placed on one side of the bottle, by indirectly pulling it out with two levers attached to a rack and pinion, and so on) depending on which type of lower-level mechanism is being used to generate the original power (see Shapiro 2004).

A related but distinct ontic role for abstraction in mechanistic explanation consists in isolating features that are shared by mechanisms that occur within radically different systems and may even occur at different levels of organization. This form of analysis yields a kind of mechanism schema (Machamer et al. 2000). Mechanism schemata involve deliberate omissions of details, capturing a mechanistic structure that many different systems have in common. As above, a schema denotes a specific set of causal powers that is found in many different systems, where it may be embedded in different supersets—thus providing a distinct form of multiple realizability.

Graph-theoretic models provide a common way of representing mechanism schemata in neuroscience (see Levy and Bechtel 2013). Such models are often used to represent types of neural circuits that crop up in different parts of the nervous system. The particular microstructural details—the types of neurons, neurotransmitters, receptors, and so on—are omitted in order to isolate organizational similarities that are invariant over these details.

For instance, lateral inhibition can be represented by a simple set of input nodes with inhibitory projections toward the neighbors of their corresponding output nodes (fig. 1). When many input nodes have overlapping receptive fields, those that receive the greatest stimulation inhibit the neighbors of their corresponding output nodes, creating a “winner-take-all” processing stream. This form of circuitry is found in many different neural regions and is particularly ubiquitous in peripheral sensory processing—for example, in the retina and in the somatosensory periphery—because it allows greater precision in transduction of sensory signals.

Such patterns in neural circuitry offer a distinct but similar form of generality to that found in moving from features of particular rat hippocampi to rat hippocampi in general, or rodent or mammalian hippocampi. In the case of hippocampi, identifying the relevant mechanism type that explains the phenomenon of interest (rat navigation, rodent navigation, or mammalian navigation) relies heavily on the structural similarity (i.e., the shared components) of those different systems, whereas in the case of lateral inhibition,
identifying the relevant mechanism type relies more on organizational similarity of the different systems.

Interim conclusion: abstraction plays two ontic roles within mechanistic explanation, namely, to explain a phenomenon at a suitable degree of generality and to identify all and only the features of the mechanism at all levels of organization. Two kinds of abstraction are needed to fulfill these roles: abstraction to general types of component, causal power, and organizational structure, and abstraction from lower to higher levels of organization.

4. Conclusion. Contrary to what is often asserted or implied, mechanistic explanation does not require maximal detail. Multilevel mechanistic explanation—even ideally complete mechanistic explanation—mandates several legitimate kinds of abstraction. Of course, actual models are not ideally complete explanations: they contain abstractions as well as simplifications and idealizations that play epistemic roles in addition to ontic roles. Such models still provide (partial) mechanistic explanations.

Sometimes mechanistic explanations are abstract schemata that cover many specific instantiations, sometimes they are abstract schemata that cover many kinds of realizing mechanisms, sometimes they are abstract schemata that cover similar mechanistic structures in different kinds of systems and possibly at different levels of organization, sometimes they are either abstract sketches or schemata that involve deliberate omissions of known details, and sometimes they are sketches that omit unknown details. These types of abstraction are essential to mechanistic explanation both in practice and in principle.

This conclusion has five important corollaries.
First, abstract constitutive explanations are legitimate and mechanistic; therefore, contrary to common assertions (e.g., Shapiro, forthcoming), abstract constitutive explanations are neither distinct nor autonomous from mechanistic explanations. All instances of mechanistic explanation involve mutual constraints between higher-level and lower-level analyses. Omission of details presupposes some sufficient understanding of those details to know what is and is not relevant. In other words, analyses that deliberately omit known details are obviously not autonomous from those details. It is only in cases in which relevant details are unknown that those details must be investigated and then either added or deliberately omitted in order to explain how the system actually performs a particular operation. Putatively nonmechanistic constitutive explanations are either sketches of mechanisms or mechanism schemata that describe aspects of one or more mechanistic levels. As one of us has argued elsewhere, mechanism sketches can be fleshed out into one level of a mechanism and integrated into multilevel mechanistic explanations (Piccinini and Craver 2011).

Second, some abstract explanations denote an important aspect of the objective causal structure of the world. Specifically, they denote a system’s complex components, subsets of causal powers, and organizational relations that are operative (and thus explain a phenomenon) at one or more relevant levels of organization and produce a phenomenon with a suitable degree of generality. Therefore, some forms of mechanistic abstraction are part and parcel of any adequate ontic conception of explanation.

Third, multiple realizability is perfectly compatible with multilevel mechanistic explanation precisely because mechanistic explanation involves abstraction to a specific subset of causal powers that is shared by many different systems, which subset can be embedded in different supersets in different systems (Piccinini and Maley 2014).

Fourth, computational explanation involves abstraction while nonetheless remaining mechanistic (Piccinini 2015).

Fifth, abstract constitutive explanations at higher levels can be integrated with less abstract explanations at lower levels to form multilevel mechanistic explanations; the adoption of this integrationist explanatory strategy is an important aspect of the cognitive neuroscience revolution (Boone and Piccinini 2016).

REFERENCES


