An Abductive Theory of Constitution*

Michael Baumgartner† and Lorenzo Casini‡

Abstract

The first part of this paper finds Craver's (2007) *mutual manipulability* theory (MM) of constitution inadequate, as it definitionally ties constitution to the feasibility of idealized experiments, which, however, are unrealizable in principle. As an alternative, the second part develops an abductive theory of constitution (NDC), which exploits the fact that phenomena and their constituents are unbreakably coupled via common causes. The best explanation for this common-cause coupling is the existence of an additional dependence relation, *viz.* constitution. Apart from adequately capturing the essential characteristics of constitution missed by MM, NDC has important ramifications for constitutional discovery—most notably, that there is no *experimentum crucis* for constitution, not even under ideal discovery circumstances.

1 Introduction

According to mechanistic theories of explanation, the upper (macro) level behavior Ψ of a system S is explained by carving out the lower (micro) level mechanism *constituting* that behavior (Glennan 1996; Machamer et al. 2000; Craver 2007). Hence, a theory of mechanistic explanation presupposes a theory of constitution providing criteria that identify those of S's spatiotemporal parts whose activities are constitutively relevant to S's Ψ-ing. The most popular theory of constitution, due to Craver (2007), purports to furnish such criteria by drawing on conceptual and methodological resources that have proven valuable in analyzing and discovering causation—notwithstanding the fact that constitution and causation are very different relations (Craver and Bechtel 2007).

Since the time of Mill (1843), one of the dominant approaches to uncovering causation consists in intervening on causes (in controlled environments) to test whether they make a difference to their purported effects. As is well-known, Woodward (2003) has built his influential interventionist theory of causation on the fundament of this experimental protocol. While causation is a unidirectional difference-making relation—i.e. causes change their effects, but not vice versa—and holds among mereologically independent entities, Craver (2007) argues that constitution is a bidirectional or mutual difference-making relation among wholes and some of their spatiotemporal parts. Correspondingly, he proposes a theory of constitution that adds a parthood and a mutuality tweak to Woodward's interventionist theory of causation. Subject to Craver's (2007, 153) *mutual manipulability* theory (MM), the behavior Φ of a spatiotemporal part X of S constitutes S's Ψ-ing

*Forthcoming in *Philosophy of Science*.
†Dept. of Philosophy, University of Geneva. Email: michael.baumgartner@unige.ch
‡Dept. of Philosophy, University of Geneva. Email: lorenzo.casini@unige.ch

iff it is possible to (ideally) intervene (from the bottom up) on X's Φ-ing such that S's Ψ-ing changes, and (from the top down) on S's Ψ-ing such that X's Φ-ing changes. Craver claims that MM provides an adequate analysis of constitution and that it furnishes a solid foundation for evidence-based constitutional discovery.

The first part of this paper takes issue with both of these claims (§2). MM is far from providing an adequate analysis of constitution, as it is in fact inapplicable to the very structures it is designed to account for. MM definitionally ties constitution to the feasibility of idealized experiments. Such experiments are, however, unrealizable *in principle*. Macro-level phenomena and their constituents are so tightly intertwined that they can only be manipulated via common causes (cf. Baumgartner and Gebharter 2015). Furthermore, less rigorous but feasible experimental setups inevitably generate confounded data that systematically underdetermine the inference to constitutive relations. Hence, MM cannot possibly ground a viable methodology for constitutional discovery.

Since constitution is a non-causal form of dependence—as commonly assumed in mechanistic theorizing—one cannot simply tweak a successful account of causation to obtain a successful account of constitution. Rather, constitution must be defined within a theoretical framework that reflects its distinctly non-causal nature. Furthermore, the inference to constitution can neither in theory nor in practice proceed along the lines of the inference to causation. The main reason is that, while there exist ideal experimental designs allowing for the generation of unconfounded data that conclusively establish the existence of causal relations, no such experimental designs exist for the inference to constitution. Even data generated under ideal discovery circumstances can always equivalently be accounted for in terms of a model that features constitutive dependencies and a model without any such dependencies. Hence, the inference to constitution is inherently underdetermined by experimental evidence (§3).

As an alternative to MM, the second part of the paper then develops an abductive theory of constitution, which exploits the fact that phenomena and their constituents are unbreakably coupled via common causes (§4). The existence of an additional dependence relation, *viz.* constitution, is the *best explanation* for this unbreakable common-cause coupling. Hence, pace Craver, the defining feature of constitution is not the *possibility* of top-down and bottom-up interventions on a mechanism and the *existence* of corresponding (mutual) difference-making scenarios, but the *impossibility* of such interventions and the *nonexistence* of such difference-making scenarios.

Our abductive theory has important ramifications for the inference to constitution that any viable method of constitutional discovery has to take into account (§5). In particular, to establish constitutive dependencies, it does not suffice to wiggle the macro level of a mechanistic system such that the micro level changes and vice versa—as stipulated by MM. Rather, an extended series of experiments is needed that explore the whole space of possible ways of *breaking* the coupling of macro and micro levels. Only if all of these tests are *unsuccessful*, an inference to constitution is warranted. Moreover, in light of its inherent empirical underdetermination, such an inference is ultimately grounded in pragmatic considerations concerning explanatory power, and is not forced upon the modeler by the evidence.

2 Inadequacy of MM

Before reviewing Craver's MM, we must render transparent two crucial background assumptions of our argument and introduce our notation.

According to the first assumption, which is compellingly substantiated by Craver and Bechtel (2007), constitution must be sharply distinguished from causation. Causation holds among mereologically independent entities such that causes temporally precede their effects, and it is a unidirectional form of dependence in the sense that effects depend on their causes but not vice versa. By contrast, constitution holds among wholes and their parts,[1] that is, among spatiotemporally overlapping entities, and it is bidirectional in the sense that parts depend on the wholes and vice versa. Although some authors are skeptical of the distinction between causation and constitution (e.g. Ross and Ladyman 2010; Leuridan 2012), the distinction is standardly accepted by representatives of theories of mechanistic explanation. As this is the theoretical context of our paper, we shall subsequently assume that constitution is a distinctly non-causal form of dependence.

The second assumption likewise belongs to the canon of mechanistic theorizing; it states that the relation between a mechanism's upper and lower level is to be analyzed in terms of non-reductive supervenience (Glennan 1996, 61-62; Eronen 2011, ch. 11). More specifically, relative to a given a mechanistic organization of the constituents, phenomena supervene on their constituents, meaning that every change in a phenomenon is necessarily accompanied by a change in its constituents (Craver 2007, 153). Moreover, phenomena are not reducible—in particular, not identical—to their constituents.

Phenomena and their constituents are types of behavior exhibited by specific entities on macro and micro levels, respectively. To represent such behaviors, we adopt the following notational conventions. Entities on macro levels are represented by individual constants S, S_1, S_2, etc., and macro behaviors by variables Ψ, Ψ_1, Ψ_2, etc. Micro-level entities are symbolized by X, X_1, X_2, etc., and micro behaviors by variables Φ, Φ_1, Φ_2, etc. (Craver 2007, 153-60). Moreover, we refer to the behaviors of specific entities—S's Ψ-ing or X's Φ-ing—by means of *specific* variables—$\Psi(S)$ and $\Phi(X)$. Contrary to generic variables, which represent behaviors as exhibited by any entities, specific variables represent behaviors of specific entities (Spohn 2006). For instance, $\Phi(X) = \phi_i$ means that entity X exhibits the behavior ϕ_i. As we—like Craver—are only concerned with behaviors

[1] Although the relation of parthood raises numerous metaphysical questions, authors working on mechanistic explanation typically sidestep the topic (Harbecke 2010 is a commendable exception), as they employ a metaphysically "thin" notion of parthood. This notion—which we shall employ, too—is the notion of containment in a phenomenon's spatiotemporal extension, a notion formally defined by the axioms of *Ground Mereology* (Casati and Varzi 1999, ch. 3).

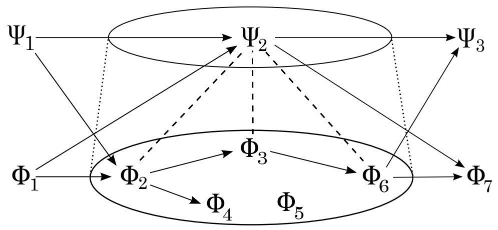

Figure 1: An example of a phenomenon Ψ_2 with three constituents, Φ_2, Φ_3 and Φ_6, and two non-constituting parts, Φ_4 and Φ_5. Dashed lines represent constitution, directed edges symbolize causation, and the dotted lines stand for spatiotemporal overlap.

of specific entities, we can dispense with generic behavioral variables altogether, which, in turn, allows us to treat the specificity of specific variables implicitly, and to abbreviate our notation by simply writing Ψ for $\Psi(S)$ and Φ for $\Phi(X)$.

Phenomena typically have a multitude of spatiotemporal parts, only a proper subset of which are constitutively relevant to them. Likewise, they tend to be involved in manifold causal interactions with non-constituents. Figure 1 provides an illustration: the phenomenon Ψ_2 in the macro-level ellipse has five parts in the micro-level ellipse; only three of them, *viz.* Φ_2, Φ_3, and Φ_6, are constituents of Ψ_2; both the phenomenon and its constituents are involved in numerous inter- and intra-level causal relationships. That abstract structure can, for instance, be interpreted in terms of the mechanism underlying a cruising car, such that Ψ_2 represents the phenomenon of the cruise, Φ_2 the running engine, Φ_3 the transmission of momentum to the axle, and Φ_6 the turning wheels. Moreover, the car has two parts that are not constitutively relevant to its movement, *viz.* the air conditioning, Φ_4, which is also operated by the engine, and the ashtray, Φ_5, which is causally detached from the causal mechanism constituting the car's cruise. The purpose of a theory of constitution is to provide criteria that distinguish constituents from non-constituting parts and from their causes and effects.

According to Craver's (2007) MM, constitution is a difference-making relation that is adequately analyzed by suitably supplementing the resources of the currently most popular difference-making theory of causation: Woodward's (2003) interventionism. In a nutshell, interventionism stipulates that a variable $\mathcal{X}$ is a cause of another variable $\mathcal{Y}$ iff it is possible to (ideally) intervene on $\mathcal{X}$ in such a way that $\mathcal{Y}$ changes, when all off-path causes of $\mathcal{Y}$ are fixed (cf. Woodward 2003, 59).[2] As constitution, contrary to causation, is a bidirectional dependence relation among parts and wholes, unidirectional manipulability as in interventionism does not suffice to establish constitutive relevance. Therefore, Craver adds a parthood and a mutuality constraint: constituents are spatiotemporal parts of phenomena and both are mutual difference-makers of each other. More explicitly:

> [...] to establish that X's Φ-ing is relevant to S's Ψ-ing [where X is a spatiotemporal part of S] it is sufficient that one be able to manipulate S's Ψ-ing

[2]We slightly adapt Woodward's notation to avoid confusion with our own notation.

> by intervening to change X's Φ-ing (by stimulating or inhibiting) and that one be able to manipulate X's Φ-ing by manipulating S's Ψ-ing. To establish that a component is irrelevant, it is sufficient to show that one cannot manipulate S's Ψ-ing by intervening to change X's Φ-ing and that one cannot manipulate X's Φ-ing by manipulating S's Ψ-ing. (Craver 2007, 159)

MM provides a sufficient condition for constitutive relevance and a sufficient condition for constitutive irrelevance, which jointly amount to a sufficient and necessary condition for constitutive relevance—that is, to a complete definition of constitution.[3] The core notion in its definiens is the modal notion of *manipulability*, which Craver (2007, §4.8.3) cashes out in terms of the existence of a possible ideal intervention as defined by Woodward (2003, 98). An *ideal intervention* on Ψ with respect to Φ is a variable $\mathcal{I}_\Psi$ taking one of its values, $\mathcal{I}_\Psi = i_n$, and thereby *surgically* fixing the value of Ψ without having an impact on Φ that is not mediated via Ψ and without being correlated with any other (off-path) causes of Φ. In sum, MM amounts to the following (where the specificity of the variables is made explicit):

(MM) $\Phi(X)$ is constitutively relevant to $\Psi(S)$ iff (i) X (resp. X's Φ-ing) is a spatiotemporal part of S (resp. S's Ψ-ing); (ii) there exists a possible ideal intervention $\mathcal{I}_\Phi = i_m$ on $\Phi(X)$ w.r.t. $\Psi(S)$ that is associated with a change in $\Psi(S)$; (iii) there exists a possible ideal intervention $\mathcal{I}_\Psi = i_n$ on $\Psi(S)$ w.r.t. $\Phi(X)$ that is associated with a change in $\Phi(X)$.

One of the main selling points of MM is that it is directly operationalizable experimentally. The most straightforward way of establishing the *possibility* of mutually intervening on Φ and Ψ is to furnish *actual* bottom-up and top-down interventions of types $\mathcal{I}_\Phi = i_m$ and $\mathcal{I}_\Psi = i_n$. If, and only if, such experiments reveal mutual difference-making among wholes and parts, the latter are proven to be constituents of the former. To use Figure 1 as an illustration, Φ_2 (the car's engine) is conclusively shown to be a constituent of Ψ_2 (the car's movement) by performing one intervention on Φ_2 (e.g., taking one's foot off the accelerator) that is associated with a change in Ψ_2 (the car's deceleration) as well as one intervention on Ψ_2 (e.g., adding external friction) that is associated with a change in Φ_2 (the engine working harder). By contrast, if it is established (inductively or otherwise) that there are no such mutual difference-making interventions on parts Φ_4 (the air conditioning) and Φ_5 (the ashtray) or on causes (Φ_1, Ψ_1) and effects (Φ_7, Ψ_3) of the mechanism, MM identifies these variables as non-constituents of Ψ_2.

Although MM has considerable intuitive appeal, the remainder of this section will show that MM does not amount to an adequate theory of constitution—not because it merely fails in some intricate cases but because it fails in *all* cases. In short, the reason is that MM is unsatisfiable in principle, for there cannot exist ideal interventions on upper and lower levels of a mechanism that are associated with changes on the other level.

[3] Some authors (e.g. Couch 2011, 382; Kaplan 2012, 560) misread MM as only providing a sufficient condition for constitutive relevance. Textual evidence clearly contradicts that assessment.

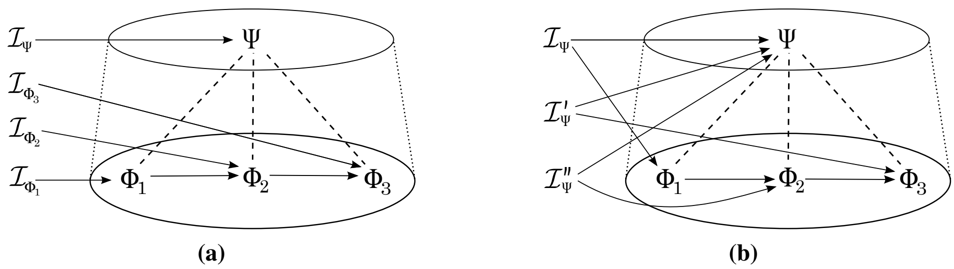

Figure 2: Model (a) depicts the impossible surgical interventions required by MM; model (b), by contrast, features the possible fat-handed interventions.

To see this, consider the simple mechanism in Figure 2a: Ψ has the three constituents Φ_1, Φ_2, Φ_3, which we assemble in what we subsequently call a *constituting set* $\mathbf{\Phi} = \{\Phi_1, \Phi_2, \Phi_3\}$. MM entails that, for the elements of $\mathbf{\Phi}$ to be constituents of Ψ, it is necessary that there exist intervention variables $\mathcal{I}_\Psi, \mathcal{I}_{\Phi_1}, \mathcal{I}_{\Phi_2}, \mathcal{I}_{\Phi_3}$ as depicted in Figure 2a that can induce changes on the system's other level. Thus, assume (for reductio) that $\mathcal{I}_\Psi$ is an intervention variable for Ψ w.r.t. Φ_1 such that changing Ψ via $\mathcal{I}_\Psi$ is associated with a change in Φ_1 (when all off-path causes of Φ_1 are held fixed). In that case, Woodward's interventionism, which constitutes the theoretical background of MM, entails that $\mathcal{I}_\Psi$ is a cause of both Ψ and Φ_1. This can be realized in one of two ways: either $\mathcal{I}_\Psi$ causes Ψ and Φ_1 along *one* causal path, say, $\mathcal{I}_\Psi \longrightarrow \Psi \longrightarrow \Phi_1$, or along *two* paths, $\Psi \longleftarrow \mathcal{I}_\Psi \longrightarrow \Phi_1$. The former option is excluded since the instances of Ψ and Φ_1 spatiotemporally overlap (i.e. the corresponding values of the variables represent spatiotemporally overlapping behaviors), which entails that their relationship is non-causal. Hence, $\mathcal{I}_\Psi$ causes Ψ and Φ_1 along two different paths, meaning that $\mathcal{I}_\Psi$ is a common cause of Ψ and Φ_1.[4] That, in turn, entails that $\mathcal{I}_\Psi$ does not surgically cause Ψ and is, therefore, not an intervention variable for Ψ w.r.t. Φ_1, which contradicts our initial assumption, thereby reducing it to absurdity.[5]

[4]Note that by a *common cause* we simply mean a cause with two (or more) parallel effects—independently of whether there is an additional dependence relation between these effects. A common cause is standardly represented by two (or more) exiting arrows (but see Woodward 2015, 331, for an alternative representation).

[5]An anonymous reviewer suggested that the consequence that $\mathcal{I}_\Psi$ is a common cause of Ψ and Φ_1—and thus a non-surgical cause of Ψ—could be avoided on the basis of a theory of causation requiring that causally analyzed variable sets exclusively contain *causally distinct* variables, which Ψ and Φ_1, arguably, are not. Clearly though, interventionism as developed in Woodward (2003, 2015) and as implemented by Craver (2007) does not impose such a restriction—and with good reasons, in our view. Whether $\mathcal{I}_\Psi$ is an intervention variable for Ψ w.r.t. Φ_1 does not depend on whether $\mathcal{I}_\Psi$ surgically causes Ψ relative to some suitably chosen variable set, but on whether *what is represented by* $\mathcal{I}_\Psi$ surgically causes Ψ *in the world*—which is not the case under the non-reductive physicalist assumption that Ψ and Φ_1 are non-identical. Moreover, a theory of causation explicitly imposing that variables be causally distinct would not be applicable to variable sets relevant for mechanistic explanations, and worse even, it would run an obvious circularity risk.

This result can easily be generalized. Due to the non-causal nature of the relationship between phenomena and their parts, every cause of a mechanism's upper or lower level that is associated with a change on the other level is a common cause of the corresponding occurrences on the two levels. In other words, whatever makes a difference on both levels of a mechanism does so along different causal paths and, hence, is not a surgical intervention.

Furthermore, as the constituents $\mathbf{\Phi}$ of Ψ realize Ψ on the micro level, they form the supervenience base of Ψ in the mechanistic context of Figure 2a, which entails that every change in Ψ (occurring in some spatiotemporal region) is *necessarily* accompanied by a change in $\mathbf{\Phi}$ (i.e. in some Φ_i in $\mathbf{\Phi}$, occurring in the same spatiotemporal region). Every cause inducing a change in Ψ necessarily also brings about a change in at least one Φ_i in $\mathbf{\Phi}$ and is, thus, a common cause of Ψ and Φ_i. More concretely, every cause of Ψ has the structural properties of either $\mathcal{I}_\Psi$, $\mathcal{I}'_\Psi$, or $\mathcal{I}''_\Psi$ in Figure 2b. And more generally, every cause of a mechanism's macro level is necessarily a common cause of the phenomenon and at least one of its constituents. Phenomena and their constituents can only be manipulated with a fat hand (cf. Baumgartner and Gebharter 2015).[6] Overall, the types of interventions required by MM cannot possibly exist for any mechanistic system. MM is hence unsatisfiable, which means that constitutive relations as defined by MM are inexistent, which again entails that friends of mechanistic explanations who rely on MM chase a chimera. This result reduces MM to absurdity.

3 Underdetermination of constitutional inference

The source of this fatal deficiency of MM is easily pinpointed: MM definitionally ties the notion of constitution to the possibility of surgical top-down and bottom-up interventions that target *one* level of a mechanism and thereby change the other level, where in fact it is only possible to induce changes on upper and lower levels of a mechanism by fat-handedly targeting *both* levels on separate causal paths. Contrapositively put, whenever surgical interventions that target a first variable and induce changes in a second one are possible, these variables are not linked in terms of constitution but in terms of causation—as is duly entailed by the interventionist theory of causation (Woodward 2003). An obvious conclusion to draw is that constitution should not be analyzed in terms of surgical (or ideal) mutual manipulability. If the mutual manipulability idea is to get off the ground at all, it must be cashed out in terms of *non-surgical* interventions of some sort.

Indeed, prompted by problems of the original version of interventionism with macro-to-micro causation, Woodward (2015) has recently offered a modified variant of his theory, *interventionism**, which comes with a correspondingly modified

[6]A fat-handed intervention is an intervention that causes its effects along two (or more) different paths (Scheines 2005, 931-32). Similarly to us, Romero (2015) and Eronen and Brooks (2014, 194) have recently argued that it follows from non-reductive physicalism and interventionism that interventions on phenomena are necessarily fat-handed (or common causes).

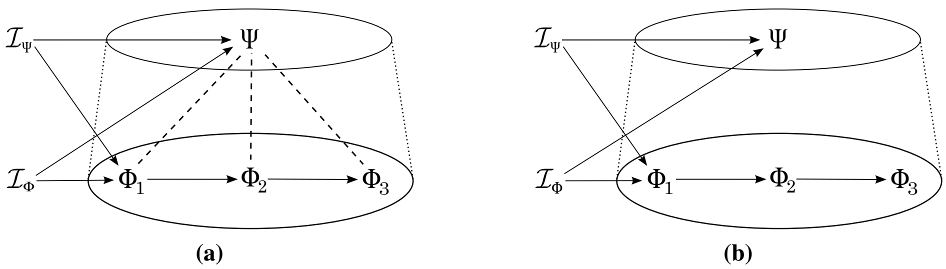

Figure 3: Two empirically indistinguishable mechanistic models.

notion of an intervention.[7] He (2015, 334) now defines an intervention on Ψ w.r.t. Φ to be a variable $\mathcal{I}_\Psi$ taking one of its values, $\mathcal{I}_\Psi = i_n$, and thereby fixing the value of Ψ without having an impact on Φ that is not mediated via Ψ *or via a variable* Γ*, which is related in terms of supervenience to* Ψ *or* Φ, and without being correlated with any other (off-path) cause Γ of Φ *such that* Γ *is not related in terms of supervenience to* Ψ *or* Φ. Against that background, $\mathcal{I}_\Psi$ can pass as an intervention variable for Ψ w.r.t. Φ even if $\mathcal{I}_\Psi$ is connected to Φ along causal paths that do not go through Ψ but through variables related to Ψ by supervenience. That is, according to interventionism*, common causes of macro and micro levels of mechanisms as in Figure 2b count as interventions, because the macro level supervenes on the micro level. If such interventions are moreover associated with changes on both levels, which in light of the fat-handedness of these interventions will consistently be the case, MM turns out to be applicable to mechanistic systems and to entail that different levels of such systems are related in terms of constitution.

That is, giving up the surgicality requirement for interventions allows the mutual manipulability framework to steer clear of the reductio argument of the previous section. Nonetheless, the fact remains that all of the non-surgical interventions on one level of a mechanism that are associated with changes on the other level are fat-handed, which, as this section will show, has far-reaching consequences for the inference to constitution and—a fortiori—for theories that analyze constitution in terms of the *existence* of (possible) manipulations of mechanisms.

Fat-handed interventions generate confounded data, which, in turn, greatly diminishes the inferential leverage delivered by them (cf. e.g. Scheines 2005). Data produced by a common cause of two target variables are uninformative as regards the relationship between these variables. To see this, consider the fat-handed interventions depicted in Figure 3a. If such interventions bring about correlations between the upper and lower levels, these correlations can be fully accounted for by the mere fact that the two levels are wiggled with a fat hand. Hence, there is no need at all to stipulate the existence of additional constitutive dependencies. Model

[7] Woodward himself sees no discontinuity between (2015) and the theory in (2003). However, as core definitions of the theory change between (2003) and (2015), we do not consider it inappropriate to refer to the latter as modifying the former.

3a, which features constitutive dependencies, and model 3b, which does not, imply the very same correlations under manipulations via $\mathcal{I}_\Psi$ and $\mathcal{I}_\Phi$. As all manipulations that induce changes on both levels of a mechanistic system are fat-handed, this finding, again, can be generalized: relaxing the constraints imposed on interventions along the lines of interventionism* entails that the mutual manipulability of macro and micro levels can always be accounted for by the mere fat-handed nature of corresponding manipulations. Mutual manipulability via common causes does not provide a rationale for inferring constitutive relations.[8] For every model featuring constitution there exists a pure common-cause model that entails the very same correlations under manipulations and, hence, cannot be distinguished from the former model empirically.

Nonetheless, if two common causes of Ψ and Φ_1 (such as $\mathcal{I}_\Psi$ and $\mathcal{I}_\Phi$ in Figure 3) that count as interventions by the standards of interventionism* yield changes on both levels of the system, MM—interpreted against the background of interventionism*—infers that Φ_1 is constitutively relevant to Ψ. And Craver is not alone in contending that a few suitable experiments (in a controlled laboratory context) can afford *conclusive* experimental evidence for constitution. Harbecke (2015), for instance, proposes a variant of Mill's (1843) method of difference that he claims to be apt for evidence-based constitutional discovery. In a nutshell, the idea is that if in one of two test situations, which are homogeneous with respect to instantiations of unmeasured constituents of a scrutinized phenomenon Ψ, a change is induced on a mereological part Φ of Ψ such that Ψ changes its value while both Φ and Ψ remain unchanged in the other of the two test situations, it can be inferred that Φ is a constituent of Ψ. However, the fat-handedness of all experimental manipulations of different levels of mechanistic systems inevitably yields that attempts at generating conclusive evidence for constitutive relations are bound to fail. Even in laboratory contexts that are perfectly homogeneous with respect to unmeasured factors that (causally or constitutively) determine a phenomenon Ψ, it is impossible to generate unconfounded data on constitutive relations. As a result, the inference to constitution is systematically underdetermined by evidence.

What is more, empirical underdetermination affecting constitutional inference differs from empirical underdetermination as is known from causal inference. Empirical data can often be accounted for in terms of different causal models that fare equally well with respect to all parameters of model fit (cf. e.g. Spirtes et al. 2000, 59-72; Eberhardt 2013). Such underdetermination ultimately stems from the complexity of causal structures in the world we live in and from our limited capacities for controlling background influences in ordinary discovery contexts. The resulting noise in typical real-life data yields that the latter, ever so often, do not unambiguously reflect underlying causal structures. But this common form of empirical underdetermination can be resolved in ideal discovery contexts. Causes and

[8] Harinen (2014), who also points out that MM imposes unsatisfiable requirements on interventions, fails to see that relaxing the requirements in the vein of interventionism* deprives interventions on mechanisms of their inferential value.

effects are mereologically independent entities. It is possible to surgically intervene on a cause with respect to its effect. Moreover, since it takes time for causal influence to be transmitted from the cause to the effect, an effect can be suppressed via suitable interventions even after the cause has occurred, that is, causal interactions can be broken. As a result, cause-effect pairs can, at least in principle, be isolated from confounding background influences. It follows that contexts of causal discovery can be idealized to such a degree that crucial experiments become possible that produce unconfounded data providing conclusive evidence for causal dependencies. The paradigm example of such an ideal discovery context is the experimental setting envisaged in Mill's method of difference, which also underlies interventionist approaches to causal inference. If a variable $\mathcal{X}$ is associated with changes in a (mereologically independent) variable $\mathcal{Y}$, when all other causes of $\mathcal{Y}$ are fixed and all further required background assumptions are warranted (e.g. that $\mathcal{Y}$ does not change in an uncaused manner, i.e. miraculously), it conclusively follows that $\mathcal{X}$ is causally relevant to $\mathcal{Y}$ (Hofmann and Baumgartner 2011). That is, there exist ideal discovery circumstances in which causal relations can receive unambiguous empirical support.

Such ideal discovery circumstances cannot possibly exist for constitutive relations. Since constituents realize phenomena on the micro level, manipulating phenomena is tantamount to manipulating their constituents. It is impossible to surgically intervene on phenomena, break constitutive dependencies, and isolate phenomenon-constituent pairs. As a result, even in ideal discovery contexts in which all unmeasured relevant factors for a scrutinized phenomenon Ψ are (assumed to be) fixed and all further required background assumptions are warranted (e.g. that the spatiotemporal parts of Ψ have been correctly identified), it is impossible to produce unconfounded data furnishing conclusive evidence for constitutive relations. The reason is that in constitutional discovery—as opposed to causal discovery—data confounding is introduced by the very experimental manipulations intended to uncover constitution. Therefore, even data generated under ideal constitutional discovery circumstances can always be accounted for both by a model featuring constitutive dependencies and by a model without such dependencies. Or differently, even if the hypothesis "Φ is a constituent of Ψ" is experimentally tested in isolation (i.e. such that the whole theoretical background is taken to be beyond doubt), no evidence can be produced demonstrating that this hypothesis is true and its negation false. Contrary to the case of causation, there cannot exist an *experimentum crucis* for constitution. Being experimentally underdetermined is an *inherent* feature of constitutive relations.

This shows that analyses of constitution in terms of the existence of suitable experimental manipulations of mechanistic systems—be it of the surgical or non-surgical type—are beyond repair. In particular, the basic idea behind Craver's MM, *viz.* to account for constitution by supplementing the resources of the most popular difference-making theory of causation, interventionism, by a parthood and a mutuality tweak, is not just misguided because of unrealistic surgicality requirements but because phenomena and their constituents simply are not difference-makers of one another. Rather, they share common difference-makers in their mutual causal past, that is, they are unbreakably coupled via common causes. The fundamental differences between causation and constitution yield that these two relations must be theoretically accounted for in fundamentally different terms and uncovered by following fundamentally different methodological protocols.

4 An abductive alternative

In this section, we propose an alternative theory of constitutive relevance, which avoids MM's problems by further developing the main finding of the previous sections: the characteristic feature of constitution is not the *possibility* of surgical top-down and bottom-up interventions on mechanisms, but the *impossibility* of such interventions. Whatever makes a difference on both levels of a mechanism necessarily does so along different causal paths, because constituents are spatiotemporal parts of phenomena and, hence, not themselves causally related to the latter. Therefore, while causal dependencies can be broken by means of suitable surgical interventions, there do not exist surgical interventions that could break constitutive dependencies. Rather, constitution relates macro and micro levels in such a way that they are *unbreakably coupled via common causes*.

To render this idea precise, it must first be emphasized that the common-cause coupling that marks constitution is characterized by an asymmetry between macro and micro levels, which stems from the fact that a phenomenon Ψ of a given mechanism supervenes on its constituents in a constituting set $\mathbf{\Phi} = \{\Phi_1, \ldots, \Phi_n\}$, but not vice versa. In consequence, every cause of Ψ is necessarily associated with a change in at least one element of $\mathbf{\Phi}$ and, thus, it causes the latter change on a path that does not go through Ψ. That is, every cause of Ψ necessarily is a common cause of Ψ and at least one element of $\mathbf{\Phi}$. The same, however, does not hold for causes of the constituents in $\mathbf{\Phi}$. Supervenience does not exclude the possibility of causing changes in the supervenience base that are not associated with a change in the supervening property. For instance, if two values ϕ_n and ϕ_m of Φ_1 realize the same value ψ_k of Ψ, causes that induce a change from $\Phi_1 = \phi_n$ to $\Phi_1 = \phi_m$ (or vice versa) are invariably associated with $\Psi = \psi_k$. Such causes can count as surgical and, hence, the micro level may be surgically manipulated. But since surgical micro-level causes are not associated with changes on the macro level, they are non-revealing with respect to constitutive relations—and thus irrelevant for an analysis of constitution. The micro-level causes that are of relevance to account for constitution are the ones that *are* associated with changes on the macro level, and it does hold that these causes are common causes of the micro and macro level.

The condition of common-cause coupling needs further refining, for, as stated, it may also be satisfied by variable sets containing parts of a phenomenon that do not constitute it. To illustrate, reconsider the mechanism in Figure 1. The variable Φ_4, which represents a car's AC, is an effect of the constituent Φ_2, the car's running engine, but not itself a constituent of the phenomenon Ψ_2, the car's cruise.

In Craver's (2007, 143) jargon, the AC is a "sterile effect" of the mechanism, that is, a constitutively irrelevant downstream effect of a constituent. Many common causes affect both the AC and the moving car. For instance, starting the engine initiates both the AC and the car's movement, and, conversely, stopping the engine terminates both the AC and the movement. And yet, the AC is not a constituent, for an obvious reason: the influence of the common causes of the cruising car and its AC is always mediated via another spatiotemporal part of the car, namely its engine, which *is* a constituent of the cruise. Or consider the causally and constitutionally isolated variable Φ_5, which represents the behavior of the car's ashtray. Many causes of Ψ_2 will also cause changes in Φ_5. For example, an accident can both warp the ashtray and stop the car. But that an accident can be a common cause of a deformed ashtray and a terminated cruise does not show that the ashtray is a constituent of the car's movement. The reason is, again, obvious: the accident not only warps the ashtray but also the engine, which *is* a constituent of the car's movement.

To ensure that a constituting set $\mathbf{\Phi}$ exclusively contains constituents, and in particular that it excludes non-constituents such as sterile effects and isolated parts, we require that $\mathbf{\Phi}$ be *redundancy-free* in the sense that no proper subset of $\mathbf{\Phi}$ is common-cause coupled with the phenomenon. That is, for every proper subset $\mathbf{\Phi}'$ of $\mathbf{\Phi}$, some cause of the phenomenon exists that fails to be associated with a change in $\mathbf{\Phi}'$ and, instead, is associated with a change in an element of $\mathbf{\Phi}$ outside of $\mathbf{\Phi}'$. Non-constituents such as sterile effects or isolated parts can be eliminated from a set of parts of the phenomenon without breaking the common-cause coupling of the phenomenon and the remaining parts: any change in the phenomenon still is associated with a change in some element of the resulting subset. More concretely, Φ_4 and Φ_5 are not part of a constituting set $\mathbf{\Phi}$ of the phenomenon Ψ_2 in Figure 1 because the set $\mathbf{\Phi}'' = \{\Phi_1, \Phi_2, \Phi_3, \Phi_4, \Phi_5\}$ is not redundancy-free: it contains a proper subset, $\mathbf{\Phi} = \{\Phi_1, \Phi_2, \Phi_3\}$, that is common-cause coupled to Ψ_2, meaning that every cause of Ψ_2 is a common cause of Ψ_2 and at least one element of $\mathbf{\Phi}$.

Overall, we contend that a first defining feature of a constituting set $\mathbf{\Phi}$ of a phenomenon Ψ is that it is *common-cause coupled with* Ψ *in a redundancy-free manner*, meaning that both of the following two conditions are satisfied. First, every cause of Ψ is a common cause of Ψ and at least one Φ_i in $\mathbf{\Phi}$. Second, no proper subset $\mathbf{\Phi} \setminus \{\Phi_i\}$ is common-cause coupled with Ψ—more precisely, for any Φ_i in $\mathbf{\Phi}$, there exists at least one cause of Ψ, which is not a common cause of Ψ and any Φ_j in $\mathbf{\Phi} \setminus \{\Phi_i\}$.

However, many variable sets $\mathbf{V}_i$ including only causally related variables may comprise a variable V_1 and a subset $\mathbf{V}'_i$ of $\mathbf{V}_i$, which does not include V_1, such that V_1 and $\mathbf{V}'_i$ are common-cause coupled (i.e. such that every cause of V_1 in $\mathbf{V}_i$ is a common cause of V_1 and some element of $\mathbf{V}'_i$) and $\mathbf{V}'_i$ is redundancy-free. Hence, the criterion of common-cause coupling needs to be supplemented by a further criterion, which discriminates between constitutional and non-constitutional (i.e. coincidental) common-cause couplings. We contend that the identifying feature of constitutional common-cause couplings is their *unbreakability*. That is, if the set of analyzed variables is expanded, coincidental common-cause couplings may be broken whereas constitutional common-cause couplings will persist across all variable set expansions. More concretely, if a variable V_1 and a set $\mathbf{V}_i'$ coincidentally happen to be common-cause coupled within a set $\mathbf{V}_i$, expansions of $\mathbf{V}_i$ by further causes of V_1 are bound to unveil, say, surgical causes of V_1 that do not induce changes in $\mathbf{V}_i'$—to the effect that the common-cause coupling disappears. This is excluded in cases of constitution. The common-cause coupling of a phenomenon Ψ and a constituting set $\mathbf{\Phi}$ is not a mere contingency of a modeled set $\mathbf{V}$, which contains Ψ and $\mathbf{\Phi}$ along with a given number of their common causes; rather, it is a structural necessity of the relationship between Ψ and $\mathbf{\Phi}$, which not only holds relative to $\mathbf{V}$ but also relative to *every expansion* of $\mathbf{V}$.

That means that, even though the constitutional model 3a and the pure causal model 3b in Figure 3 are indistinguishable relative to data on the variables in the set $\mathbf{V}_3 = \{\mathcal{I}_\Psi, \mathcal{I}_\Phi, \Psi, \Phi_1, \Phi_2, \Phi_3\}$, they are not equivalent in their implications for what happens under expansions of $\mathbf{V}_3$. Model 3a with its constitutive dependencies entails that the common-cause coupling of Ψ and its constituents cannot be broken by expanding $\mathbf{V}_3$, whereas model 3b does not have any such implications. Rather, according to model 3b it is to be expected that, sooner or later, surgical causes of Ψ and Φ_1 to Φ_3 will be found that break their coupling.

These different implications render it possible to choose between constitutional and pure common-cause models. Gradually expanding $\mathbf{V}_3$ (in a series of follow-up studies) will yield one of two outcomes: (I) the common-cause coupling of Ψ and $\mathbf{\Phi} = \{\Phi_1, \Phi_2, \Phi_3\}$ is broken or (II) it is not broken. In case of outcome (I), the attempt to model the relationship between Ψ and $\mathbf{\Phi}$ constitutionally is empirically rejected. If, say, a follow-up study reveals a surgical cause of Ψ, $\mathbf{\Phi}$ is shown not to be a constituting set of Ψ, meaning that the common-cause model prevails. By contrast, outcome (II) gives preference to modeling $\mathbf{\Phi}$ as a constituting set of Ψ, notwithstanding the fact that the correlations in every expansion of $\mathbf{V}_3$ can likewise be reproduced by a mere common-cause model. The reason is that a constitutional model not only reproduces the empirical correlations but also *explains* why the common-cause coupling of Ψ and $\mathbf{\Phi}$ is not broken. That is, in case of outcome (II), a constitutional model is preferable over a pure common-cause model because it exceeds the latter in explanatory power. While it is a structural necessity of the constitutional model 3a that Ψ and $\mathbf{\Phi}$ remain common-cause coupled in all expansions of $\mathbf{V}_3$, *viz.* that outcome (II) obtains, the common-cause model 3b provides no reason whatsoever why surgical causes of Ψ cannot be found. The inference to constitution is thus inherently *abductive*: constitutional models are preferable over pure causal models because they explain both the highly correlated behavior of phenomena and their constituents as well as the *impossibility to de-couple* them.[9]

[9]Our proposal on how to model the coupling of phenomena and constituents bears certain similarities to Causey's proposal on how to interpret biconditional dependencies used as bridge laws in theory reductions. For Causey (1977, chs. 2, 5), those dependencies express attribute identities if all attempts at causally explaining them have failed—for identity then is the only explanation left standing (Causey 1977, 98-99). (We thank an anonymous reviewer for indicating this parallel to us.)

By letting a *complex instance* of a variable set $\mathbf{\Phi}$ designate the occurrence or process (in a particular spatiotemporal region) represented by a complex value assignment $\Phi_1 = \phi_1, \ldots, \Phi_n = \phi_n$ to all elements of $\mathbf{\Phi}$, we can now introduce the definitional details of our proposed *No De-Coupling* (NDC) theory of constitutive relevance (where the specificity of the variables is again made explicit):

(NDC) $\Phi_1(X_1)$ is constitutively relevant to $\Psi(S)$ if, and only if, there exists a variable set $\mathbf{V}$ containing $\Psi(S)$ and a proper subset $\mathbf{\Phi}(\mathbf{X}) = \{\Phi_1(X_1), \ldots, \Phi_n(X_n)\}$, such that:

(1) **Parthood**. For every complex instance of $\mathbf{\Phi}(\mathbf{X})$, there is an instance of $\Psi(S)$ such that the former is part of the latter.

(2) **Coupling**.

(i) Every cause of $\Psi(S)$ in $\mathbf{V}$ is a common cause of $\Psi(S)$ and at least one $\Phi_i(X_i)$ in $\mathbf{\Phi}(\mathbf{X})$;

(ii) for no $\Phi_i(X_i)$ in $\mathbf{\Phi}(\mathbf{X})$ does $\mathbf{\Phi}(\mathbf{X}) \setminus \{\Phi_i(X_i)\}$ comply with (2.i).

(3) **No De-Coupling**. The Coupling of $\mathbf{\Phi}(\mathbf{X})$ and $\Psi(S)$ cannot be broken (invalidated) by expanding $\mathbf{V}$.

NDC differs in a number of crucial ways from MM. First, it replaces MM's mutual manipulability conditions by Coupling and No De-Coupling. This replacement ensures that NDC avoids the problems of MM: while MM defines constitution in terms of the *possibility* of top-down and bottom-up interventions, Coupling and No De-Coupling essentially define it in terms of the *impossibility* of such interventions. Coupling and No De-Coupling capture what we take to be the characteristic feature of constituents, *viz.* that they are unbreakably linked to their phenomenon via common causes.

Second, on a related note, while MM cashes out constitution broadly on a par with causation, NDC defines it in stark opposition to causation. More concretely, according to difference-making theories of causation, of which interventionism and interventionism* are popular instances, a causal relation among $\mathcal{X}$ and $\mathcal{Y}$ is analyzed in terms of the *existence* of (possible) probabilistic or counterfactual difference-making scenarios. As §2 has shown, MM likewise renders constitution dependent on the existence of (possible) mutual difference-making scenarios. That is, both causation and constitution are rendered as existentially defined relations. In consequence, that some entities are related in terms of causation or constitution promises to be conclusively verifiable by exhibiting the existence of required difference-making scenarios. By contrast, NDC defines constitution in terms of the *nonexistence* of (possible) surgical causes. As a negative existential is the same as a universal negation, NDC renders constitution as a universally defined relation. According to NDC, it thus holds that a constitutive relation cannot be conclusively verified; rather, it can only be inductively corroborated.

In that light, NDC-defined constitution is much closer to non-causation or causal irrelevance than to causation. To establish that $\mathcal{X}$ is causally irrelevant to

$\mathcal{Y}$, according to difference-making theories, requires establishing the nonexistence of a (possible) scenario where $\mathcal{X}$ makes a difference to $\mathcal{Y}$. No (finite) data sample could ever conclusively establish this. Due to the fact that it is a universally defined relation, causal irrelevance—just as NDC-defined constitution—can only be inductively corroborated. Of course, NDC does not yield a notion of constitution that is co-extensional with causal irrelevance. After all, constitution is a robust dependence relation, whereas the extension of causal irrelevance encompasses many independent entities. Still, the extension of the notion of NDC-constitution is a proper subset of the extension of the notion of causal irrelevance (as defined by difference-making theories). This squares nicely with our initial background assumption—taken from the canon of mechanistic theorizing—that constitution is a distinctly non-causal form of dependence (Craver and Bechtel 2007).

A third manifest difference between NDC and MM is that the former defines constitutive relevance of Φ_i to Ψ with recourse to *sets* of variables, whereas, according to the latter, that relation is defined in terms of the pair $\langle \Phi_i, \Psi \rangle$ alone. That is, while MM renders constitutive relevance as intrinsic property of the pair $\langle \Phi_i, \Psi \rangle$, NDC renders it as extrinsic property of that pair, which depends on whether Φ_i is contained in a constituting set of Ψ. This does justice to the fact that, contrary to causal dependencies, pairwise constitutive dependencies cannot be isolated from their context; rather, a constituent Φ_i is an indispensable element of a system that figures as supervenience base realizing Ψ in a given mechanistic context—to the effect that the phenomenon cannot exist without the constituents. At the same time, the set-relativity of NDC-defined constitution does not turn it into a relativized notion. Whether Φ_i is a constituent of Ψ does not depend on the existence of a particular constituting set but on the existence of *any* such set. Moreover, No De-Coupling ensures that constitutive relations are constant across variable set expansions. They either hold relative to all expansions of a set that is common-cause coupled to a scrutinized phenomenon or they do not hold at all.

5 Constitutional discovery in a new light

As indicated in §2, one of the most attractive features of MM is that it is straightforwardly operationalizable methodologically: to establish that a spatiotemporal part is a constituent of a phenomenon, it is sufficient to produce one successful top-down and one bottom-up intervention each. We have seen, however, that MM does not ground a viable method of constitutional discovery. For reasons of space, we must postpone a methodological operationalization of NDC to another occasion. Still, we want to emphasize that, against the background of NDC, constitutional discovery appears in a very different light than against the background of MM.

The role NDC attributes to experimental manipulations of mechanisms differs fundamentally from the role attributed to them by MM. MM calls for top-down and bottom-up manipulations whose intended purpose is to reveal mutual difference-making. By contrast, NDC—being formulated in terms of common-cause coupling instead of mutual difference-making—does justice to the fact that there cannot exist top-down and bottom-up manipulations in the first place, that is, manipulations that indirectly induce changes on one level of a mechanism by virtue of directly changing the other level. Rather than for mutual difference-making, NDC requires testing for unbreakable common-cause coupling, which is considerably more challenging than the simple test designs demanded by MM.

Whereas, according to MM, one successful top-down and one bottom-up manipulation warrant an inference to constitution, corroborating the unrestricted universal quantifiers in Coupling and the impossibility operator in No De-Coupling calls for a whole battery of severe tests. In a nutshell, the ways of manipulating a phenomenon Ψ and the elements of a candidate constituting set $\mathbf{\Phi}$ must be systematically altered while the baseline set of modeled variables $\mathbf{V}$ is gradually expanded, in order to check whether, outside of $\mathbf{V}$, there exist surgical causes that break the common-cause coupling of $\mathbf{\Phi}$ and Ψ.

That is, testing for the unbreakability of a common-cause coupling calls for maximally diverse manipulations in maximally diverse background conditions. Of course, it will often be infeasible to conduct experiments on the whole space of common causes of a systems's levels, and all variable set expansions are bound to be finite. Therefore, after a finite number of expansions and severe but unsuccessful attempts at breaking recovered common-cause couplings, the satisfaction of Coupling and No De-Coupling by $\mathbf{\Phi}$ and Ψ must be inductively inferred. This, in turn, licenses an abductive inference to the constitutive relevance of every Φ_i in $\mathbf{\Phi}$.[10] By contrast, if variable set expansions reveal a surgical cause of Ψ that does not target any element of $\mathbf{\Phi}$, the common-cause coupling of $\mathbf{\Phi}$ and Ψ is falsified. Note, however, that this does *not* also falsify the constitutive relevance of a particular Φ_i in $\mathbf{\Phi}$ for Ψ. Even if $\mathbf{\Phi}$ is not a constituting set, Φ_i might still be contained in another set $\mathbf{\Phi}'$, which is unbreakably common-cause coupled to Ψ. That is, while the common-cause coupling of $\mathbf{\Phi}$ and Ψ is conclusively falsifiable, constitutive relevance of one particular Φ_i in $\mathbf{\Phi}$ is not; rather, that Φ_i fails to be constitutively relevant to $\mathbf{\Phi}$ can only be inductively corroborated by means of an extended unsuccessful search for a constituting set comprising Φ_i. Overall, as to NDC, *every* constitutional inference—be it to constitution or to non-constitution—inevitably involves an inductive leap.

Independently of how exactly a viable method of constitutional discovery will eventually look like, we contend that the following methodological consequence of NDC must be respected by any such method. Even under ideal discovery circumstances, no finite number of experimental manipulations can be sufficient for conclusively warranting a constitutional inference. In particular, there does not exist a design for an *experimentum crucis* for constitution. Rather, establishing constitutive relations requires an extended test series exploring the whole space of

[10] There is no universal rule determining how much a variable set needs expanding before an inference to constitution as defined by NDC is justified. Rather, the justification depends on the particularities of a given research context (e.g., the size of the original baseline set, the nature of the mechanism under scrutiny, or the amount of available prior knowledge about that mechanism).

possible ways of breaking the coupling of macro and micro levels. Only if these tests are unsuccessful, an inference to constitution is warranted. And since the evidence for the unbreakability of common-cause couplings is never conclusive, constitution can only be inductively corroborated.

6 Conclusion

If, as is standardly assumed in mechanistic theorizing, constitution is a non-causal form of dependence, and phenomena are non-reductively supervening on their constituents, constitution cannot be adequately accounted for in terms of mutual difference-making relations, along the lines of Craver's (2007) mutual manipulability theory (MM)—which likewise invalidates all accounts of mechanistic explanation based on MM. The reason is that different levels of a mechanism can only be manipulated simultaneously, on different causal paths. In other words, the top-down and bottom-up interventions required by MM are impossible, because phenomena and their constituents can only be manipulated with a fat hand. In light of the inevitable fat-handedness of interventions on mechanisms the inference to constitution is inherently underdetermined by evidence: for every constitutional model there exists an empirically equivalent pure common-cause model.

In that light, this paper developed an abductive theory of constitution (NDC), according to which constitution is a dependence relation that best explains why phenomena and some of their spatiotemporal parts are unbreakably coupled via common causes. Rather than in terms of the possibility of mutual difference-making scenarios, NDC spells out constitution in terms of the impossibility of such scenarios. Against that background, constitutional inference and discovery appear in a new light. Subject to NDC, establishing that Φ is a constituent of Ψ amounts to establishing that Φ is contained in a set $\mathbf{\Phi}$ that is unbreakably common-cause coupled to Ψ; this, in turn, calls for a whole battery of experiments exploring the space of possible manipulations of the scrutinized system. The abductive inference to constitution is warranted only if all attempts to de-couple $\mathbf{\Phi}$ and Ψ have failed.

An obvious question remains: is it possible to faithfully reconstruct constitutional reasoning in science in terms of NDC? Craver (2007) contends that MM provides a faithful reconstruction of scientific practice, and he has undertaken considerable efforts to interpret real-life studies on the basis of MM. We have done none of that sort for NDC here. Still, our results demonstrate that MM cannot ground a viable method of constitutional discovery. Hence, if we grant practicing scientists that they uncover constitutive relations based on some viable method—whichever this may be—MM cannot faithfully reconstruct their practice. How well NDC fares in this respect is thus the crucial follow-up question to be answered in future work.

Acknowledgements We thank the audiences of the annual conference of the British Society for the Philosophy of Science, Manchester, 2–3 July 2015, and of the European Society for the Philosophy of Science, Düsseldorf, 23–26 September 2015. We are especially grateful to Beate Krickel for very helpful comments and discussions, and to three anonymous reviewers for comments on earlier versions of the paper. This research was generously supported by the Swiss National Science Foundation, grant no. PP00P1_144736/1 for MB and grant no. CRSII 1_147685/1 for LC.

References

Baumgartner, M. and A. Gebharter (2015). Constitutive Relevance, Mutual Manipulability, and Fat-Handedness. *British Journal for the Philosophy of Science*. doi: 10.1093/bjps/axv003.

Casati, R. and A. C. Varzi (1999). *Parts and Places. The Structures of Spatial Representation*. Cambridge, MA: MIT Press.

Causey, R. L. (1977). *Unity of Science*. Dordrecht: Reidel Publishing.

Couch, M. B. (2011). Mechanisms and Constitutive Relevance. *Synthese 183*(3), 375–88.

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

Craver, C. and W. Bechtel (2007). Top-down Causation Without Top-down Causes. *Biology & Philosophy 22*(4), 547–63.

Eberhardt, F. (2013). Experimental Indistinguishability of Causal Structures. *Philosophy of Science 80*(5), 684–96.

Eronen, M. (2011). *Reduction in Philosophy of Mind: A Pluralistic Account*. Frankfurt (Main): Ontos.

Eronen, M. I. and D. S. Brooks (2014). Interventionism and Supervenience: A New Problem and Provisional Solution. *International Studies in the Philosophy of Science 28*(2), 185–202.

Glennan, S. (1996). Mechanisms and the Nature of Causation. *Erkenntnis 44*(1), 49–71.

Harbecke, J. (2010). Mechanistic Constitution in Neurobiological Explanations. *International Studies in the Philosophy of Science 24*(3), 267–85.

Harbecke, J. (2015). The Regularity Theory of Mechanistic Constitution and a Methodology for Constitutive Inference. *Studies in History and Philosophy of Biological and Biomedical Sciences 54*, 10–19.

Harinen, T. (2014). Mutual Manipulability and Causal Inbetweenness. *Synthese*. doi: 10.1007/s11229-014-0564-5.

Hofmann, U. and M. Baumgartner (2011). Determinism and the method of difference. *Theoria 26*, 155–176.

Kaplan, D. M. (2012). How to Demarcate the Boundaries of Cognition. *Biology and Philosophy 27*(4), 545–70.

Leuridan, B. (2012). Three Problems for the Mutual Manipulability Account of Constitutive Relevance in Mechanisms. *British Journal for the Philosophy of Science 63*(2), 399–427.

Machamer, P. K., L. Darden, and C. F. Craver (2000). Thinking About Mechanisms. *Philosophy of Science 67*(1), 1–25.

Mill, J. S. (1843). *A System of Logic*. London: John W. Parker.

Romero, F. (2015). Why There Isn't Inter-level Causation in Mechanisms. *Synthese 192*(11), 3731–55.

Ross, D. and J. Ladyman (2010). The Alleged Coupling-Constitution Fallacy and the Mature Sciences. In R. Menary (Ed.), *The Extended Mind*. Cambridge, MA: MIT Press.

Scheines, R. (2005). The Similarity of Causal Inference in Experimental and Non-Experimental Studies. *Philosophy of Science 72*(5), 927–40.

Spirtes, P., C. Glymour, and R. Scheines (2000). *Causation, Prediction, and Search* (2nd ed.). Cambridge: MIT Press.

Spohn, W. (2006). Causation: An Alternative. *British Journal for the Philosophy of Science 57*(1), 93–119.

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

Woodward, J. (2015). Interventionism and Causal Exclusion. *Philosophy and Phenomenological Research 91*(2), 303–47.