

# Merleau-Ponty's Implicit Critique of the New Mechanists

Benjamin Sheredos

---

## Abstract

I argue (1) that what (ontic) New Mechanistic philosophers of science call mechanisms would be material *Gestalten*, and (2) that Merleau-Ponty's engagement with *Gestalt* theory can help us frame a standing challenge against ontic conceptions of mechanisms. In short, until the (ontic) New Mechanist can provide us with a plausible account of the organization of mechanisms as an objective feature of mind-independent ontic structures in the world which we might *discover* – and no ontic Mechanist has done so – it is more conservative to claim that mechanistic' organization is instead a mind-dependent aspect of our epistemic strategies of mechanistic explanation.

---

*Cite only the published version: this is a pre-print of an article published in Synthese. The final authenticated version is available online at:*

<https://doi.org/10.1007/s11229-018-02006-7>.

*This interim version of the article contains portions that were omitted from the final version, and lacks portions/revisions that were added in response to reviewer feedback.*

## Acknowledgments

I thank Rebecca Hardesty and Jason Winning for very helpful feedback on earlier drafts. [Re: the final version: I thank the editors and the anonymous reviewers for helpful feedback.] Some early work leading up to §2 of this paper was presented at the Gateway Graduate Conference at UMSL in March, 2014, and I'd like to thank the organizers: Matt DeStefano and Elliot Risch. Acknowledgments regarding the themes of §4.1 are made in Sheredos (2017). Finally, Bill Bechtel deserves thanks for initially steering me into and through the mechanism literature.

---

\*E-mail address: sheredos@ucsd.edu

## 1. Introduction

In early work, New Mechanistic philosophers of biology seemed to promote what Wesley Salmon (1984a) called an ontic view of explanation. In an ontic view (roughly), the explanation of a phenomenon is, most fundamentally, the mind-independent ontic structure(s) in the world that causes, constitutes, or is otherwise responsible for the occurrence of the phenomenon. Ontic Mechanists hold that mechanisms are such structures. This view has been further articulated and promoted by Craver (2007, 2014). An alternative is what Salmon called an epistemic view, in which (roughly) explanations are, most fundamentally, mind-dependent strategies of making phenomena intelligible, predictable, or understandable. According to robust epistemic Mechanists, mechanisms are theoretically constructed as part of such a strategy (cf. Bechtel & Abrahamsen 2005, Bechtel 2015).

Here I invoke the history of philosophy to articulate what is today a novel challenge to ontic Mechanists. First (§2) I argue that the ontic Mechanists' ontological commitment to mechanisms is equivalent to the Gestaltists' commitment to what they called "physical *Gestalten*" – what I call "material *Gestalten*" (Köhler 1920; Koffka 1936).<sup>1</sup> I show that Mechanists endorse a form of holism regarding mechanisms' organization and constitution. Just such holism, I argue, is what the Berlin Gestaltists highlighted when distinguishing material *Gestalten* from mere "and-summations." Mechanisms, if they exist in the world, are material *Gestalten*.

I then (§3) show that this ontological commitment has been poorly understood by Mechanists themselves. In short, there is consensus that constitutive relations within organized mechanisms (between parts and whole) should be understood non-causally, but no ontological analysis of them has attained consensus. This leads to a puzzle: how can ontic Mechanists justify commitment to the holistic organization of mechanisms, when leading accounts of Mechanistic explanation seem offer no specification of what it is, exactly, to which we are allegedly ontologically committed?

Finally (§4), I adapt Merleau-Ponty's argument against the Gestaltists to challenge the ontic Mechanists' understanding of our ontological commitment to material *Gestalten*. Merleau-Ponty's own argument involved claims which many authors today may not endorse, but I disentangle his core criticism.

---

<sup>1</sup>I focus on the Berlin School of *Gestalt* theory, not the Austrian school of von Ehrenfels and Meinong: see Smith (1988, p.150).

According to Merleau-Ponty's core criticism, we do not *discover* material *Gestalten* through scientific research, and conducting scientific research does not itself *incur* any ontological commitment to organized *Gestalten*. Instead, (a) we assume in advance that some of the systems we study should be understood as organized wholes, and (b) that assumption is never discharged by some scientific demonstration that it holds true.

I claim that Merleau-Ponty helps us offer a diagnosis of why it has proven so difficult for the New Mechanists to provide any clear ontological account of mechanisms' organization: this ontological commitment cannot be *spelled out in* an ontic view of Mechanistic explanation, since it is *presupposed by* the ontic Mechanist from the outset. Moreover, I claim Merleau-Ponty helps us frame a moderate position which lets us evade the problem of analyzing mechanisms' organization. Absent any objective analysis of mechanisms' organization, we may more moderately maintain that such organization is simply a subjective feature of how we apprehend mechanistic phenomena: it is a mind-dependent feature of epistemic strategies of explanation, not a mind-independent feature of the world. The basic claims required for this view are essentially already expressed in strong forms of the epistemic view of Mechanistic explanation (cf. [Bechtel 2015](#)). Thus, exploring the historical engagement between Merleau-Ponty and *Gestalt* theory helps us frame a standing challenge against ontic conceptions of mechanisms.

## 2. Mechanisms are Material *Gestalten*

In this section I argue that ontic mechanisms are material *Gestalten*. Bringing this to light requires completing three tasks. First (§2.1) I review a common conception of *Gestalten* as only perceptual-phenomenal entities. Second (§2.2) I re-introduce the under-appreciated notion of material *Gestalten*, and clarify the role they were intended to play in the Gestaltists' psycho-physical explanation of perceptual experience. As I shall clarify, the Gestaltists endorsed two claims as characteristic of all *Gestalten*:

***Holistic Individuation*** (HI): There exist parts and/or properties that are individuated only when they occur in some whole (i.e., as dependents parts of some *Gestalt*).

***Holistic Explanation*** (HE): There exist wholes (i.e., *Gestalten*) whose parts exhibit behaviors or properties which are only

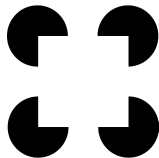
explicable given properties of the whole.<sup>2</sup>

The Gestaltists held that HI and HE characterized not only the organized whole-objects of perceptual experience, but also the organization of material structures. Relations between these different varieties of *Gestalten* were key to the Gestaltist’s proposed model of psycho-physical explanation. I do not intend to promote the Gestaltist’s own view of psycho-physical explanation, but discuss it in some detail to clarify the basic features of material *Gestalten*: the question of whether there exist material *Gestalten* can be decoupled from the explanatory purposes to which Gestaltists put them.

Finally I clarify foundational commitments regarding the organization of mechanisms, and I argue that the New Mechanists are likewise committed to HI (§2.3) and HE (§2.4): mechanisms are *Gestalten*.

### 2.1. Perceptual-phenomenal Gestalten

In contemporary philosophy, “*Gestalten*” are typically understood as certain perceptual-phenomenal structures.<sup>3</sup> For example, Grush<sup>4</sup> provides the “illusory contours” figure below, then invokes a perceptual-phenomenal conception of *Gestalten* when he comments:



“...the explanation would appear to be that something about the stimulus conditions triggers the perceptual system to construct a representation other than the one that accurately reflects what is being presented... The various Gestalt criteria can be given explanations in these terms... What this suggests is that the content of perceptual experience is not a mere passive registration of external events, but involves some degree of active construction...” (Grush, 2008, 150).

<sup>2</sup>I am indebted to Chudnoff (2013) for helping me see the distinction between HI and HE, though he does not frame it quite as I do.

<sup>3</sup>As illustration, the *Cambridge Dictionary of Philosophy* (1999) provides no general entry for “*Gestalten*,” referring the reader to (i) an entry on figure-ground phenomena in perception, and (ii) a short bio of Köhler.

<sup>4</sup>I do not accuse Grush himself of having an over-narrow view of *Gestalten*. I simply borrow his claims to illustrate what such a narrow view would look like.

We can make limited headway in understanding the original conception of *Gestalten* using this case. First, recall the Gestaltists' thesis of HI:

***Holistic Individuation*** (HI): There exist parts and/or properties that are individuated only when they occur in some whole (i.e., as dependents parts of some *Gestalt*).

HI is the claim that the *individuation* of some entities is dependent upon their role in a whole. Here “individuation” is meant in a metaphysical sense, not a cognitive or epistemic sense (Lowe, 2003). Epistemically, we individuate objects by distinguishing them from others – for disambiguation, I call this “identification.” Metaphysically, an object is individuated by what makes it the object it is, distinct in its being from others. In our perception of this illusory contours figure, there is a straightforward sense in which HI holds true. Viewing the figure, we tend to experience a white square, flanked at each corner by occluded black circles. Experiencing the figure of a square requires experiencing its (illusory) *edges*. What is perceptually “filled in” appears (fleetingly) as edges-of-a-square, not free-floating “lines.” The individuation of the edges as *edges-of-a-square* requires their participation in a whole square, in accordance with HI. I take it that this application of HI is relatively straightforward.

HE is another matter:

***Holistic Explanation*** (HE): There exist wholes (i.e., *Gestalten*) whose parts exhibit behaviors or properties which are only explicable given properties of the whole

Applied here, HE means that experiencing apparent edges is in some sense *explained* by properties of one's whole experience when viewing this figure. Grush sketches the more common approach: we explain the experience of edges by appeal to properties of the stimulus and of sub-personal processing.

How might the Gestaltist invoke HE? As Chudnoff clarifies, the Gestaltists pursued a kind of phenomenological explanation, captured by the (somewhat cryptic) doctrine of the Law of *Prägnanz* (2013, p.176).<sup>5</sup> The Law of *Prägnanz* holds that our perceptual experience tends to be as “good” (regular, symmetric, simplistic) as it can be, even if this requires that our experience be an inaccurate representation of stimuli. For the illusory contours

<sup>5</sup>I depart from Chudnoff in what follows: he discusses only phenomenal *Gestalten*.

figure, suppose an experience of a white square occluding four black circles is simpler than an experience of four un-occluded Pac-Man shapes oriented toward the center of the figure. If that were so, and if the Law of Prägnanz held good, we could invoke it to provide a kind of explanation (an answer to a why-question). Why do we experience illusory edges? Because that is part of the simplest whole-experience we can have, faced with this stimulus. On this view, “a property of the whole [experience] – greater simplicity in organization – explains a property of the part” (Chudnoff, 2013, p.176).

## 2.2. Material Gestalten

In my view, the foregoing is not the full account of the Gestaltists’ explanation of this case, and this is because a narrow, perceptual-phenomenal understanding of *Gestalten* does not fully capture the notion.

In his *Principles of Gestalt Psychology*, Koffka introduced *Gestalt* theory as a new worldview which promised, among other things, (a) a new understanding of the nature of facts, theories and science, (b) a comprehensive integration of quantity, order, and meaning (value) into a unified ontology, and (c) a novel resolution of the materialist-vitalist debate in early 20<sup>th</sup> century biology.<sup>6</sup> The notion of material *Gestalten* can be introduced in connection with this last issue. The Gestaltist aims to resolve the materialism-vitalism debate by showing that the kind of “order” which the vitalist claimed to uniquely characterize living systems is not unreal, not the result of some “vital force,” but rather (in a basic form) already present in inanimate physical systems (Koffka 1936, p.17). This approach was pursued emphatically in Köhler’s (1920) *Physical Gestalten in Rest and Stationary State*: the notion of *Gestalten* was extended to the material world itself, to characterize orderly systems at large. To illustrate this view, I sketch the Gestaltist’s complete view of psycho-physical explanation.

Recognizing material (more specifically, physical) *Gestalten* is construed as affirmatively answering the question “Are there physical whole-states or whole structures in which the parts are not mere and-summations [*Undverbindungen*]... of elementary individual states and individual structures?”

---

<sup>6</sup>In the materialist-vitalist debate, the materialist was often a *classical* mechanist: these must be contrasted with the New Mechanists who are my interlocutors here.

(Köhler, 1920, p.25/42).<sup>7</sup> We can specify what this question (and its answer) amounts to by clarifying the concept of an “and-summation”:

“An aggregate [*Zusammen*] of ‘parts’ or ‘pieces’ is a pure ‘summation’ [*eine reine “Summe”*] only when its constituents may be added together one after another without thereby causing any alteration in any of them; and conversely, a summation is that kind of aggregate from which any one or more units may be removed without any effect either on the ones remaining or on the ones removed” (*ibid.*, p.25/42).

Where a collection is an “and-summation,” the properties of the whole are merely summative, and can be understood in terms of: the intrinsic properties of the first item, and the intrinsic properties of the second item, and the intrinsic properties of the third... and so on. For example, the weight of a collection is equal to the weight of all the parts taken together. If all the collection’s parts and properties work like this, it is a mere and-summation.

*Gestalten*, in contrast, exhibit interdependence: a kind of conditioning or determination of parts’ properties occurs in virtue of their role in the whole. Köhler’s flagship examples of physical *Gestalten* concern electrostatic charges in circuits – cases chosen for their presumed applicability to neural systems. A simple case concerns a charged conductor. When charge is applied to an insulated conductor, the charge (*i*) distributes itself over the outer surface of the conductor, and (*ii*) attains equilibrium. The specific distributions which count as equilibria are determined by physical properties of the conductor. In a charged conductor which is, as a whole, at equilibrium, “it is impossible to decrease, increase, or displace any part of this charge alone; for with any such change there occurs a reaction throughout the entire natural structure” (Köhler, 1920, p.28/56). “In a word the structures of static charges upon conductors of given shape are physical *Gestalten*” (Köhler, 1920, p.28/68).

HI applies, because the charge at any point *p* on a conductor *c* is only individuated (it only has the determinate properties it has) because the *whole*

---

<sup>7</sup>Citations to Köhler usually provide (*i*) pagination from the only existing partial translation, due to Ellis (1967), followed by (*ii*) pagination from the first German edition. Ellis employs significant editorializing, and while it is usually apt, I occasionally amend his translation. Occasionally I cite a passage which was not included in Ellis’ translation, and there provide only the page-number from the original.



conductor is at equilibrium.<sup>8</sup> The determinate amount of charge at any point  $p$  could be varied by applying more or less charge to the whole conductor and allowing it to attain equilibrium again.

HE also applies. We can answer the question of *why* the charge at  $p$  is as it is by saying the *whole* conductor has attained equilibrium, given the applied charge. Explanation need not stop here, of course; we can proceed to explain *how* the conductor attains equilibrium in terms of the physical properties of the conductor, the movements of ions, etc. This does not diminish or counteract the holistic explanation in terms of equilibria.

Köhler thought this model applied wholesale to the brain. Neurons were regarded as poorly-insulated conductors. When charged, a *field* of electrostatic forces would surround the neuron, forming a *Gestalt*. A number of such fields in interaction (e.g., in a region of the brain) would form a yet-larger whole-field: another, larger *Gestalt*. The whole brain would exhibit a field composed of myriad lesser *Gestalten*.<sup>9</sup>

Köhler suggested further that a naturalistic explanation of any experience (e.g., the perceptual-phenomenal experience of seeing an illusory figure) would depend upon their being some region of the brain exhibiting a material *Gestalt* of electrostatic forces whose structure was *isomorphic* to the structure of the experience. The perceptual tendency towards good form, captured by the Law of Prägnanz, would be mirrored by the electrostatic tendency towards equilibrium. This doctrine of isomorphism, in its simplest form, claimed: “*actual consciousness resembles in each case the real structural properties of the corresponding psycho-physical processes*” (Köhler, 1920, p.38/193).

My argument does not require viewing this as a workable approach to psycho-physical explanation. I am concerned, initially, with the idea of a material *Gestalt* in its own right: it is a separate question whether this bit of *Gestalt theory* can be put to work in a naturalistic *Gestalt psychology* in the way Köhler envisioned. Likewise, it is irrelevant to my argument whether electrostatic phenomena should, in the final analysis, be counted

---

<sup>8</sup>Again, “individuation” is used in a metaphysical sense: see §2.1 above.

<sup>9</sup>I omit some historical details: initially Köhler held a more chemically-based view of neuronal function.



as material *Gestalten*.<sup>10</sup> I am concerned only with what material *Gestalten* were supposed to be, as characterized by HI and HE. With this sketched, I turn to the New Mechanists.

### 2.3. Mechanisms as Gestalten: HI

New Mechanistic accounts are heterogeneous, and there are many subtly distinct conceptions of mechanisms on offer. However, they all emphasize the organization of mechanisms as an important feature. While it has been overlooked, New Mechanists' claims about organization either flirt with or explicitly commit to the view that mechanisms are material *Gestalten* which satisfy the two conditions HI and HE. I approach HI in this sub-section, then discuss HE in the next.

Machamer, Darden, and Craver offered one influential characterization of a mechanism:

“Mechanisms are entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions” (2000, p.3).

When a mechanism operates, a phenomenon of interest is produced. The aim of mechanistic explanation is to exhibit the “productive continuity” of the organized mechanism, making intelligible how the entities and activities which compose it jointly operate to produce the phenomenon.

Beyond this, Machamer, Darden, and Craver say little to characterize mechanisms' organization. The organization of entities and activities “determines the ways in which they produce the phenomenon,” and “entities must often be appropriately located, structured, and oriented, and the activities in which they engage must have a temporal order, rate, and duration” (*ibid.*, p.3). If we dig deeper, we find them flirting with HI. The issue concerns the individuation of mechanisms and their parts. They maintain:

“[On one hand:] Mechanisms are identified and individuated by the activities and entities that constitute them, by their start and finish conditions, *and by their functional roles*.”

---

<sup>10</sup>Likewise, I do not claim that all material *Gestalten* are mechanisms. So it is not a problem if Glennan (1996, p.54) is right that electromagnetic fields cannot be mechanistically explained, in his sense.

[On the other hand, in the next breath:] Functions are the roles played by entities and activities in a mechanism. *To see an activity as a function is to see it as a component in some mechanism...* Functions... should be understood in terms of the activities *by virtue of which entities contribute to the workings of a mechanism*" (Machamer et al., 2000, p.6, my emph.).

The first claim suggests that entities, activities, and their functions are individuated *prior* to entering into an organization to compose a mechanism. The parts' individuation does not depend upon the whole mechanism's organization, rather the whole mechanism is itself individuated on the basis of the entities composing it.<sup>11</sup> The second claim, by contrast, suggests that no entity or activity has a function unless it is *already* a component of a mechanism which is organized to produce a phenomenon. This latter claim runs close to an endorsement of HI: no entity or activity possesses a function unless it is part of an organized whole which produces a phenomenon, for its function just is its contribution to the workings of the whole, and it cannot make such a contribution unless it is situated within that whole.

Other New Mechanists likewise flirt with HI. Consider Glennan:

"My analysis can be summarized by a definition...

(M) A mechanism underlying a behavior is a complex system which produces that behavior by the interaction of a number of parts according to direct causal laws.

Notice that (M) is a definition of a 'mechanism underlying a behavior' rather than a mechanism *simpliciter*. One cannot even identify a mechanism without saying what it is that the mechanism does. The boundaries of the system, the division of it into parts, and the relevant modes of interaction between these parts depend upon what behavior we seek to explain" (Glennan, 1996, 52).

Elsewhere Glennan reasserts that "mechanisms are not mechanisms *simpliciter*, but mechanisms *for* behaviors" (2002, S344). This claim – sometimes called Glennan's Law – has been widely endorsed by New Mechanists

<sup>11</sup>MDC might here use "individuate" in an epistemic sense – see §2.1 above. This will not disturb my argument.

(cf. Glennan 2017, p.23-24; Illari & Williamson 2012, esp. *fn.2* and p.134). If that is so, then plausibly – as Glennan seems to suggest in the tail-end of the passage above – the parts must likewise be understood not as “parts” *simpliciter*, but as parts-of-a-whole-mechanism-*for* some behavior. This flirts with HI: the properties and interactions of parts that are genuinely relevant to the production of the behavior are features they exhibit only while *in* the organization of the whole system they compose. Take “the same parts” out of that organization, and they no longer interact and condition each other’s properties. Whatever they might still do, they contribute nothing to the production of the behavior: they are no longer parts-of-a-mechanism-*for* that behavior.

This “quick route” to HI through Glennan’s Law might seem unconvincing. In early work, Glennan might seem to resist it:

“...to prevent (M) from being vacuous... the parts of a mechanism must have a kind of robustness and reality apart from their place within that mechanism. It should in principle be possible to take the part out of the mechanism and consider its properties in another context. Care must be taken so that parts are neither merely properties of the system as a whole nor artifacts of the descriptive vocabulary. I shall summarize these restrictions by saying that parts must be objects” (Glennan, 1996, p.53).

Note, however, that this does not fully evade HI. To treat parts as “objects” in this sense requires only that they have *some* continued existence, and exhibit *some* stable properties, outside of the organization of a mechanism. This is consistent with the possibility that the parts may have other properties only *within* the organization of a mechanism, *per* HI. In what follows, I argue that mechanism parts are claimed to have just such holistically-individuated properties, and that these properties are central to the (ontic) New Mechanists view.

The underlying issue here concerns the extent of “decomposability” which is required of mechanisms, or the extent to which organization is permitted to condition parts’ properties. Both Craver (2007, p.135) and Bechtel & Richardson (1993, p.25) appeal to the work of Wimsatt (1986, 1997) to characterize mechanisms’ organization. Wimsatt articulated four conditions which characterize different varieties of aggregativity. A system is a mere aggregate, it is completely decomposable, if it meets all four conditions:

1. Intersubstitutability of parts: rearranging and swapping parts doesn't influence the properties of the whole.
2. Qualitative similarity: the properties of the whole remain qualitatively similar (if quantitative, varying only in value) with addition or subtraction of parts.
3. Stability through decomposition and re-aggregation of parts: the properties of the whole remain invariant even if previously-independent parts are fused, and even if previously-linked parts are decoupled.
4. Minimal interactions among parts: there are no cooperative or inhibitory interactions among parts which influence properties of the whole.<sup>12</sup>

If a system is a mere aggregate, the specific organization of the parts contributes nothing significant to the features of the whole system: the whole is neatly decomposable into parts whose operations are self-determined. What Wimsatt calls a “mere aggregate” is quite clearly co-extensive with what the Gestaltists called an “and-summation.” Where mere aggregativity (and-summation) is violated, we have a material *Gestalt*.<sup>13</sup>

According to New Mechanists, a prototypical mechanism is precisely not a mere aggregate. Craver (2007, pp.135-136) argues that mechanisms possess an “active organization” in virtue of which they violate all four conditions for mere aggregativity: a mechanism is not just “literally the sum of its parts” (Craver 2007, p.135; see also Glennan 2017, p.23). Bechtel & Richardson (1993, pp.26-32) claim that pursuing mechanistic explanation begins with *fal-lible heuristics* of decomposition and localization, which presume that *some* degree of decomposability is possible, hence that the mechanism under investigation possesses some form of aggregativity. However, they endorse the

<sup>12</sup>Wimsatt's (1986) original presentation was adapted by Bechtel & Richardson (1993, p.25). Wimsatt's (1997, p.S376) more concise presentation was adapted by Craver (2007, p.135). My adaptation here draws inspiration from all these sources.

<sup>13</sup>I am pursuing a conceptual argument that mechanisms' organization makes them material *Gestalten*. There may also be a more direct historical connection. Wimsatt (1986, p.289) acknowledges that his view of aggregativity was likely influenced by disagreements with Ernest Nagel's “Wholes, sums, and organic unities” (1952; rpt. in *The Structure of Science*). Nagel's piece explicitly opposes Köhler's conception of physical *Gestalten*. When New Mechanists borrow Wimsatt's anti-Nagelian view to clarify the nature of mechanisms' organization, they implicitly borrow, second-hand, a Gestaltist view.

view that (New) Mechanistic explanation does not require decomposition and direct localization (*ibid.*, cf. pp.227ff.), and can be pursued even where a system is only “minimally decomposable” – i.e., where conditions 1-4 are violated (*ibid.*, p.26). Just to the extent that conditions 1-4 are violated by mechanistic systems, HI applies to them: the failure of decomposability arises precisely because some parts’ properties are individuated only when they occur in the right kind of whole.

We can clarify in greater detail precisely how the Mechanist implicitly endorses HI regarding such systems. Craver serves as a useful guiding example. A mechanistic explanation of how a system (“S”) exhibits a behavior (“ $\psi$ -ing”) proceeds by clarifying how its parts (“Xs”) exhibit some lesser behavior(s) (“ $\varphi$ -ing”), and how the behaviors of the parts contribute to the behavior of the whole. To provide any such explanation, we must reckon with the *active organization* of the parts, and must deny that they form a mere aggregate. The parts:

“...act and interact with one another in such a way that the  $\psi$ -ing of S is more than just a sum of  $\varphi$ -properties. In fact, the  $\varphi$ -properties of a working mechanism are not just properties; *they are the activities of and interactions among the entities in the working mechanism*. The different components act in cooperation or competition, and they do so with some components and not with others. It matters which Xs  $\varphi$  with which others, and it matters how they interact. This is why the parts of mechanisms often cannot be reorganized randomly... added or subtracted at will... or taken apart and put back together again... without disturbing their corporate ability to  $\psi$ ” (Craver, 2007, p.136, my emph.).

Notice:  $\varphi$ -properties or parts’ activities are fully-individuated as such (they are what they are) only when the *whole* mechanism is working to produce its behavior and the parts are interacting in the organization which composes the mechanism. HI applies. Take “the same parts” out of that organization, or impede the working of the mechanism, and – whatever you might learn about the parts – you will not observe the full set of  $\varphi$ -properties they exhibit *in* the mechanism, since these  $\varphi$ -properties arise only in the context of the whole, active organization. Glennan also endorses this in recent work: “For some entity that is part of a mechanism to actually contribute to the production

of the phenomenon for which the mechanism is responsible, it must produce changes in other parts of the mechanism; there is no production without interaction” (2017, 21-22). The interactions of the parts is what renders mechanisms organized wholes rather than mere aggregates. This is why parts cannot simply be studied in isolation, but must be *recomposed* to provide a complete explanation of how they produce the phenomenon of interest (Bechtel & Abrahamsen, 2009, 2013).

Commitment to HI is in fact encoded again in leading accounts of the “components” or “working parts” of a mechanism. Craver’s account will be a recurring topic of discussion in §3 below, and I shall say more about it there. For now, not everything within a mechanism counts as a working component relevant to the mechanism’s operation. Following Salmon<sup>14</sup> Craver calls this the issue of “constitutive explanatory relevance:” the issue of which of the items spatially contained within a mechanism genuinely contribute to its organized operations, in virtue of which the *explanandum* phenomenon occurs. Craver offers what he calls the mutual manipulability criterion of constitutive relevance, and with it, a practical strategy for identifying those component parts which contribute to a mechanism’s operation:

“...a part is a [working] component in a mechanism if one can change the behavior of the mechanism as a whole by intervening to change the component *and* one can change the behavior of the component by intervening to change the behavior of the mechanism as a whole” (2007, p.141).

The practical orientation here clearly hides ontological commitment to the claim that working components and whole mechanism are co-individuated:

---

<sup>14</sup>Craver (2007, p.8, *fn.*9; p.140, *fn.*19) insists that he follows Salmon’s (1984a, p.297) usage, and does not intend to invoke any other metaphysical conception of constitution. Salmon clarifies his view as follows: “A constitutive explanation is thoroughly causal, but it does not explain particular facts or general regularities in terms of causal antecedents. The explanation shows, instead, that the fact-to-be-explained is constituted by underlying causal mechanisms. Many cases of physical reduction qualify as constitutive explanations. When, for example, we explain optical phenomena in terms of Maxwell’s electromagnetic theory, the explanation is constitutive. Light waves *are* the electromagnetic waves (in a particular part of the spectrum) treated by Maxwell’s theory” (Salmon, 1984b, p.270). Here a metaphysical claim of constitution is treated as a claim of identity (the whole mechanism *is* the organization of the working parts). This is *stronger* than the claim of co-individuation I focus on in this paper.

the component, *qua component*, is thus-and-so only when the whole mechanism is thus-and-so, and *vice versa*. That is *why* an intervention to change one changes the other. Glennan supports these same points in more recent work:

[A]ll mechanisms are mechanisms for some phenomenon... If we characterize a mechanism as a thing S (entity or system) that is  $\psi$ -ing, what counts as part of the mechanism's phenomenon (the  $\psi$ -ing) depends upon what is actually contributing to the production of the phenomenon. Take for instance a car. A car is a system with many mechanisms, but pre-eminently a car is a system for driving on roads. While the car has many parts, not all parts contribute to its driving. The wheels matter but the rear-view mirror does not. Those parts of the car that are implicated in the mechanism are its *working parts*... A similar point may be made about the activities involved... The engine, for instance, is a working part of the car's mechanism for turning the wheels, but produces heat as it rotates the crankshaft. The rotation, but not the heat, is a *working activity* in the mechanism (2017, pp.23-24).

And a similar point, we can add, should be made about the *interactions* between parts. The *working interactions* between parts are what constitute the mechanism's "active organization," in virtue of which they parts jointly contribute to the production of the *explanandum* phenomenon. Other interactions are irrelevant (e.g., light might be reflected from the rear-view mirror onto the steering wheel.)

In sum, the Mechanist aims to explain the operation of a whole mechanism by appeal to the working components, their working activities, and their working interactions. But those components, activities, and interactions are co-individuated with the whole: they are what they are only when the whole is what it is. Thus, HI applies to mechanisms. And it applies non-trivially: HI applies to precisely those features of mechanisms' working parts which arise through the distinctive organization of the mechanism. To the extent that a system is a mechanism, HI must, to some degree, apply.

#### 2.4. Mechanisms as Gestalten: HE

Now, is the Mechanist likewise committed to HE? The sorts of explanations which New Mechanists have championed are roughly the inverse of



HE. Mechanistic explanation is construed as a decompositional or reductive strategy which aims to show how the parts of the mechanism work together to produce a phenomenon (cf. [Bechtel & Richardson 1993](#), p.17; [Machamer et al. 2000](#), p.21; [Bechtel & Abrahamsen 2005](#), p.425 [Craver 2007](#), p.133; [Glennan 2017](#), p.223). In contrast, HE applied to mechanisms would be the claim that some properties of the parts are only explicable with reference to the whole mechanism. Clearly, *for the purposes of mainstream Mechanistic explanation*, HE is irrelevant.<sup>15</sup> Yet, for *other* explanatory purposes, HE appears quite suitable. I offer an illustration.

A common example in the New Mechanist literature is the mechanism of action potential generation in a neuron. We offer a mainstream Mechanistic explanation of how an action potential is generated by adverting to activities of the components – e.g., the joint operations of voltage-gated Na<sup>+</sup> and K<sup>+</sup> channels. But suppose we ask (after an action potential occurs in an intact neuron): why was the Na<sup>+</sup> channel open? *Ceteris paribus*, a perfectly respectable answer seems to be an application of HE: the Na<sup>+</sup> channel was open because the *whole* neuron was firing an action potential. This holistic explanation is available precisely because the opening of the Na<sup>+</sup> channel is a working activity of a working component of the mechanism-for the action potential and is of constitutive relevance for the production of the action potential, such that (*ceteris paribus*) if the Na<sup>+</sup> channel opens, the neuron fires an action potential, and *vice versa*. Where the events occur fully-individuated *as* the phenomenon of interest and the working activity that contributes to it, the two are co-individuated and co-constitutive.

This is not a causal explanation, for all the usual reasons that constitution is not causation (cf. [Craver 2007](#), pp.153-154). It is what Glennan would call a non-causal explanation, a “*why-but-not-what-or-how explanation*” ([2017](#), p.223-224). One who prefers causal explanations will likely find this application of HE deeply unsatisfying. Of course, explanation need not stop here, and the New Mechanist is free to pursue their preferred style of explanation. I do not see that this diminishes holistic explanations of the

---

<sup>15</sup>This is as good as any a place to indicate one more connection between New Mechanism and Gestaltism. In their broadly “reductive” approach, New Mechanists resist the idea that classical “laws” can be cited to *derive* the whole mechanism’s behavior from the behavior of the parts. They thus regard mechanisms as exhibiting what were called, in analyses of Gestaltism, “emergent wholistic properties” ([Rescher, 1953](#), p.327), or “D-G-T underivable attributes” ([Rescher & Oppenheim, 1955](#), p.93).

sort I've just described.

Still, let me offer two arguments to dispel the air of triviality that might surround HE.<sup>16</sup> While initial statements of New Mechanism suggested ontological commitment to mechanistic explanations as things in the world, there has more recently been a split over whether mechanistic explanation should be fundamentally regarded as epistemic or ontic. The epistemic view holds (roughly) that, most fundamentally, explanation is a matter of making some phenomenon predictable, intelligible, or understandable. The ontic view holds that, most fundamentally, explanations are things in the world that cause, constitute, or are otherwise naturally responsible for the occurrence of some phenomenon. Both views should recognize that HE, applied to mechanisms, can offer a legitimate explanation.

Start with the epistemic view. It's a contingent matter of fact that mechanistic explanations tend to be pursued in cases where scientists have identified a phenomenon, made a provisional attempt to circumscribe the mechanism-for that phenomenon, and are now laboring to identify its working parts, activities, and interactions. Under these conditions, we are familiar enough with the phenomenon to target it as our *explanandum*, and we cast our hypotheses about parts with reference to it. Under these conditions, HE will appear unsatisfying: we already *know* (somewhat) what the whole mechanism is doing (but not *how* it does it), so attempting to explain why a part does what it does by invoking the whole's behavior may seem to get us nowhere. However, we could conceivably discover that some system is a working part of a newly-discovered, larger mechanism. That is: we previously understood some of the features and activities of the part, but come to understand it *as* a working part in a previously unidentified mechanism-for a phenomenon. In that case, there would be real novelty involved in coming to understand that the part does what it does when the whole does what it does. We would have *increased* understanding of the part if we could understand its behavior in terms of the behavior of the whole. "Going up a level" can be, in its own way, explanatory (Bechtel & Abrahamsen, 2005, p.426). In the usual course of events, we *begin* our attempts at mechanistic explanations with the preconception that the significant parts will be working parts of the whole. We might then overlook the epistemic explanatory value of HE, but

---

<sup>16</sup>For clear (footstamping) accusations of triviality against this sort of explanation (simply because it is not decompositional) see e.g. Allport 1924 & Hamlyn 1951.

it is still there.

Consider next the ontic view. I've argued that (early and ontic) New Mechanists are committed to the claim that a whole mechanism-for a phenomenon is co-individuated (in a metaphysical sense) with the working parts, activities, and interactions which compose it. The whole mechanism's  $\psi$ -ing is co-constitutive with the working parts' joint  $\varphi$ -ing. The working parts only exhibit the full list of  $\varphi$ -activities when they work together to produce the phenomenon of interest, and when they do so, the phenomenon is produced. (Likewise: an electrical field attaining equilibrium is co-constituted with a certain charge being attained at each point in the field.) Now the ontic theorist maintains that an ontic explanation is what causes, constitutes, or is otherwise responsible for the production of some phenomenon. If the phenomenon of interest is some behavior of the whole mechanism, then we can pursue a standard mechanistic explanation in terms of the working parts which constitute it. If the phenomenon of interest is some behavior of a working part, however, and *if that part's exhibiting that behavior is co-constitutive with the whole exhibiting its behavior*, then the behavior of the whole is equally well an ontic explanation of the behavior of the part. Because "constitutive relevance is symmetrical in a way that etiological (that is, causal) relevance is not," the whole's behavior constitutes, or is otherwise (i.e., non-causally) responsible for – it is an ontic explanation of – the behavior of the part (Craver, 2007, 153). This is not the sort of ontic explanation that ontic Mechanists have given pride of place, but it is nonetheless, by their own lights, an ontic explanation.<sup>17</sup>

Both camps, then, should recognize HE as a legitimate form of explanation, whether or not it is especially central to Mechanists' favored explana-

---

<sup>17</sup>HI, the ontic theorist would urge, is *why* HE holds, such that we can answer a why-question about the part by adverting to the whole. This was essentially Köhler's view. Uncovering the New Mechanists' commitment to HE thus addresses an issue which Fagan (2012, p.459; 2015) seemed to raise as an objection. If constitution is symmetrical, Fagan thinks it cannot account for the "bottom-up" or decompositional direction of mechanistic explanation: why shouldn't we also explain parts' behaviors by appeal to the whole mechanism's? Answer: we could, and we should not seek to rule out in advance that we could. We should simply restrict the notion of (mainstream) "mechanistic" explanation to the decompositional direction: mechanistic explanations in this sense are not holistic, though wherever they are available, a holistic explanation is as well.

tions.<sup>18</sup> Note: I have not claimed that the mechanists are *correct* in implicitly endorsing HI and HE. I have only sought to show that they are committed to the claims, and thus to the claim that mechanisms – if they are real – are material *Gestalten*.

### 3. Recent Attempts to Understand Constitutive Organization

Mechanisms' organization is clearly central to their explanatory value. Despite this, until recently Mechanists said little to clarify mechanisms' organization. Indeed, one might object that the applicability of the term “material *Gestalten*” to mechanisms simply trades on an under-specified conception of their organization. In this section I pursue two aims simultaneously. First, I emphasize that the burden of providing a philosophical analysis of the nature of mechanisms' organization is widely recognized to rest squarely on the Mechanists' shoulders: if the identity of mechanisms and *Gestalten* is cause for alarm, the complaint should be lodged with the Mechanists, not with me. Second, I examine a recent in-house debate amongst Mechanists, showing that there is little consensus regarding how Mechanists' might carry this burden. In §4 hereafter, I show that this in-house debate was anticipated decades ago by Merleau-Ponty in his claims against the Gestaltists, and draw upon his work to offer a diagnosis of the debate.

#### 3.1. Craver's Mutual Manipulability Criterion

One influential attempt to clarify mechanisms' organization was Craver's “mutual manipulability” criterion, introduced in §2.3 above. As I read him, Craver sought to clarify a *sufficient* criterion for concluding that some part of a mechanism is a working component, and is of constitutive relevance for the production of the phenomenon of interest. The account was explicitly offered as a novel adaptation of Woodward's (2002, 2003) manipulationist account of causal relevance, modified to address constitutive relevance.

Woodward's account, roughly, is that we can say an item X is causally relevant to an item Y in conditions C if there is an ideal intervention on X in C (i.e., an intervention *exclusively* on X which neither affects any causal intermediaries between X and Y, nor directly effects Y) which would change the value of Y (compare Craver 2007, pp.94-95). Or, still more roughly: X is

---

<sup>18</sup>Doing so would *increase* mechanisms' explanatory import, beyond decompositional *how* explanations.

causally relevant to Y if wiggling X (alone) would wiggle Y. Craver utilizes this account to clarify the norms of causal explanation he discerns in neuroscientific practice, particularly in the experimental strategies neuroscientists undertake to determine the causal relevance of entities in a mechanism (*ibid.*, pp. 98-104).

Some distinctive experiments that neuroscientists perform are what Craver calls *interlevel experiments*, including “bottom-up” and “top-down” experiments (see 2007, p.146ff. for details). Scientists use these experiments to determine whether a part of a mechanism is a *working* part which is of constitutive relevance to the production of a whole mechanism’s *explanandum* phenomenon (*ibid.*, p.144.). The inspiration for Craver’s mutual manipulability account comes from considering the practicalities of experimentation: “the close analogy between causal experiments [i.e., those performed to uncover causal relations] and interlevel experiments suggests that the manipulability account of etiological [or causal] relevance might provide a model for thinking about constitutive mechanistic relevance” (*ibid.*, p.152). *Mutual manipulability*, as we saw it in §2.3 above, relies on a bidirectional or symmetrical version of Woodward’s conception of manipulability. In short, if X is a part of Y, and if wiggling X (alone) wiggles Y, *and* if wiggling Y (alone) wiggles X, then Craver suggests we have sufficient reason to conclude that they are of constitutive relevance to each other. And conversely: where we can demonstrate a lack of mutual manipulability, we have sufficient reason to conclude there is a lack of constitutive relevance (*ibid.*, p.159). This, he suggests, is how neuroscientists actually do determine whether a putative part of a mechanism is a *working* part of it.

The mutual manipulability criterion was offered as a “starting point” for understanding mechanisms’ constitutive organization (Craver, 2007, 160). It has not proven to be a fixed starting point, and its merits are a subject of continued debate.

### 3.2. Organization, Post-Craver

In this sub-section I review the ongoing debate regarding how to understand mechanism’s constitutive organization. I discuss this debate in some detail for two reasons. First my aim in §4 below is to draw upon the work of Merleau-Ponty to offer a diagnosis of the debate, and it is crucial to my argument that the diagnosis does not rely on a simplified or straw-manned variant of the debate, but rather captures all its significant features. Second, it is likewise crucial to my argument and diagnosis that a sober view of the

debate reveals that ontic Mechanists – though they largely agree that constitutive relevance should not be understood causally – have reached little consensus regarding what their ontological commitment to organized mechanisms amounts to, or where that commitment comes from. This can only be shown with a detailed overview of the debate. Readers familiar with the debate might skip to §3.3 for an over-all summary, and selectively revisit this subsection as they see fit.

The continued debate about how to understand parts' constitutive relevance to mechanisms' organization can largely be understood as arising from a perceived tension between three claims,<sup>19</sup> all of which have been traced to Craver's work:

- C1. Woodward's account of ideal interventions (as stated in earlier work and as sketched in §3.1 above) offers a suitable ontological analysis of the relation of causal relevance.<sup>20</sup>
- C2. Mutual manipulability (as stated by Craver and as sketched in §3.1 above) offers an ontological analysis of the relation of constitutive relevance, through an application of Woodward's account of causal relevance.
- C3. The constitutive relations between levels of a mechanism (e.g., working part and whole) are not causal (e.g., because they are synchronic).

If C1 and C2 are true, then constitution is a (bidirectional) causal relation, and C3 is false. Yet upholding C3 seemed to be the intended motivation for the mutual manipulability criterion: what Craver sought was a model for thinking about constitutive relevance relations *as distinct from* causal relations. So if C1 and C3 impose constraints on an ontological analysis of constitutive relevance, then mutual manipulability seems not to be a coherent ontological analysis of constitutive relevance.

A major choice-point in the debate concerns C2. Some authors evade the tension by rejecting C2 as false. One way to do so is to view mutual manipulability as only an *epistemic* criterion for when it is apt to *believe* that a part is a working part. On this view, Craver's appropriation of the formal

---

<sup>19</sup>My exegesis in this subsection owes much to [Romero \(2015\)](#) and [Krickel \(2018\)](#).

<sup>20</sup>It should be noted that Woodward does not explicitly endorse this view: cf. [Woodward 2003](#), p.7. Still, this is a popular reading, invoked by a number of participants in the debate.

apparatus of Woodward's interventionism to express mutual manipulability does not provide an *ontological* analysis of what mechanistic constitution *is* (cf. Couch 2011, pp.381-382; Fagan 2012, p.459; Glennan 2017, p.44).<sup>21</sup> In that case, it is an open problem to go beyond mutual manipulability, by offering a genuine ontological analysis of constitutive relevance. Couch proposes appropriating a distinct analysis of causation to understand constitution. He draws upon Mackie's (1974) well-known analysis of a cause as an insufficient but nonredundant part of an unnecessary but sufficient (INUS) condition. On Couch's analysis, a working part of a mechanism is "an insufficient but nonredundant part of an unnecessary but sufficient mechanism" for a phenomenon (2011, p.384). Couch's account is often-cited, but rarely discussed. Glennan cryptically suggests that it offers only an epistemic criterion (2017, p.44, *fn.22*). (As I discuss further below, it is difficult to locate an alternative ontological analysis in Glennan's work.) Krickel dispenses with Couch by branding his a "regularity account" of constitutive relevance – 2018, *fn.3*, a characterization that seems to have started with Harbecke (2015) – and sets it aside to pursue an interventionist account.

I suggest that Couch's account may have been poorly understood precisely because it flirts with HI and HE. Couch insists that his is a truly ontological account, which "provides an explanation of what the [constitutive] relevance relation is" (Couch, 2011, p.387). The relation he has described is "being an insufficient but nonredundant part of an unnecessary but sufficient mechanism." One could have difficulty seeing that this helps us understand constitutive relevance. A sufficient mechanism is one that is sufficient to produce the phenomenon of interest. It does this because its relevant parts work together to jointly produce the phenomenon. So much is largely a matter of definition. Now comes the key question: what is the nature of those parts which are the relevant parts? Couch's answer: they're the ones that, when

---

<sup>21</sup>As a matter of interpreting Craver, I would reject C2 along these lines, and find little support for it in his work. Baumgartner & Gebharder suggest Craver offers mutual manipulability as a "necessary and sufficient condition" for the occurrence of relations of constitutive relevance, and claim that unspecified "textual evidence [in Craver] clearly contradicts" the view that he offers only an epistemic criterion (2015, p.738, *fn.4*). While Craver's offering a "merely" epistemic criterion is somewhat discordant with the bulk of his text, this seems clearly to be where he winds up: cf. Craver 2007, bottom of p.159. It is in any case more charitable to suppose this is his view, if the alternative is to accuse him of conflating constitution and causation precisely where he aims to distinguish them.



non-redundantly working together, compose a mechanism which is sufficient to produce the phenomenon (though each of them individually is insufficient). We might ask again: but what is the nature of the parts which *makes* them fit for this role? Couch's account seems to offer no general answer to this question. It's nothing about the *parts* themselves, considered in *isolation*, which makes them constitutively relevant. Rather it is their participation (of the right sort) in a whole (of the right sort).

Other authors do not deny C2 – they hold that Craver's mutual manipulability criterion *is* an ontological analysis of mechanistic constitution (Leuridan 2011, §5; Fagan 2012, pp.458-459; Baumgartner & Gebharder 2015; Romero 2015, §2.3; Gallagher 2018, p.5). The question then is whether it is suitable as it stands. Leuridan is alone in suggesting that we abandon C3: so-called “constitutive” relations (as *per* C2) turn out to be bidirectional causal relations (as *per* C1) (Leuridan, 2011, p.424). On this view, the concept of causal relevance turns out to be ambiguous between “intralevel etiological relevance (or causal relevance *sensu strictu*) and interlevel relevance” (*ibid.*, p.424). Leuridan is aware that this does not fully resolve the issue: “There *is* a difference between interlevel relations [i.e., part-whole relations, in general] and bidirectional etiological relations... But my proposal makes clear that the difference in question is not as strict as Craver and Bechtel claim it to be and that as yet we lack the means to spell it out” (*ibid.*, p.424).

Most authors seek to uphold C3: they regard the distinction between causal relevance and constitutive relevance to be non-negotiable. One family of views along these lines invoke the idea of so-called “fat-handed” interventions. Romero (2015) suggested that we should not understand interlevel experiments as involving anything like Woodward's ideal interventions, in which *only* a part is manipulated and the result is a change in the whole, or *only* the whole is manipulated and the result is a change in the part. Rather, both are “fat-handed” (somewhat un-selective) manipulations: a manipulation of a working part simultaneously manipulates the whole, and a manipulation of the whole simultaneously manipulates at least one working part. This is to be expected if (as I have argued is true for New Mechanists) the two are regarded as co-constitutive. Romero adopts a stronger ontological claim of identity here, saying “a mechanism's activity [i.e., the *explanandum* phenomenon] and its components arranged and working in the

right way are the same physical event” (2015, 3746).<sup>22</sup> Ultimately, Romero does not offer a worked out ontological interpretation of constitutive relevance (cf. 2015, p.3749, p.3750). Since the fat-handed interventions involved in mutual manipulability are not ideal interventions, Romero can afford to be agnostic regarding C1 (*ibid.*, p.3749). Since fat-handed interventions in interlevel experiments produce *simultaneous* changes in part and whole, the relation between part and whole should not be considered a causal relation (*ibid.*, 3747). In this way C3 is upheld.

Baumgartner & Gebhartner (2015) pursue a similar strategy of appealing to fat-handed interventions, but also incorporate recent amendments that Woodward (2015) has made to his interventionist account. They can thus be read as rejecting C1, since (following Woodward himself) they adopt a revised conception of ideal interventions. They endorse C2, but reject mutual manipulability as an acceptable ontological account. They seek to uphold C3. Notably, however, they do not provide an alternative ontological analysis of constitution: their aim is rather to clarify how fat-handed interventions, understood in the right way, provide “abductive evidence for constitutive dependences” (*ibid.*, p.753). Theirs is a novel epistemic criterion, pursued because they think that mutual manipulability is incoherent – both as an epistemic and as an ontological criterion (*ibid.*, §3). In a follow-up paper, Baumgartner & Casini (2017) have advanced an ontological account of constitutive relevance. In short, constitutive relevance is that relation between parts and whole in virtue of which the two cannot be causally decoupled, such that fat-handedness is a necessary consequence: the impossibility of mutual manipulability through ideal interventions becomes the defining mark of constitutive relevance, and thus fat-handedness can serve as an epistemic standard for inferring constitutive relevance.

Fat-handedness is sometimes viewed as descriptively inadequate to characterize scientists’ interlevel experiments. Thus another response to the tension between C1-C3 is to try to maintain the spirit of all three claims, with suitable tweaks. This is the strategy pursued by Krickel (2018). Krickel takes C3 to be non-negotiable. She follows Baumgartner & Gebhartner in accommodating Woodward’s recent revisions, altering the conception of ideal intervention, and thus rejecting the letter of C1. In her view, the real solution depends on clarifying the *relata* that stand in relations of causal and

---

<sup>22</sup>Compare Salmon’s remarks in *fn.14* above.

constitutive relevance. In the sketches above, we supposed that a relation of constitutive relevance held between (a) some working part(s) and (b) a whole mechanism producing an *explanandum* phenomenon. Krickel's account relies on a more fine-grained view of the phenomenon as composed of a number of temporal parts, each of which occurs as phases of the whole mechanism's prolonged activity of exhibiting the phenomenon. If we get fine-grained in this way, the interventionist can say that there is a *single* entity (the mechanism, constituted as it is) whose spatial and temporal moments are mutually manipulable: (i) some working part(s), working at  $t_1$ , causally contributes to the production of a temporal part of the *explanandum* phenomenon at  $t_2$ , such that ideally intervening on the former at  $t_1$  would causally influence the latter at  $t_2$ , and (ii) this same temporal part of the phenomenon causally contributes to changes in the same working part(s) at  $t_3$ , such that ideally intervening on the former at  $t_2$  would causally influence the latter at  $t_3$ . This mutual manipulability does not, however, render *constitution* relations causal: the causal relations all hold between fine-grained spatio-temporal moments of *one* entity (the organized mechanism). So the causal relations which underlie mutual manipulability – though they *are* interlevel causal relations – are all diachronic, and are distinct from the synchronic constitution relations between part and whole. Yet if we emphasize the “temporal heterogeneity” of the phenomenon, we can always (ideally) directly manipulate one and not the other, and thus preserve the possibility of ideal interventions, avoiding limitation to fat-handed interventions.<sup>23</sup> The ontological analysis of constitutive relevance is then that a phenomenon and working-part are of constitutive relevance to each other iff the foregoing causal relations of mutual manipulability hold between their fine-grained spatiotemporal moments.<sup>24</sup>

There is one final view worth noting, though it is difficult to cast in terms of C1-C3. Fagan (2012, 2015) has proposed to understand the organization of mechanisms on analogy with Bratman's (1992) analysis of joint or socially-

---

<sup>23</sup>It is not clear to me how Krickel's account avoids the claim that interventions on working parts or the phenomenon are – considered *at a moment* – fat-handed: they must be so, if the two are co-constitutive. Her account does not analyze mutual manipulability or constitutive relevance *as* fat-handed interventions, but I do not see how one could intervene to alter some working part's working, at  $t_1$ , without simultaneously altering the temporal part of the phenomenon occurring at  $t_1$  (and *vice versa*) on her account.

<sup>24</sup>Gallagher (2018) cites Krickel's work as inspiration for an enactivist account of “causal constitution” in his enactivist theory of cognition.

shared intentional action. To borrow her illustration: you and I might each have an individual intention to walk, and might wind up walking in the same direction; or alternatively, we might have a *shared* intention, and might walk *together* in genuinely social action. “By analogy with the social action case, joint activities in biological contexts can be analyzed as follows:

- (**J**) Components  $x_1$  and  $x_2$  jointly  $\varphi$  if and only if:
- (i)  $x_1$  has properties that mesh with those of  $x_2$  and vice versa,
  - (ii)  $x_1$  and  $x_2$  form a complex  $x_1x_2$  in virtue of their meshing properties,
  - (iii) complex  $x_1x_2$   $\varphi$ 's and
  - (iv) uncomplexed  $x_1$  and  $x_2$  do not  $\varphi$ ” (Fagan, 2015, p.75).

The notion of “meshing” is meant to capture those features of working components which enable them to work together. In the social case, this would crucially include shared intentions. In purely biological contexts, “meshing” is a placeholder for some non-intentional something which can analogously unite the activities of biological entities into a joint activity of  $\varphi$ -ing – “the notion of meshing properties here needs further clarification, obviously” (*ibid.*, p.74). Applying (**J**) to mechanisms, we claim that a number of meshing parts together constitute the mechanism itself, and their joint activity constitutes the mechanism’s exhibiting the *explanandum* phenomenon. We can explain the *explanandum* phenomenon on the basis of these constitution relations. The explanation is intended to be “bottom-up” or decompositional: “the meshing properties required for interaction are the basis of the entire explanation” (*ibid.*, p.74 – contrast the discussion of Couch above, and see also *fn.17* above). Glennan’s most recent account is similar in spirit. Glennan suggests that working parts have “affinities... capacities to ‘seek out’ and engage in interactions with other parts,” and that “All mechanistic organization depends to some degree upon the existence of affinities – as parts must have capacities to interact with other parts” (2017, p.121). Glennan offers this as a kind of extension of the concept of affinities between chemical reactants, yet does not offer a very clear ontological analysis of affinities in general. Much of the work done by Glennan’s concept of affinities is intuitive, relying on an analogy with social affinities between persons – note the scare-quoted use of intentional or anthropomorphic language to describe parts’ affinities.

Fagan’s formal apparatus looks like a step beyond an intuitive grasp on social action, but does not get us very far. Fagan sees part of the value of

(**J**) as “making the *organization* of interacting components central to mechanisms,” but – see §2.3 above – this does not seem to me to be a novel development (2015, p.75).<sup>25</sup> (**J**), applied to mechanisms, appears to be an alternative description of the basic features of mechanistic organization: the constitutively relevant working parts of a mechanism are held to be the ones that work together (because they “mesh”) to constitute the mechanism’s production of the *explanandum* phenomenon. This is “active organization” as described in §2.3 above. The debate I’ve been reviewing has concerned whether and how we can provide an ontological analysis of constitutive relevance. In Fagan’s account, we swap the problem of constitutive relevance for the problem of analyzing “meshing,” yet we are given no ontological analysis of it. The analogy to social intentional action seems especially inapt as a starting point: what non-intentional feature of biological parts could be relevantly analogous to shared intentions? The envisioned ontological analysis would seem to be capable of succeeding only where it enforces a breakdown of the analogy which is offered to promote it.<sup>26</sup>

### 3.3. The Present Debate, in Sum

Let’s take stock. The ongoing debate about constitutive relevance is largely a series of rival attempts to understand constitution non-causally. The debate was initiated by a reading of Craver (2007) as ontologically analyzing constitutive relevance as mutual manipulability – i.e., as bidirectional causation, in terms of Woodward’s interventionism. Few suggest that constitutive relevance should be understood causally in this way, and those that do admit that this is only a starting point in an ontological analysis of constitutive relevance (Leuridan, 2011). Most pursue a non-causal analysis of constitutive relevance – though there is a tendency toward renewed attempts to adapt pre-existing analyses of causation (Couch, 2011, Harbecke, 2015), and also a tendency toward metaphors which do not seem to advance our

---

<sup>25</sup>I confess I find much of Fagan’s motivation difficult to discern, and suspect we must have very different readings of, e.g., Craver. Fagan (2012, 2015) offers extended arguments, purportedly against Craver (2007), for distinguishing causal relations (and explanations) from constitutive relations (and explanations) – which I think Craver had intended to do from the outset.

<sup>26</sup>More recently, Fagan sketches the relations of meshing which make mechanisms’ parts “fit together” as “involv[ing] causal, spatiotemporal, and mutual binding relations” (2016, p.1096). Compare the scanty remarks on mechanism’s organization offered by MDC, discussed at the start of §2.3 above.

understanding of the ontology of constitutive relevance (Fagan, 2015, Glennan, 2017). Some analyses abandon the possibility of ideal interventions on parts and wholes (and thus mutual manipulability), allowing only fat-handed interventions on working parts and wholes (Romero, 2015). This route risks abandoning an ontological analysis of constitutive relevance, offering a merely epistemic criterion (Baumgartner & Gebharter, 2015), unless we turn the tables and analyze constitutive relevance as what necessitates fat-handedness (Baumgartner & Casini, 2017). If we get very fine-grained in our analysis, we might uphold mutual manipulability while also distinguishing causal relations from constitutive relations (Krickel, 2018). Some accounts seem to enforce a decompositional approach, in which properties of the parts, considered in isolation, are ultimately what explain constitutive relevance (Fagan, 2015, Glennan, 2017), whereas others may take a more holistic approach (Couch, 2011).

In short, as Kaiser & Krickel have summarized it recently: “the notion of mechanistic constitution has gained increasing attention... But so far no approach has been developed that has found common acceptance” (2016, p.746, *fn.3*). According to the ontic Mechanists, Mechanistic explanation commits us ontologically to *something* – they do not agree what – called “constitutive relevance,” and leading accounts of Mechanistic explanation must be *supplemented* – they do not agree how – to provide some account of constitutive relevance, and of the ontology of organized mechanisms.

#### 4. A Merleau-Pontian Diagnosis and Null Hypothesis

In this section my aim is to show that Merleau-Ponty’s claims about *Gestalten* anticipate and diagnose much of the debate I’ve just reviewed. I first (§4.1) introduce the reader to Maurice Merleau-Ponty’s engagement with the Gestaltists, particularly in *La Structure du comportement* (“*SC*”) (1963/1943).<sup>27</sup> In Merleau-Ponty’s view, *Gestalten* are not mind-independent entities, and they possess their unique organization only as mind-dependent objects of (perceptual) consciousness. As a result, their organization cannot be adequately analyzed in a causal-objective framework. I underscore that Merleau-Ponty’s own arguments against the Gestaltists rely on premises that will likely not win wide acceptance. However, I next (§4.2) show how his ar-

---

<sup>27</sup>For further elaboration of the reading of *SC* I invoke here, see Sheredos (2017).

guments might plausibly be extended, specifically to resist ontological commitment to the material *Gestalten* which the New Mechanists have sought to characterize. In short, we could adopt a strong epistemic (as opposed to ontic) conception of mechanisms, treating them as idealized theoretical entities, as Bechtel (2015) has urged. Finally (§4.3) I argue that this view would offer a diagnosis of the Mechanists' failure to reach consensus on an objective, ontological analysis of constitutive relevance: there is no such analysis to be had. I thus suggest that a Merleau-Pontian conception of *Gestalten* offers us a null hypothesis: until an adequate, objective analysis of mechanism's organization is provided (from an ontic standpoint) we may more modestly view their organization as mind-dependent (epistemic).

#### 4.1. Merleau-Ponty on Gestalt Theory

In Merleau-Ponty's view, it is phenomena of *order* which centrally inspired the Gestaltists to invoke the category of "form" or "structure" or "*Gestalt*." In doing so, the Gestaltists opposed a classical conception of scientific realism, which viewed the world as ultimately composed of physicalistic, ontic "atoms," which stand in unidirectional causal relations. Each atom's capabilities depend ultimately on its intrinsic properties. In the classical view, order does not exist, and there are no wholes which possess any distinctive features or capabilities that are not reducible to those of juxtaposed atoms. The classical view is thus incapable of recognizing – let alone properly explaining – orderly phenomena. A central example in Merleau-Ponty's discussion concerns the orderly (active and adaptive) character of intelligent behavior. The classical view would require us to treat all behavior as arising through reflex-elicitations, "the action of a defined physical or chemical agent on a locally defined receptor which evokes a defined response by means of a defined pathway" (*SC*, pp.8-9/6-7).<sup>28</sup>

With regard to behavior, the basic issue is that the classical view, as sketched above, can provide no account of how a number of reflexes might be blended or combined to constitute an instance of orderly behavior. A popular "solution" had been to divide the nervous system into a hierarchy, and to posit some higher mechanism to regulate (inhibit or control) reflexes. Merleau-Ponty is skeptical of this route, for two related reasons. First, he is

---

<sup>28</sup>Citations provide the page-number in Fisher's (1963) English translation, followed by the page number in the later French editions (those with a preface by de Waelhens).



opposed to any hierarchical conception of the nervous system (*SC*, p.22/19, citing [Goldstein 1934](#)). Thus he remarks:

“[T]he classical conception is maintained [by positing regulation] only if the regulation is *localized* in certain devices comparable to reflex arcs. But it does not seem to be exclusively bound up with cerebral activity... nor does it seem explicable moreover in each place by automatic devices of association or disjunction... Depending on the case, each part of the nervous system can in turn appear to be inhibiting or inhibited...

In the final analysis, inhibition and control do not explain nerve functioning. They themselves presuppose a process which regulates their distribution” (*SC*, pp.31-32/31, my emph.).

Merleau-Ponty takes the neuroscientific data he reviews (and which I omit) to suggest there is no hierarchical organization in the nervous system, no reflex-like “control circuit” which *automatically* regulates lesser circuits. This is the first reason why Merleau-Ponty thinks the classical view will not work. The concept of the reflex failed to describe the nervous system in the first attempt, since it could not account for orderly behavior. The classical theorist then re-posed the reflex “further in,” in the guise of an automatic control circuit. If this too fails, the classical theorist must again scramble to actually *apply* the concept which is intended to get their explanation off the ground. Merleau-Ponty quips that this strategy will

“have to be reinitiated indefinitely; and the solution will always be deferred, never furnished, until the moment when a principle which *constitutes* the order [*un principe qui constitue l'ordre*] instead of undergoing it has been introduced...” (*SC*, p.33/33).

This latter claim helps us see the second, deeper reason why Merleau-Ponty thinks the classical view will not work. It is because a classical view simply cannot account for orderly phenomena (forms, structures, *Gestalten*). Order cannot be *causally produced* in the classical view: the output of any causal process will always just be a new juxtaposition of atoms. As a result, the puzzle is not to find the right *kind* of regulation (inhibition, control, etc.) to *cause* order. Instead, as Merleau-Ponty phrases it: “the problem of order has no meaning if we make it a second problem of causality” (*SC*, p.50/53). Merleau-Ponty thinks we need no new *kind* or *concept* of causation in order

to explain order – “causation” is not ambiguous, and has no secret sense we must work out. (Contrast Leuridan’s (2011) claims, reviewed on p.23 above).<sup>29</sup> We must rather seek to understand a wholly novel, acausal kind of co-dependency involved in the *constitution* of order (*SC*, p.154/167).

This is not Merleau-Ponty’s full view: we have not yet clarified his disagreement with the Gestaltists. But we can already spell out how Merleau-Ponty’s view of *Gestalten* anticipates much of the dispute regarding mechanism’s organization. The notion of order we’ve been discussing is precisely the sort of organization which renders some systems *Gestalten* rather than mere and-summations. As Merleau-Ponty puts it – flirting with the idea of both merely physical and biological *Gestalten* – “in a soap bubble as in an organism, what happens at each point is determined by what happens in all the others. But this is the definition of order” (*SC*, p.131/141-142). Just this sort of interdependence of working parts is what the New Mechanists have called a mechanisms’ “active organization” (§2.3 above). Merleau-Ponty’s insistence that the constitution of order cannot be explained by (classical) causation anticipates the widespread insistence by Mechanists that constitutive relevance should not be analyzed as a causal relation (§3.2 above). In addition, Merleau-Ponty’s criticism against the classical atomist echoes some other claims by New Mechanists. As discussed above, both Fagan (2012, 2015) and Glennan (2017) seem to endorse a reductive or compositional approach to constitutive relevance, according to which it is ultimately certain features that parts possess in isolation (e.g., their “meshing properties” or “affinities”) which will enable us to analyze constitutive relevance. Merleau-Ponty makes clear a basic constraint on any such approach: we will not be able to account for the orderly organization of mechanisms in terms of the parts unless we have an ontologically rich conception of the parts, which invests them with some principle which enables them to *constitute* the order

---

<sup>29</sup>As I have argued elsewhere (Sheredos, 2017) there has been persistent confusion in the secondary literature on Merleau-Ponty on this point. A popular “enactivist” reading of Merleau-Ponty supposes that he supports the claim that “a dynamical gestalt [is] composed of processes that unfold over time, and [is] characterized by recursive reciprocal causal relations” (Gallagher, 2017, p.10). The New Mechanists’ debate over constitutive relevance – especially Krickel (2018) – has likewise been pressed into service to clarify the enactivist’s core thesis of “causal constitution” (Gallagher, 2018). By my lights, the causal-enactivist conception is at odds with both the aims of New Mechanists, and Merleau-Ponty’s own view – at least as this is expressed in *SC*.

which is our *explanandum*. The classical view, with its limited ontological resources, could not make such a view workable.

We can go further in drawing connections by considering one of Merleau-Ponty's central claims about the relationship between an organism's "reactions" and its surrounding world (*milieu* or "situation"). He understands these as two moments of the indivisible *Gestalt* which is perceived behavior. This bears emphasis: for Merleau-Ponty, an organism's behavior (as we perceive it) is not "a thing in-itself (*en soi*) which would exist, *partes extra partes*, in the nervous system or in the body; rather... behavior [is] an embodied dialectic which radiates over a milieu immanent to it (*SC*, pp.161/174). Both the organism and its *milieu* are "parts" of one whole perceived *Gestalt*: behavior is not localized to a classical, physical body.<sup>30</sup> The acting organism's body, in this sense, is not a classical juxtaposition of atoms, but rather an object of perception: what Merleau-Ponty calls a "phenomenal body" (*SC*, p.156/169). With this in mind, consider the following claim about behavior:

"Situation and reaction are linked internally by their common participation in a structure in which the mode of activity proper to the organism is expressed. Hence they cannot be placed one after the other as cause and effect: they are two moments of a circular process [*d'un processus circulaire*]... If behavior is a 'form,' one cannot even designate in it that which depends on each one of the internal and external conditions taken separately, since their variations will be expressed in the form by a global and indivisible effect" (*SC*, pp.130-131/140-141).

The claim made here, in reference to behavior, is indicative of Merleau-Ponty's conception of all *Gestalten*. The claim is that any variation in the "parts" of a *Gestalt* is co-constitutive with an alteration of the whole form. This is tantamount to the core notion of "fat-handedness," and the claim that manipulating any working part of a mechanism (*Gestalt*) co-constitutively manipulates the whole (§3.2 above). As is clear from the passage, Merleau-Ponty likewise insists that these relations are not causal.

<sup>30</sup>This *Gestalt* is thus also temporally extended, like a melody – a central example of the Gestaltists besides the more clearly spatial cases, such as ambiguous figures.

These connections, I hope, make it clear that Merleau-Ponty's decades-old claims about *Gestalten* effectively summarize the basic "tension" (§3.2 above) which has driven the New Mechanists' debate over constitutive relevance. Merleau-Ponty's concern, like that of most New Mechanists, was how to understand the constitution of order or organization in a non-causal way. Here Merleau-Ponty's disagreement with the Gestaltists comes to the fore.

The Gestaltists reified material *Gestalten*, abandoning a classical worldly ontology of physicalistic atoms in favor of an ontology of *Gestalten*. This, in Merleau-Ponty's view, was an error: *Gestalten* were rather only to be understood as mind-dependent phenomena which we experience in perceptual consciousness. He offered several arguments for this view. One argument took the Gestaltists to task for failing to offer an adequate ontological analysis of material *Gestalten*. Köhler's (1920) discussion of physical (electrostatic) *Gestalten* was a target here. In Merleau-Ponty's view, Köhler's reliance on intuitive metaphors was telling:

"That in the final analysis form [i.e., *Gestalten*] cannot be defined in terms of reality but in terms of knowledge, not as a thing of the physical world but as a perceived whole, is explicitly recognized by Koehler when he writes that the order in a form "rests... on the fact that each local event, one could almost say, 'dynamically knows' the others."\* It is not an accident that, in order to express this presence of each moment to the other, Koehler comes up with the term 'knowledge.' A unity of this type can be found only in an object of knowledge... This unity is the unity of perceived objects. [Thus e.g.] A colored circle which I look at is completely modified in its physiognomy by an irregularity which removes something of its circular character and makes it an imperfect circle" (*SC*, pp.143-144/155, footnoting Köhler 1920 p.180 at the asterisk).

Merleau-Ponty here makes three claims.

- M1. If this kind of anthropomorphic analogy is the best we can offer to analyze *Gestalten*, we do not have an adequate realistic, ontological analysis of them;
- M2. the *reason why* we are led to invoke this sort of analogy is that we are (dimly, but correctly) implicitly regarding *Gestalten* as mind-dependent phenomena we experience in perceptual consciousness, because

M3. the sort of unity which is imperfectly expressed by these anthropomorphic analogies is solely possessed by perceived wholes.

I take it most of us would grant M1. Merleau-Ponty's point here is relevant to some New Mechanists: see the discussion in §3.2 above of Fagan's (2015) conception of "jointness," on analogy with social action, and Glennan's (2017) claim that parts have "affinities" to "seek out" other parts to interact with. These analogies are plausibly no substitute for a realistic, ontological analysis of constitutive relevance and mechanistic organization.

M2 and M3 are more contentious, and they connect with Merleau-Ponty's second argument against the Gestaltists. I provide a sketch of it here, and will elaborate further in §4.2 below. Merleau-Ponty's key contention was that we do not discover material *Gestalten* in the advancement of science. We begin with an intuitive apprehension of perceptual-phenomenal *Gestalten* – for example, a perception of a phenomenal body exhibiting behavior. Our intuitive grasp on the phenomenon, through perception, pre-dates our hard-won scientific understanding of it, and the whole purpose of scientific endeavor is to make the perceived world more precisely intelligible (*SC*, p.145/157). While we can come to understand a perceptual-phenomenal *Gestalt* more precisely through scientific research, we do not discover that entities in the mind-independent world are *really*, in themselves, material *Gestalten*. Both universal laws of nature (*SC*, p.139/150) and mathematical formalisms (*SC*, p.143/155) might have some applicability to perceived phenomena, enabling us to understand them more precisely, but neither licenses us in reifying *Gestalten* as mind-independent entities. Because the aim of scientific investigation is to make phenomena *intelligible*, Merleau-Ponty thinks it is incoherent to regard them as mind-independent things in themselves: a *Gestalt* should be regarded as a moment of a dialectic, not as an "element of the world" which exists independent of consciousness (*SC*, pp.142/153). This division between perceived phenomena and elements of the world is the root of Merleau-Ponty's claim M3, and this in turn supports his claim M2.

Many contemporary thinkers would likely not uphold Merleau-Ponty's division between elements of the world and perceived phenomena. Moreover, most Mechanists would likely not agree that our starting-point for mechanistic explanation is a simple *perception* at all: prototypical cases in the New Mechanistic literature in philosophy of biology are microscopic mechanisms and *explanandum* phenomena (e.g., within single cells). As a result, Merleau-Ponty's own arguments against material *Gestalten* will not win

many adherents. However, I believe we can see our way to a plausible extension of Merleau-Ponty's argument. The core of Merleau-Ponty's argument against the Gestaltists is that the theoretical understanding of the world science provides does not reveal any mind-independent entities in the world as having the organization of *Gestalten*. The Gestaltists had themselves sometimes suggested this. In his analysis of electrostatic *Gestalten*, Köhler was forced to admit that no mathematical formula on its own dictates that its intended referent is a material *Gestalt*: the same quantitative apparatus can always be used in a way that leaves out the qualitative characteristics of a *Gestalt* – for example, arbitrary values could be assigned to any formula (*SC*, p.141/152). Köhler thus advocated devising some new mathematical sign which, appended to a given formula, would indicate that it should be given a qualitative interpretation as representing a material *Gestalt* (Köhler, 1920, p.105). The key issue here – the issue of whether and how our theoretical understanding of some system (e.g., as represented in a mathematical formula) might justify the reification of it as an organized whole – can be divorced from Merleau-Ponty's specific commitment to M2 and M3. In fact, a similar view has already been considered by some New Mechanists.

#### 4.2. Extending Merleau-Ponty's view

Bechtel (2015) has argued that biological mechanisms (the mainstay of Mechanistic explanation) are typically “scale-free,” meaning not only that there is no well-delineated temporal window in which activities occur, but also that there is no well-delineated spatial region in which the mechanism can be said to exist. He proposes that the “mechanisms” (and their parts, and their activities) which are hypothesized to exist in offering a Mechanistic explanation should be viewed as idealizations, where an idealization “involves the introduction of simplifying falsehoods” – e.g., the falsehood of discrete spatiotemporal boundaries (*ibid.*, p.85). A Mechanistic explanation can only be accurate to a “first approximation,” and the idealizations involved always make it the case that the hypothesized mechanism “fails to give a fully correct account of the phenomenon occurring *in nature*” (*ibid.*, p.85). It is worth repeating in full the take-home message Bechtel offers readers:

“[T]he mechanisms proposed are posits of the scientists developing the explanation. They do not exist in nature as well delineated entities. The goal of mechanistic explanation is not to represent mechanisms as they exist independently of scientists.

Rather, it is to show what phenomena, the parts and operations selected by the scientists, operating in the time-window they consider, can largely account for. While this may limit the aspirations of both scientists pursuing mechanistic explanations and philosophers characterizing their project, it does not challenge the value of pursuing mechanistic explanation and in the process idealizing mechanisms by delineating boundaries that do not exist in nature” (*ibid.*, p.86.)

Bechtel’s is an especially robust variant of an epistemic approach to Mechanistic explanation. The aim of scientific explanation, on this view, is to make phenomena more precisely intelligible. As seen in the passage above, this need not require uncovering the way the world really is. The ontic approach to explanation, by contrast, maintains that epistemic explanation is parasitic upon identifying so-called “ontic explanations” – the things in the world which cause, constitute, or are otherwise responsible for phenomena. The scale-freeness of mechanisms makes things difficult for the ontic camp: if we cannot *identify* the spatiotemporal location of mechanisms, it is difficult to (a) incur ontological commitment to them as “ontic explanations,” and (b) demand such ontological commitment as a success-condition of epistemic explanation.

Bechtel’s argument has two *foci*. (1) The number and character of interactions between mechanisms and their environments makes it difficult to identify the mechanism’s spatial boundaries, such that any precise hypothesis about its boundaries will be an idealization (2015, §3). (2) Following Marom (2010), the time-scales of mechanistic processes (e.g., parts’ activities, or the whole mechanism’s  $\psi$ -ing) are not characteristically well-defined, such that making any claim to the effect that these events occur within a well-defined temporal window will involve idealization (2015, §5).

I suggest that Merleau-Ponty’s basic claim against the *Gestaltists* may be understood along similar lines. His claim is that our scientific understanding of the world always involves idealization. Specifically, the theoretical division of perceived *Gestalten* into “parts” which may be to some extent understood in isolation (e.g., they can be represented using distinct variables in mathematical formulae) is an idealization, involving a falsehood of decomposability or aggregation. This idealized understanding is not nothing, but once we have attained it, it can offer no grounds for reifying material *Gestalten*. (Recall that here, Köhler suggesting appending some novel symbol to



mathematical formulas to enforce a qualitative interpretation of them as representing a *Gestalt*, as if to say: “and *this* simplifying idealization represents a real, material *Gestalt* – though nothing in it shows that it does.”)

We might disagree with Merleau-Ponty’s own specific view that the starting point for scientific investigation is always perception, and that *only* perceived wholes exhibit *Gestalt* structure. This does not seem to diminish his main point against the Gestaltists: if the Gestaltists implicitly pursue analytical decompositions of some system which (by their own lights) ought to be understood holistically, then they labor under idealization. The idealized understanding they acquire is not nothing, but they cannot simply append an endorsement of holism as an *afterthought*, reifying organization on a whim, if their decompositional strategy has done nothing to investigate it. If scientific investigation can offer only an idealized understanding of allegedly material *Gestalten* which *does not itself capture* their holistic character, then it cannot on its own incur ontological commitment to the mind-independent existence of material *Gestalten*.

If the Gestaltists admit they they have *presupposed* from the outset that the system of study is an organized whole, then some story must be told about where that presupposition comes from. Here Merleau-Ponty’s answer is perception, and his claim is precisely that the Gestaltists have misunderstood perceptual-phenomenal *Gestalten*, which are mind-dependent phenomena of naïve perception, and have effectively reified them as mind-independent material *Gestalten* (*SC*, p.144/156). Setting Merleau-Ponty’s appeal to perception aside, the core problem is this: if the Gestaltists have *presupposed* that the systems of study are organized wholes, they must then admit that a scientific understanding of that system does not offer us a means to *discover* material *Gestalten* at all.

#### 4.3. Merleau-Ponty’s Implicit Critique of the New Mechanists

With all this in place, Merleau-Ponty’s implicit critique of many New Mechanists is fairly straightforward: it is a special application of his core critique of the Gestaltists, and what I shall say in this subsection to spell it out closely mirrors what I said in the last.

On the one hand, New Mechanists have been championing a decompositional style of explanation, which labors under the idealization (or heuristic) that mechanisms’ parts can be studied in isolation. On the other hand, New Mechanists have sought to append an endorsement of holism (in their

claims about mechanisms' organization and their parts' constitutive relevance). These are not strictly incompatible, but they are disjoint: the decompositional strategy does nothing to clarify whence an endorsement of holism should arise. If we do indeed take this second step then, as I argued in §2, we embrace ontological commitment to material *Gestalten*. Yet, as I reviewed in §§3.2 & 3.3, no ontological analysis of mechanisms' organization has attained consensus. At present, many New Mechanists are promoting ontological commitment to certain holistic, organizational features of mechanisms, when (a) the core strategies of Mechanistic explanation offer us little to understand those features, and (b) little has been said to even enable us to *identify* the features to which ontological commitment is encouraged. Mechanists cannot simply reify organization as an afterthought, on a whim. If Mechanistic explanation can offer only an idealized understanding of those alleged ontic structures called mechanisms which *does not capture* their holistic character, then it simply cannot incur ontological commitment to the mind-independent existence of organized mechanisms.

If Mechanists admit that they have *presupposed* from the outset that mechanisms are organized wholes, then some story must be told about where that presupposition comes from. Minimally, Mechanists must then admit that mainstream Mechanistic explanation has not offered us a means to *discover* mechanisms' mind-independent organization at all.

It is possible that future work might resolve these issues. An ontological analysis of mechanistic organization might be proposed which quickly gains favor, and which makes clear how scientists' research strategies ("interlevel experiments") enable them to *discover* organized mechanisms. Yet caution is suggested by the wide variety of approaches toward constitutive relevance that have been taken up in the short time since it became a focus of inquiry. Understanding Merleau-Ponty's critique of the Gestaltists suggests a conservative null hypothesis: until such time as a suitable, realistic, ontological analysis of constitutive relevance and mechanistic organization is offered which addresses these concerns, it is more modest to suppose that mechanisms, *qua* organized unities, need not really exist in themselves, in a mind-independent nature. We instead begin from an *assumption* that mechanisms are organized wholes.

A story must then be told – *by Mechanists* – about where that assumption comes from. But on this view, that assumption need not be discharged by some empirical demonstration that there exist ontic structures called mechanisms which, in themselves, possess some unique organization. The assump-

tion rather does its work as part of an epistemic strategy of explanation: it specifies in advance some constraints on how we seek to understand phenomena when we pursue a Mechanistic explanation. On this view, the debate over constitutive relevance amounts to some first forays into elaborating a framework for how we might make mechanisms' (presupposed) organization more precisely intelligible. Whether or not such a framework is successfully clarified, our pre-scientific apprehension of mechanisms' organization will always pre-date our idealized scientific understanding of it.

## References

- Allport, F. H. (1924). The group fallacy in relation to social science. *American Journal of Sociology*, 29(6), 688–706.
- Audi, R. (1999). The cambridge dictionary of philosophy.
- Baumgartner, M. & Casini, L. (2017). An abductive theory of constitution. *Philosophy of Science*, 84(2), 214–233.
- Baumgartner, M. & Gebharder, A. (2015). Constitutive relevance, mutual manipulability, and fat-handedness. *The British Journal for the Philosophy of Science*, 67(3), 731–756.
- Bechtel, W. (2015). Can mechanistic explanation be reconciled with scale-free constitution and dynamics? *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 53, 84–93.
- Bechtel, W. & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 36(2), 421–441.
- Bechtel, W. & Abrahamsen, A. A. (2009). Decomposing, recomposing, and situating circadian mechanisms: three tasks in developing mechanistic explanations. In H. Leigeb & A. Hieke (Eds.), *Reduction and Elimination in Philosophy of Mind and Philosophy of Neuroscience* (pp. 173–186).
- Bechtel, W. & Abrahamsen, A. A. (2013). Thinking dynamically about biological mechanisms: Networks of coupled oscillators. *Foundations of Science*, 18(4), 707–723.
- Bechtel, W. & Richardson, R. C. (1993). Discovering complexity.
- Bratman, M. E. (1992). Shared cooperative activity. *The philosophical review*, 101(2), 327–341.

- Chudnoff, E. (2013). Intellectual gestalten. In U. Kriegel (Ed.), *Phenomenal Intentionality* (pp. 174–193). Oxford University Press.
- Couch, M. B. (2011). Mechanisms and constitutive relevance. *Synthese*, 183(3), 375–388.
- Craver, C. F. (2007). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. Oxford University Press.
- Craver, C. F. (2014). The ontic account of scientific explanation. In S. O. P. D. Kaiser, M.I. & A. Hütteman (Eds.), *Explanation in the Special Sciences: The Case of Biology and History* (pp. 27–52). Springer.
- Ellis, W. D. (1967). *A source book of Gestalt psychology*. New York, NY: Humanities Press.
- Fagan, M. (2012). The joint account of mechanistic explanation. *Philosophy of Science*, 79(4), 448–472.
- Fagan, M. B. (2015). Collaborative explanation and biological mechanisms. *Studies in History and Philosophy of Science Part A*, 52, 67–78.
- Fagan, M. B. (2016). Interventionist omissions: A critical case study of mechanistic explanation in biology. *Philosophy of Science*, 83(5), 1082–1097.
- Gallagher, S. (2017). *Enactivist Interventions: Rethinking the Mind*. Oxford University Press.
- Gallagher, S. (2018). New mechanisms and the enactivist conception of constitution. In M. Guta (Ed.), *Consciousness and the Ontology of Properties* (pp. 207–220). New York: Routledge.
- Glennan, S. (2002). Rethinking mechanistic explanation. *Philosophy of science*, 69(S3), S342–S353.
- Glennan, S. (2017). *The New Mechanical Philosophy*. Oxford University Press.
- Glennan, S. S. (1996). Mechanisms and the nature of causation. *Erkenntnis*, 44(1), 49–71.
- Goldstein, K. (1934). *The Organism: A Holistic Approach to Biology Derived from Pathological Data in Man* (trans: Kurt Goldstein, 1939 ed.). New York, NY: American Book Company.
- Grush, R. (2008). Temporal representation and dynamics. *New Ideas in Psychology*, 26(2), 146–157.
- Hamlyn, D. W. (1951). Psychological explanation and the gestalt hypothesis. *Mind*, 60(240), 506–520.
- Harbecke, J. (2015). The regularity theory of mechanistic constitution and a methodology for constitutive inference. *Studies in History and Philosophy*

- of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 54, 10–19.
- Illari, P. M. & Williamson, J. (2012). What is a mechanism? thinking about mechanisms across the sciences. *European Journal for Philosophy of Science*, 2(1), 119–135.
- Kaiser, M. I. & Krickel, B. (2016). The metaphysics of constitutive mechanistic phenomena. *The British Journal for the Philosophy of Science*, axv058.
- Koffka, K. (1936). *Principles of Gestalt psychology*. New York: Harcourt, Brace.
- Köhler, W. (1920). *Die physischen Gestalten in Ruhe und im stationären Zustand*. Braunschweig: Der Philosophischen Akademie Erlangen; Druck von Friedr. Vieweg & Sohn.
- Krickel, B. (2018). Saving the mutual manipulability account of constitutive relevance. *Studies in History and Philosophy of Science Part A*, 68, 58–67.
- Leuridan, B. (2011). Three problems for the mutual manipulability account of constitutive relevance in mechanisms. *The British Journal for the Philosophy of Science*, 63(2), 399–427.
- Lowe, E. J. (2003). Individuation. In M. J. Loux & D. Zimmerman (Eds.), *The Oxford Handbook of Metaphysics*.
- Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of science*, 67(1), 1–25.
- Mackie, J. L. (1974). *The Cement of the Universe*. Oxford University Press.
- Marom, S. (2010). Neural timescales or lack thereof. *Progress in neurobiology*, 90(1), 16–28.
- Merleau-Ponty, M. (1943). *Le Structure du Comportement* (6th (1967) ed.). Paris, FR: Presses Universitaires de France.
- Merleau-Ponty, M. (1963). *The Structure of Behavior*. Boston, MA: Beacon Press. Trans. Alden L. Fischer.
- Nagel, E. (1952). Wholes, sums, and organic unities. *Philosophical Studies*, 3(2), 17–32.
- Rescher, N. (1953). Mr. madden on gestalt theory. *Philosophy of Science*, 20(4), 327–328.
- Rescher, N. & Oppenheim, P. (1955). Logical analysis of gestalt concepts. *The British Journal for the Philosophy of Science*, 6(22), 89–106.
- Romero, F. (2015). Why there isnt inter-level causation in mechanisms. *Synthese*, 192(11), 3731–3755.
- Salmon, W. . (1984a). Scientific explanation: Three basic conceptions. In

- PSA: Proceedings of the biennial meeting of the philosophy of science association*, volume 1984, (pp. 293–305). Philosophy of Science Association.
- Salmon, W. . (1984b). *Scientific explanation and the causal structure of the world*. Princeton University Press.
- Sheredos, B. (2017). Merleau-Ponty's immanent critique of Gestalt theory. *Human Studies*, 40(2), 191–215.
- Smith, B. (1988). *Foundations of Gestalt Theory*. Munich and Vienna: Philosophia Verlag.
- Wimsatt, W. C. (1986). Forms of aggregativity. In N. P. J. A. Donagan & M. V. Wedin (Eds.), *Human Nature and Natural Knowledge* (pp. 259–291). Springer.
- Wimsatt, W. C. (1997). Aggregativity: Reductive heuristics for finding emergence. *Philosophy of Science*, 64, S372–S384.
- Woodward, J. (2002). What is a mechanism? a counterfactual account. *Philosophy of Science*, 69(S3), S366–S377.
- Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press.
- Woodward, J. (2015). Interventionism and causal exclusion. *Philosophy and Phenomenological Research*, 91(2), 303–347.