This is a section of doi:10.7551/mitpress/14630.001.0001

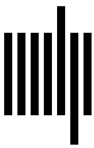
Context Changes Everything

How Constraints Create Coherence

By: Alicia Juarrero

Citation:

Context Changes Everything: How Constraints Create Coherence
By: Alicia Juarrero
DOI: 10.7551/mitpress/14630.001.0001
ISBN (electronic): 9780262374774
Publisher: The MIT Press
Published: 2023



The MIT Press

1

What Went Wrong? The Backstory

The controversy surrounding the concept of identity illustrates how philosophical and scientific presuppositions that go without saying entangle us in never-ending mazes.

In the history of ideas, claims that *identity* consists of a defining essence have a long backstory. In fact, one could arguably claim that it is among the oldest backstories in philosophy, a backstory that involves the paradox of the One and the Many. What makes different entities (e.g., dachshunds, boxers, chihuahuas, Great Danes) the same, that is, dogs? What makes *anything* continue being the same thing? How is it that something can persist as itself—despite changes wrought by either developmental processes, the environment, or internal malfunction?

Issues with identity are not new; the problem was first posed in the sixth century BCE. The classical version went as follows: If all the planks of the ship of Theseus are replaced over time, is it still "the ship of Theseus," even when not one of the original planks remains? Does anything persist as the same thing, the ship of Theseus, despite such significant alterations? How is change compatible with remaining the same? What does *same* mean, in this context or any other?

In 2019, Chinese scientists inserted a human gene into the brains of eleven monkeys; those monkeys then outperformed normal monkeys in tests of short-term memory. Even more striking, the modified monkeys' brain developed into an organ more like a human's than like that of the unmodified monkeys (Regalado 2019). Are these animals now partly human? In 2021, an international team of scientists injected human stem cells into primates (Subbaraman 2021). The interspecies chimeras grew for two weeks before being destroyed. If a human being's body parts are gradually replaced with silicon components until no original organic material remains, would they still be human, or a cyborg? If all

my memories disappeared, or if my beliefs suddenly became entirely different, would I still be me? What is it to be me? Just as the Ship of Theseus is one of the oldest backstories in philosophy, tales of aliens taking over the bodies of earthlings are among the oldest themes in science fiction.[1]

Today's controversy about identity is not just an old story; it is also an oft-told story (Toulmin 1990; Juarrero 1999). However, the central theme of this book is not identity, nor will this book be solely a survey of these debates. The problem of identity happens to be a noteworthy philosophical topic that encapsulates in a single controversy three categories that have been egregiously responsible for framing our views of reality in a misleading manner. Our understanding of many concepts, not only those that arise with respect to identity, continues to be misdirected by outdated and erroneous ideas about three central issues:

- Are wholes different from aggregates?
- Are context and history part of the fabric of reality?
- How do cause and effect work?

The three questions are so deeply interwoven in the way we see things that a discussion of one inevitably brings to bear the others. The problem of wholes and aggregates is the problem of Types or Kinds and Essences, the problem of context and history is the problem of Interactions and Relations, and the problem of causes and effects is, well, the huge philosophical problem of Causation.

The last one, Causation, is the most recalcitrant and pervasive of the three. It wraps around the other two: Do interactions among individual entities create real novelty in the form of coherent wholes? Can the context in which an event occurs have any influence either on whether it happens at all or on exactly what happens?

This chapter presents highlights from the history of these ideas. It describes how relations and interactions, context and history, came to be thought of as causally impotent; how seemingly coherent totalities were reduced to nothing but aggregates; and how the idea of cause and effect came to be restricted to energy-transferring processes. These controversies are problematic, not only because of what they have to say about the nature of identity; they mainly arise because of our deeply flawed views about *causality* and *coherence*—about the causes of coherence. Once we understand what went wrong with these two notions, we can begin to reimagine a different and more inclusive interpretive framework that can also make room for identity as contextually embedded coherence and interdependence. The book takes preliminary steps in that direction.

Types, Kinds, and Essences

The fable of the Scorpion and the Frog shows how "identity" has traditionally been thought to capture the essence that uniquely distinguishes something from everything else—everything else that something is not. The scorpion asks the frog to ferry it across the river safely. When the frog hesitates, fearful the scorpion would kill it, the scorpion protests that doing so would mean that both would drown. The frog accepts that reasoning. Halfway across the river the scorpion stings the frog. As they are both drowning, the dying frog asks the scorpion, "Why did you do that?" To which the scorpion replies, "I couldn't help it. It's my nature." Thinking of identity in this way suggests that if certain properties, internal to the creature, were missing, it would no longer be itself because those properties define what it is to be a scorpion. They capture, that is, a thing's fundamental essence and therefore its identity.

The idea of nature or essence as internal and standalone is noteworthy because of the unacknowledged assumptions it uncritically accepts. Essence as internal represents a refusal to acknowledge that interactions and relations play a role in a thing's nature; it also refuses to recognize that relational properties like coordination, integration, and context embeddedness are real. It ignores both the past and current circumstances. It underpins, in short, a worldview that dismisses time and place—context in general—from reality. These become passive containers instead. Such failures make it impossible to understand coherence and identity.

That backstory goes as follows.

Beginning with Thales in the sixth century BCE, pre-Socratic philosophers embarked on a search for the essence and nature of reality. What made pre-Socratics different from their precursors was their faith in reason and logic, as opposed to an unquestioning acceptance of mythic tales (Juarrero 1993). Is Water the foundation of reality, asked Thales? Is there more than one fundamental essence—perhaps, air, wind, fire, as well as water? Maybe all four, as Empedocles proposed? Or is the essence of reality change (Heraclitus)? Or an unchanging Plenum (Parmenides)? Lucretius, the first recorded *atomist*, noted that whatever turns out to be fundamental must be in-divisible—*a-tom*. Otherwise, it could be decomposed into more simple stuff and so would not be fundamental. The late Carl Sagan, host of the popular public television series *Cosmos*, lamented that the history of ideas went wrong when Socrates veered from this pre-Socratic interest in cosmological and natural questions and turned instead to ethical and normative ones.

Fast forward to the eighteenth century. Modern English atomist John Dalton noticed that chemicals could be combined into more complex substances or broken down into more fundamental elements. In a nod toward Lucretius, Dalton called the fundamental particles *atoms* and identified weight as their primary property. In that same century, French chemist Antoine Lavoisier added mass as a second primary property of atoms.

Primary and Secondary (Accidental) Properties

Modern science and philosophy consolidated the idea of primary properties as the ground of reality and therefore of identity; only primary properties constitute the essence of things. One of the features that allegedly makes primary properties primary is their ability to exist on their own, independent of others; they do not depend on anything else either to exist or to be perceived. Properties like atomic weight and mass—those that underlie solidity, extension, and quantity—are often listed as preeminent examples of primary properties. From this perspective, essence and identity are what might be called *internalist* concepts: they pick out properties that are internal to the entity in question; they are also objectively measurable and observer independent.

Once the internalist definition was adopted, concluding that interactions contribute nothing fundamental to reality followed. Relations with other entities in the environment and with the past bring context to bear, but because context was viewed as a container external to the thing's essence, it too was set aside as irrelevant to identity. The philosophical conundrum posed by the measurement problem in quantum physics originates in this framing.

Contrast the seeming ontological independence and self-sufficiency of atomic mass and weight with the context dependence of color and sound. It has been known since classical times (Democritus) that flavors, odors, color, sounds, and even temperature are inherently relational and contextual. Inserting a hand in a bucket of tepid water feels cold if it comes after inserting the hand in hot water, but it feels warm if it comes after icy water. Because they are not independent of perception, relational properties dependent on an observer were considered subjective.[2] In the language of academic philosophy, properties that arise from interactions, especially with human observers, are secondary or accidental properties. For that very reason, they are also ontologically secondary, their reality derivative.

In natural philosophy, *Kind* traditionally refers to natural kinds and is often used synonymously with *Type* and *Universals* to mean classes of

entities or processes defined by common primary properties.[3] "It's my [scorpion] nature" follows that reasoning. On this view, kinds and types pick out real, universal features in the world. *Scorpionality*, if you will, is real, the set of primary properties that makes scorpions *scorpions*. The elements of the periodic table are prototypes of natural kinds not least because they represent measurable and observer-independent properties such as the number of protons in the nucleus. All instances of element X have Y number of protons in their nucleus. This primary property constitutes their essence.

It is important to note that, even in classical times, types, kinds, and universals were understood to be multiply realized: differently embodied tokens can realize the same type. Some actual scorpions are ground dwelling; others are tree dwellers, but as scorpions, they possess the same basic primary properties. Likewise, whether isosceles, equilateral, or scalene, actual triangles are *tokens* of the type triangularity. They all realize its essential, primary properties: three-sided, enclosed, two-dimensional figures whose interior angles sum to 180 degrees. They would not be triangles if they did not.

According to Plato, even if material tokens of types ceased to exist, if this triangle I drew in the sand got washed away with the tide, the essence of triangularity would continue to exist even if no longer instantiated. Platonic Forms such as Justice, Goodness, Truth, and Beauty as well as triangularity and numbers were thought to exist in a transcendental realm, fixed and eternal. They were considered real even if no actual triangles, or written numbers, or actual cases of justice, goodness, truth, or beauty had ever existed.

So, Carl Sagan did not get it quite right: Socrates was not entirely deviating from the tradition of the pre-Socratics. Platonic dialogues such as the *Republic* (about justice) and *Meno* (about knowledge acquisition) narrate Socrates' attempts to elucidate and define the universal and eternal essences of justice, knowledge, truth, and goodness. Socrates' efforts to articulate the inherent properties of universals are therefore not entirely unlike pre-Socratic attempts to discover the essential traits of natural phenomena.

Because types and forms can be realized in multiple instances or *tokens*, each different from the next, providing a full list of actual realizations does not capture a type's essence; to do so, scholars must discover the full set of its primary properties. That is what real (as opposed to ostensive) definitions do. Even after Plato's influence gave way to Aristotle's, and well before formulating natural laws became the central aim of science, classifying natural kinds such as types of rock (igneous, sedimentary, metamorphic) or biological *taxons* such as vertebrates by discovering their primary properties consumed most scientists' time.

Multiple realizability, also called *degeneracy* in biology, where it is ubiquitous, will play a key role in this story. The terms refer to the capacity of different events and processes to realize the same function or other higher-level property. The main idea presented in this introductory chapter is that according to the received understanding of types and kinds, individual tokens of a given kind were thought to differ only in their secondary properties. Their essential, primary properties, those that identify them as a type of thing—a given species, for example—are inherent, universal, and unchanging throughout. As late as the nineteenth century, this was even thought to be so in nature; Darwin's proposal that new species originated by evolving in response to selection was controversial for that very reason.

* * *

The way philosophers thought about essences and types evolved over the course of the history of ideas. As just mentioned, Plato concluded they were real transcendent forms. Aristotle, on the other hand, did not agree that form and matter could exist independently. In opposition to Plato's disembodied realm of forms, Aristotle focused on *embodied Substance*, the unity of *in-formed matter*.

Living things are quintessential examples of Aristotelian in-formed matter. In contrast to compacted masses of pebbles, for instance, organisms "hang together" in a unique way; they embody or realize a coherence that Aristotle postulated as the essence of substance.[4] Aristotelean substances are also *independent existents*, the other hallmark of reality for both philosophy and science.[5] On this view, actual lions are the coherent unity of the form Lion with the appropriate matter. In comparison to colors, sounds, or flavors, actual lions do not require observers to exist. In lieu of the otherworldly existence of Plato's forms, the essence of Aristotelean substances was therefore embodied self-sufficiency.

From both Plato and Aristotle's realist perspective, forms and universals are real. It was not until later in the history of Western thought that universals or types came to be considered mere labels or tools with which we think and organize our concepts and actions, or with which we design and conduct scientific experiments. This perspective, known as *nominalism*, held that the generality of common nouns is our doing; properties that define common nouns do not refer to objective essences. In contrast to realism, nominalism argued that all that really exists are actual particulars.

With the rise of scientific method in the sixteenth century, observation and analysis were elevated over a priori reasoning and synthesis.

In *Novum Organum* (*New Method*), English philosopher and political leader Francis Bacon proposed a new way of reasoning. In contrast to medieval Scholasticism's focus on a priori reasoning, Bacon's scientific method was based on induction, on inferring general conclusions from fine-grained, precise observations. In *Discourse on Method* a century later, French philosopher and mathematician René Descartes laid out a technique for knowledge acquisition that returned to deduction. Nevertheless, it required thinkers to begin with the most elementary and therefore indubitable ideas instead of accepted beliefs or revealed truths. In either case (Baconian induction or Cartesian deduction), the proper way to acquire knowledge was by starting from the elements (bare observations or clear and distinct ideas) and then proceeding to reconstruct the whole.

Threats to coherence are already present in this approach. Before Descartes, medieval thinkers like Christian Thomas Aquinas, Jewish Moses Maimonides, and Muslim Ibn Rushd (Averroes) followed Aristotle in the assumption that coherent substance—unity of form and matter—is the default basis of reality. What is the source of the coherence of embodied substance? Alas, Aristotle's writings did not fully resolve the question of how embodied substance, in-formed matter, is bound into coherent and unified wholes (Gill and Lennox 1994).

In contrast, influenced by modern atomists such as French thinker Pierre Gassendi and English experimentalist Robert Boyle, modern scientists and philosophers prioritized analysis over synthesis, parts over the whole. In doing so, the philosophical problem of identity was transformed from "How does essence in-form matter and confer on it its identity?" into "How do primary properties cohere into complex wholes?"

Whence coherence? Or, more precisely, whence the cause of coherence?

In Descartes' writings, Aristotelian embodied substance sundered into mind and matter, two distinct ontological domains, each with its own substance and each defined by its distinct primary property. Mind is defined by its primary property, res cogitans, whose essential trait is consciousness, and matter, res extensa, by its primary property, spatial extension. Known as dualism, the theory maintained that both substances, mind and matter, were real.[6] It is unsurprising, therefore, that by the middle of the seventeenth century, fears about whether coherence was even possible were being voiced. In his famous poem, *An Anatomy of the World* (1611), English poet and soldier John Donne lamented, "Tis all in pieces, all coherence gone."

But doubts about coherence did not begin in the seventeenth century. Philosophers as far back as Plato were sophisticated enough not to confuse form (an entity's core type identity) with shape. Justice was a Platonic form, for example, although it obviously has no shape. If forms are transcendent, however, how are actual entities related to the form that makes them what they are, that is, that imbues them with their essential nature?

Plato's famous allegory of the cave taught that physical objects are mere shadows of forms. Over the centuries, explaining how shadows or simulacra (material tokens) "partake in" or embody transcendent forms became a cottage industry. Plato thought of matter as a corrupting influence: when ideal forms get mixed in with matter, the result is a degraded reality. A few centuries later, and to avoid the Paulist and Augustinian disposition to label all matter corrupt after the Fall, Neo-Platonists such as Plotinus identified Platonic forms with God and suggested that they emanate into matter. Forms somehow ooze into matter?[7] Philosophy searched for a principle of coherence that binds transcendent forms to matter such as to generate coherent wholes.

The generation and preservation of coherence and its emergent properties is the central subject matter of this book.

Descartes' proposal that mind and body are two separate and distinct substances, each capable of independent existence and characterized by nonoverlapping essences, raised additional questions. Juarrero 1999 focused on how intentional causation and therefore purposive action (as opposed to reflex reaction) are even possible. How can two substances from entirely different ontological domains, the extended physical and the conscious nonphysical, *interact* at all? How can a nonphysical event like the intention to do X move the body such that the ensuing behavior carries out the content of that intention, doing X? Purposive action and the mind–body problem became the crucible of coherence. This book focuses on mereological causation, that is, on how interacting entities generate wholes with novel properties and how those wholes, once they coalesce, guide behavior. In particular, it focuses on the manner of causation that generates and preserves parts–whole and whole–parts coherence.

Interactions and Relations

Hypotheses about coherent wholeness and its primary and secondary properties intersect with views about the status of interactions and relations. As noted earlier, context and interactions, including fluctuations

and perturbations, were presumed not to affect essential primary properties of fundamental elements or substances. To repeat, whether these consist of mass and atomic weight or thought and extension, primary properties were allegedly eternal, fixed, and universal, and they inhered in the entity whose essence and identity they underpin. Neither context nor history, not to mention process, alteration, or flux, could change those fundamental building blocks of reality.[8] "It's my nature." The essence of particles and substances (including minds) was internal, remained unaltered despite secondary modifications, and was capable of separate and independent existence. And, as always, relations and interactions with other particles or substances were considered neither essential nor foundational.

* * *

The field of chemistry is all about relations and transformations; chemists study interactions among molecules and charged particles. However, if relations and interactions are secondary, chemical compounds and chemical synthesis cannot then be fully real, *qua chemical*, as philosophers say. They must be nothing but aggregates of primary and essential properties of constituent physical elements. Chemical properties must be the accidental effluvia of clumped particles.

On this view, then, the apparent unity of chemical compounds, or living things, is illusory; wholes (including organisms) are mere agglomerations, the sum of their elementary parts. Material aggregates do not generate coherence; they are just masses of physical stuff. They are nothing but macroscopic clumps of elementary particles. They might appear to possess qualitatively novel capacities and properties like oxidation, life, sensation, and perception or even mind and consciousness. In reality, however, those are mere side effects—castoffs—of aggregated masses of primary properties. Those seemingly emergent properties and powers are ontologically derivative. They bring with them no powers of their own.

The principle that coherence is nothing but impotent aggregation is still presumed to apply generally, not only in chemistry. It motivates the philosophical perspective called reductionism, often characterized as "nothing but-ism." Reductionism maintains that because wholeness is nothing but particles and their secondary interactions, any seemingly novel properties of purportedly coherent wholes can be derived, in principle, if not in fact, from laws pertaining to its constituent elements. Colloquially, "Wholes are nothing but the sum of their parts." On this view, descriptions of *coherence* are either convenient labels or epistemic simplifications to which scientists resort because computing all the details of micro-interactions among primary properties takes too long.

In summary, reductionists maintain that, in principle, chemical, biological, psychological, sociocultural, or ecological processes, reactions, combinations, transformations, changes, and rearrangements of elements will all be fully accounted for by internal, context-independent, and fully quantifiable, essential—that is, primary—properties of fundamental particles.

The inevitable conclusion of the foregoing is that seemingly coherent entities like biological organisms and chemical compounds are causally powerless *as* biological and chemical wholes; if apparently coherent "wholes" are mere aggregations, their seemingly emergent properties are powerless. In the jargon of philosophy, coherence and emergent properties that characterize that apparent coherence are *epiphenomenal*, that is, causally impotent as such. The heavy causal lifting is always and only done by the fundamental elementary particles (van Gulick 2004). Accordingly, the arrow of explanation always looks downward, as the saying goes, because reality is grounded in the primary properties of elementary particles, which underwrite natural laws.

In short, misconceptions about causation cut across the previous considerations. According to the standard physics-based framework, as just described,

- Essences reside in internal and self-sufficient primary properties capable of independent existence. This implies that context and history are not fundamental features of reality.
- "Wholes" that appear to display novel and emergent properties are, at bottom, nothing but the sum of those primary properties and their interactions.

It follows that causal powers reside exclusively in fundamental particles and their primary properties. Bottom-up influences might bubble up from parts to aggregates, but with the following caveats: (1) any apparent coherence or wholeness is mere aggregation, and (2) seemingly emergent properties such as phenomenal awareness and feelings are, likewise, epiphenomenal characteristics of aggregated elementary particles.

Specifically, according to reductionism, our intuition that mental processes such as intentions and beliefs have powers to actively bring about meaningful, purposive actions is illusory. Thoughts, feelings, and intentions derive their powers and properties from biology; biology from those of chemistry; chemistry from physics. Properties that appear unique to biological organisms (such as being alive) or human beings (such as symbolic language) can, in principle, be inferred from chemical processes that

constitute them. Causal powers that seem to issue from those higher-level properties can be derived from physical properties, at least in principle. It is not quite "turtles all the way down," however. The turtle at the bottom (at the level of elementary physics) is special. The primary properties of a-toms, reality's constituents (read now quarks and electrons), are the real and most simple stuff that does the causal work and provides explanatory power. Ultimate causes reside in and issue from there. Meaning and purpose are impotent.

Such is the dream of a theory of everything, the promise of an equation that spells out the lawful correlations among microdetails and from which everything else can be derived and precisely predicted. It is this intellectual heritage, I submit, that prompts well-known philosophers like Australian David Chalmers to espouse *double-aspect monism* or its close relative, *panpsychism*, the view that even the fundamental constituents of reality must have a built-in nonphysical aspect, mentality. Without literally building in consciousness into the most basic building blocks of reality from the start, hopes of showing how novel properties emerge from material stuff are doomed to failure because material stuff can only increase in quantity. The Cosmos does not actively create qualitative novelty. Full stop.

Efficient Causality

A particular understanding of cause-and-effect relations advanced by Newtonian science gave additional support to this worldview. By restricting cause–effect relations to *efficient cause*, the ontological status of coherent wholes became even more suspect.

I turn now to the backstory on causality.

The first systematic treatise of causation was Aristotle's *Physics*,[9] which argued that everything must be explained in terms of four Causes. The first three are Material cause, the physical stuff from which something is made; Final cause, the purpose or goal the substance brings about; and Formal cause, the inherent essence that makes it that kind of substance and no other. Consider a potter shaping a pitcher or a carpenter making a chair. Material causes of the pitcher and chair are clay and wood, respectively. Final causes are pouring liquids and sitting. (In nature, the *intrinsic teleology* or final cause of, say, acorns, is to become an oak. Teleology is built into nature.) Aristotle illustrated the fourth cause, Efficient cause, with descriptions of the potter's hands transforming clay into a pitcher and the carpenter's arm wielding saws or hammers[10] on wood. More recently, the

preferred illustration of efficient causes is cue sticks striking billiard balls. Efficient causes are energetic forces acting on matter.

With the advent of modern science and philosophy, final and formal causes (which Bacon disparaged as "metaphysics") could be dismissed as "no part of science." Empiricism held that *material* cause could be accounted for through empirical observation of primary properties, and purposive or *final* cause came to be viewed as an irrelevant leftover from the discredited belief that the ultimate drive of all living things is toward perfection and God. The ghostly concept of formal cause became otiose as nominalism took hold.

As a result, modern science turned to Aristotle's *efficient cause*, the transfer of energy to matter with which forceful impacts bring about effects. After the sixteenth century, the term *causality* and its cognates came to mean exclusively efficient cause, the transfer of kinetic energy from an agent-cause to a body. Since the advent of Newtonian mechanics in the seventeenth century, cause–effect relations have been conceptualized solely in terms of motive forces. This perspective was amply rewarded with Newton's Laws of Motion and justifiably so. I hasten to state unequivocally that nothing in this work should be taken as a rejection of or argument against efficient causes or Newtonian science. When calculating planetary motions, however, adding a third planet yields a system that is chaotic and unpredictable. The laws of motion fail to predict their trajectories. Newton himself was aware of this so-called three body problem, an example of environmental influence completely transforming a system's dynamics. The puzzle was an early hint that context should not be so readily dismissed, but modern science and philosophy ignored its warning.

* * *

By definition, efficient causes exist separately from their effects. Relying exclusively on this form of causation, however, brought with it other, less well-known implications.

Aristotle's reasoning about efficient causes went as follows: since causes and their effects cannot simultaneously exist both before (as cause) and after (as effect of) themselves, causes are spatiotemporally "other than" their effects. The potter's actions start before they touch the clay. The motion of the cue stick begins before it hits the ball. The potter is spatiotemporally distinct from the pitcher; the cue stick is, likewise, "other than" the ball. The inevitable conclusion is that *self-cause* and other forms of "circular causality" are impossible because causes and effects would have to exist simultaneously: before (as agents) and after (as effects), a contradiction.

Conclusion: the view that reality rests on primary properties of elementary particles and that causality is exclusively efficient cause is at the root of modern science and philosophy's failure to explain the causal powers of systemic properties. Efficient causality cannot account for how parts become interwoven into interdependencies that bind together coherent wholes and persist despite turnover or deletion of component parts. Much less can it account for top-down causation from the emergent features of coherent wholes to those components or behavior.

Mereology

As noted, philosophers call relations from parts to wholes and wholes to parts *mereology*. According to the default ontological framework described (and that is still very much in place), parts cannot interact to produce wholes with strongly emergent properties—that is, with properties that are not reducible to the sum of the parts. Efficient causes produce mere aggregates. Emergent characteristics of wholes, as wholes, such as coordination and synchronization, cannot influence components because they could only do so as efficient causes, which is impossible since parts and wholes are not spatiotemporally distinct. And since aggregates are epiphenomenal, mereological "causation" from clumped particles to components is impossible (Campbell 1974; Kim 1989; Moreno and Mossio 2016; also see Murphy, Ellis, and O'Connor 2009). These misconceptions about mereological relations have occasioned no end of problems, many still unrecognized.

Working from the assumption that all cause is efficient cause, therefore, philosophers and scientists have resisted the idea of *recursive* or *circular causality* ever since. The problems of identity and intentional causation were perhaps the most insistent philosophical topics that wouldn't go away. Juarrero (1999) chronicled the epicyclic contortions to account for purposive behavior while refusing to accept recursive causality.

How can two separate and distinct substances from different realms interact in terms of efficient causality? How can mind activate matter? Descartes identified the pineal gland as the site where mental events such as intentions trigger purposive action. With the advent of thermodynamics in the nineteenth century, this maneuver ran into a different objection, this time involving modern science's conservation laws and principle of physical closure.

Physical Causal Closure and Overdetermination

Combine an exclusive reliance on efficient causality with the First Law of Thermodynamics, which holds that that the total amount of matter and energy in the universe is fixed. Matter can neither be created nor destroyed; it can be transformed (into energy) and back, but the total amount is always conserved. In combination, these two principles lead to a third, *causal closure of the physical*. This principle holds that (1) all states are physical states, and (2) all physical states are the effects of physical efficient causes. The unavoidable implication of these two premises is that proposing that the symbolic content of thoughts or beliefs can bring about changes in the physical realm violates conservation laws and causal closure.

The reasoning goes as follows: nonmaterial "causes" such as Aristotelian Formal and Final causes or Cartesian intentions would activate bodies by introducing themselves into the physical realm and activating them as efficient causes. In combination with the resolutely materialist presuppositions of the principle of physical closure, this insertion into the physical world would increase the universe's total matter and energy and violate conservation in the process. Conclusion: (allegedly) nonphysical thoughts, intentions, and other mental events as such cannot bring about physical effects without violating causal closure.

Purposive causation by nonphysical intentions would also imply an *overdetermined universe*, likewise prohibited by causal closure. The argument here goes as follows: causal closure requires that all events (including arms rising) be the effects of physical causes—that is, they must result from the energy transfer of efficient causes. But if intentions are Cartesian mental events and my arm rises because I intended to raise it, the resulting behavior is qualitatively different than if my arm rose because of an involuntary neuromuscular spasm or because someone pushed me. Intentional causation transforms behavior into purposive action, with all the legal and moral implications that entails (Murphy and Brown 2007; Juarrero 1999).

But if intentions as meaningful mental states can bring about purposive action, the arm motion would be caused by both my intention to raise my arm and the neuromuscular processes required by the principle of causal closure. Such double causation overdetermines the universe. Recall that on the standard understanding of efficient causes, causes and effects must be logically and spatiotemporally distinct, so on this account, intentions must be other than neurophysiological processes, which would also violate causal closure.

Conclusion: allegedly nonphysical mental events such as intentions, thoughts, and beliefs cannot cause actions as meaningful and purposive intentions to do X. These considerations reinforced the received view that systemic properties are epiphenomenal.

It follows from this reasoning that postulating top-down (downward) causation, from intentions in virtue of emergent mental properties down to behavior (Campbell 1974), violates conservation laws and the principle of the closure of the physical. Western philosophy's exclusive reliance on efficient cause and total dismissal of context bars naturalist accounts of mental states and mental causation at the root. By maintaining that identity rests on primary properties of fundamental material particles (while simultaneously excluding all forms of influence other than efficient causality and insisting that relations and interactions with context are secondary and epiphenomenal), understanding how mind and body could possibly constitute a coherent whole, much less how identity can be preserved and persist despite changes, becomes impossible.

As framed by standard philosophy of science, naturalism closed off debate about the causes of coherence and therefore about identity and individuation. Without a principle that explains the formation of coherent wholes from previously separate parts, natural science and philosophy of mind were also left without a way of accounting for either mereological relations or interactions with the environment and influence from the past. Nothing can generate outcomes that are qualitatively different than mere aggregation without violating causal closure. Reductionism followed. Human beings are nothing but atoms clumped together in time and space.

* * *

Over the centuries, several ways of resolving this impasse were proposed. The early twentieth century witnessed a flurry of philosophers who espoused an approach labeled emergentism.[11] In keeping with the reductionism required by natural philosophy, emergent properties were invariably deus ex machina contributions, literally injected from outside the physical realm into inchoate and passive matter. Emergentists such as French philosopher Henri Bergson were often labeled Vitalists in reference to the mysterious and externally provided "vital" force that they proposed animates living things. Natural scientists summarily ignored these proposals.

Other thinkers such as Scottish philosopher John Stuart Mill, Australian-born philosopher Samuel Alexander, French scientist Henri Poincaré, and Austrian biologist Ludwig von Bertalanffy cast about for a systems-theoretic approach to account for novel properties of wholes. Michael Polanyi tried to have it both ways by arguing that although ultimate boundary conditions

were set from without, once in place, they can generate life without outside influence. Charles Sanders Peirce, a practicing chemist, knew there was something more to chemical compounds than mere aggregation (Juarrero and Rubino 2010), but in the end gave up on finding an ontological answer and, in keeping with Kant's epistemological turn, settled for a semiotic and "pragmaticist" account.

For the most part, however, the reductionist principle that aggregates are nothing but the sum of their parts and that seeming coherent wholes such as organisms like you and me are epiphenomenal, powerless to exercise causality as anything other than masses of atomic primary properties, ruled the day. One can sympathize with John Donne's lamentation that all coherence is gone.

But the problems of coherence and identity did not go away. Intuitively, it is difficult to shrug off the belief that there is some logic that holds organisms together over time despite perturbations. Organisms are organized; they present an internally coherent and persistent "hanging together" that is qualitatively different from mere aggregation. What makes coherent processes in general, including living things, cohere and persist as the same type of entity despite dramatic alterations? Suggesting, as early complexity theorists proposed in the mid-twentieth century (Prigogine and Stengers 1984; Maturana and Varela 1979; Pattee 1972a), that constrained and self-organizing dynamics might generate coordination and complex organization with qualitatively novel properties went completely against the grain of the received framework. With very few exceptions (Collier and Muller 1998; Cilliers 1998; Collier and Hooker 1999), mainstream philosophers of science disregarded complexity.

Meanwhile, the world moved on. Among the issues raised in identity politics today is the role contextual factors such as race, colonialism, class, culture, sexual orientation, and age play in personal identity. Environmental and historical influences are implicated in making all living things what and who they are. In biology, the new field of epigenetics has shown that the dichotomy between nature and nurture is too simplistic. The even more recent research on the metagenome in the gut microbiome, not to mention lichens and mycorrhizae, raises even further questions about symbiotic wholes. Are human beings more symbiotic wholes than individual organisms? Just asking this begs the question of this book's central theme: what causes coherence? In any case, the physics-based worldview from which our understanding of causes and effects, coherence, and identity derives continues to assign only secondary or accidental ontological

status to context and history. It prioritizes separate and standalone existence over relations and interactions.

* * *

Mistrust in the dichotomy between essential and accidental properties surfaced in academic philosophy circles in the early twentieth century. Austrian philosopher Ludwig Wittgenstein and American philosopher Willard Van Orman Quine independently held that the distinction between primary and secondary properties was untenable. As an example, there are no necessary and sufficient primary properties that make all games *games*.[12] As students at Oxford University during World War II, Philippa Foot, Mary Midgely, Iris Murdoch, and G. E. M. Anscombe argued against the view that reality is wholly transcendent and that the faculty of moral intuition is independent of features in the world. These remarkable thinkers emphasized instead social, creative, spiritual, and biological forms of life as the source and weave of human life. Alas, the significance of their ideas is only now being recognized (Lipscomb 2021; Cumhaill and Wiseman 2022).

Since the received view of essences and type of entities was predicated on primary properties, the next philosophical step was usually a retreat to nominalism, the theory mentioned earlier to the effect that any pretense that types and kinds are real should be abandoned: common nouns like *triangle* or *scorpion* do not refer to real categories or to real entities marked by an internal principle of coherence; they are only linguistic units or epistemic cuts that frame our thoughts and investigations.

Remarkably, even though he was discussing the failure of essentialist definitions, the later Wittgenstein did not explicitly retreat to nominalism. He proposed instead that language captures "forms of life," a phrase often interpreted as the opening salvo of a social constructivist approach to language. Wittgenstein's phrase suggests to me, as it did to Anscombe, Foot, Midgely, and Murdoch, not an outright retreat from realism but an attempt to reimagine the world from the point of view of relations and circumstances, context and history.

An opening to a contextual and relational ontology, in other words.

Dismissing the influence of context is the original sin whose imprint continues to permeate many of these philosophical and scientific debates. This book will argue that recognizing that context changes everything reopens a path toward rehabilitating coherence, identity, and causation—from parts to coherent wholes and from emergent coherent wholes to parts, including intentional causation. It proposes a realistic and naturalist

philosophy based on the principle that nature self-organizes and sorts itself into real but contextual *coherent dynamics*. Call its principle of coherence its *constraint regime*.

To be consistent, Platonic forms as well as final and formal causes must in consequence be reimagined in terms of interactional, extended, and *context-dependent interdependencies*. Thus understood, relational types are real and coherent patterns of energy flow, structure, and activities that form locally from contextually constrained interactions among individuals and that, in turn, as coherent dynamics, constrain the individuals and circumstances from which they emerge. Reimagining cause-and-effect relations, especially mereological relations between parts and wholes, and the influence of context and history on those relations, will be the hinge on which this reformulation turns. By taking context seriously in this fashion, coherence and wholeness, as well as meaning, "shall be returned."

This new interpretive framework makes room for a more expansive notion of identity and coherence based on interdependence and characterized by resilience and robustness. And generated by constraints.

Even John Donne might approve.

This book was set in Sabon by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Juarrero, Alicia, author.
Title: Context changes everything : how constraints create coherence / Alicia Juarrero.
Description: Cambridge, Massachusetts : The MIT Press, [2023] | Includes bibliographical references and index.
Identifiers: LCCN 2022030581 (print) | LCCN 2022030582 (ebook) | ISBN 9780262545662 | ISBN 9780262374781 (epub) | ISBN 9780262374774 (pdf)
Subjects: LCSH: Complexity (Philosophy) | Causation.
Classification: LCC B105.C473 J83 2023 (print) | LCC B105.C473 (ebook) | DDC 117—dc23/eng/20230124
LC record available at https://lccn.loc.gov/2022030581
LC ebook record available at https://lccn.loc.gov/2022030582