This is a draft of a chapter/article that has been accepted for publication by Oxford University Press in the forthcoming book *Alternative Approaches to Causation: Beyond Difference-Making and Mechanism* edited by Yafeng Shan due for publication in 2023.

Causal Efficacy: A Comparison of Rival Views

R. D. Ingthorsson

The paper critically assesses whether four rival accounts of efficient causation—*transmission*, *mechanistic*, *powers-based*, vs. *powerful particulars accounts*—are compatible with the scientific understanding of two often discussed types of causal phenomena: (i) collisions between billiard balls, and (ii) how water dissolves salt. It is argued that only the powerful particulars view can be considered compatible with the scientific understanding, mainly because the other three characterize interactions—to varying degrees—in terms of a unidirectional exertion of influence of one thing on another, which is incompatible with the established scientific fact that all interactions are perfectly reciprocal.

Key words: *transmission accounts of causation*, *mechanistic accounts of causation*, *powers-based accounts of causation*, *powerful particulars view of causation*, *efficient causation*, *causal realism*

1. Introduction

The aim of this paper is to critically discuss four rival views about how to best make sense of causation as a real mind-independent feature of the world that involves the production of changes through the exertion of influence of something on something else. Is causation a question of (i) transmission of conserved quantities (*transmission accounts*), (ii) activities of the parts of mechanisms (*mechanistic accounts*), (iii) mutual manifestation of powers (*powers-based accounts*), or (iv) reciprocal action between powerful particulars (*powerful particulars account*)? I argue in favour of (iv). Note that in this paper I will use the term 'interaction' to denote any phenomenon where something exerts an influence on anything else, even though I will argue that the scientific understanding of interactions, the one I promote, is of *reciprocal action* between two entities, notably that whenever any entity *A* exerts any kind of influence on any other entity *B*, *B* will at the same time exert the same kind of influence on *A*, and to the same magnitude.

It bears to mention that transmission and mechanistic accounts are not always counted as causal realist accounts—i.e., as treating causation as a mind-independently real phenomenon—but only as attempts to clarify causal reasoning. This is partly because some proponents explicitly take that stance, and partly because these accounts originate in the empiricist tradition of accounting for everything in terms of what can be observed and are therefore unreflectively grouped with neo-Humean accounts. Indeed, their proponents share an aversion to the postulation of anything popularly regarded as 'unobservable' such as powers. However, I believe the thinkers mentioned in this paper do treat causation as a mind-independent and real phenomenon and therefore can count as causal realists.

I will first briefly sketch the core ideas of the different accounts before ventilating my concerns about each view, first in general terms and then with respect to the way they explain two kinds of causal phenomena: (i) collision between billiard balls, and (ii) how water dissolves salt. It will be argued that transmission, mechanistic, and powers-based accounts are, to varying degree, incompatible with certain scientifically established facts about the behaviour of physical entities, in a way that the powerful particulars view is not.

2. Transmission accounts

I will initially use Wesley Salmon's account of causation in terms of *causal processes*, *propagation,* and *production* (Salmon 1980, 1984) to illustrate the core idea of transmission accounts, since most contemporary accounts are influenced by it, such as Philip Kitcher (1989), Phil Dowe (1992), and Max Kistler (1998). Later, I will refer to Dowe (2009) for a more recent formal statement of the core tenets of transmission accounts.

It bears to mention that Salmon's account is sometimes described as "neo-mechanistic" because it was originally a revival of the idea that explanation is about identifying the causal mechanisms that produce natural phenomena (Galavotti 2022). However, he ultimately moved towards a transmission account of causal mechanisms.

According to Salmon, the persistent entities that form the basis of most scientific ontologies—particles, molecules, cells, animals, planets, etc.—are to be understood as *causal processes*; they are causal because able to exert influence on each other, and they are processes because science reveals them to be continuously changing even though they may appear not to. Ultimately, everything above the level of elementary particles, is either a continuously changing atom or made of atoms of that kind.

Salmon denies that talk of 'process' is a commitment to event ontology, i.e., that entities are structured series of events. However, he doesn't offer an alternative analysis of process; he just points to examples of things that remain the same through continuous intrinsic change (1984: 139ff). His rejection of event ontology is not a commitment to substance ontology either, but a rejection of *any* ontology that tries go beyond the observable. Of the transmission accounts that developed in Salmon's wake, some work with events as the relata of transmission (Kistler 1998) while others stick to Salmon-style causal processes (Dowe 2009).

Propagation refers to the way causal processes conserve their structure over time when not interacting with other processes and consequently conserve their ability to exert causal influence in any future interaction. *Production* refers to the way interactions between causal processes result in modifications in their structure and in their ability to exert causal influence.

Salmon resists commitment to causal powers for the same reason he resists commitment to event ontology; they are unobservable and therefore mysterious. Instead, he first uses the term 'marks' to refer to the observable properties of causal processes recognised by the sciences, and which play the same explanatory role that powers are meant to do (Salmon 1980). Later, due to the criticism from Kitcher (1989) and Dowe (1992), Salmon adopted the view that changes are produced in interacting causal processes because one transmits a conserved quantity to the other.

Salmon's view can be illustrated concretely by considering a batter hitting a baseball with a bat in such a way that it leaves the pitch and breaks a window in a neighbouring house. The bat and baseball are causal processes propagating the conserved quantity of momentum. On contact, a conserved quantity of momentum is transmitted from bat to ball. The ball then preserves its momentum as it travels across the pitch, propagating it from point A to point B. When the ball interacts with a window it is in virtue of the preserved momentum that it breaks the window.

Let us now look at a more formal statements of transmission accounts; one event-based and another based on causal processes. Kistler presents the event-based account in what he calls a 'reduction statement':

> (T) Two events a and b are causally related in the sense that one is a cause of the other if and only if there exists a conserved quantity Q of which a particular amount P is transmitted between a and b (Kistler 1998: 1)

Unfortunately, the statement is not specific enough about the relevant sense of 'event' to distinguish causal and non-causal instances of transmission, and therefore doesn't distinguish between what Salmon calls propagation and production. If events are understood Kim-style, i.e., as a particular *a* instantiating a property *F* at a time *t*, then conservation laws entail that any quantity present in any Kim-style event e_1 will be transmitted to the next event e_2 whether or not anything causal occurs between e_1 and e_2. For instance, a uniformly moving body will through any given temporal interval constitute a succession of distinct Kim-style events between which any conserved quantity possessed by any event e_i will be transmitted to a subsequent event e_j (see definition of 'world-line' below). To be fair, Kistler obviously assumes that everyone understands that (T) is only meant to apply to what happens in causal interactions between distinct entities. However, this needs to be spelled out explicitly to avoid objections by hair-splitting readers such as myself. More importantly, to make (T) more precise it seems necessary to introduce some term for the persistent particulars that are the constituents of events. There are other worries, but they apply just as much to the causal process version discussed below.

The causal process version specifies that while a causal process is continuously transmitting conserved quantities (really, *propagating* them), it is only when something is transmitted between distinct processes to *produce* change that we can talk about causal interactions. This position can be summed up in terms of three claims (for reference, see Dowe 2009: 219):

Causal process: an object that possesses a conserved quantity.

World-line: a spacetime trajectory of a causal process in which it propagates a nonzero amount of a conserved quantity between every spacetime point of that trajectory

Causal interaction/production: an intersection of world lines A and B ($A \neq B$) that involves exchange of a conserved quantity producing a modification in A and B.

Dowe's causal process version of the transmission account, resolves the worries I raised about the event-based version. However, there are other problems, most of which are also present in mechanistic and powers-based accounts. It is important for what comes later, that while an 'exchange' implies a certain kind of reciprocity between *A* and *B*—both change in proportional albeit different ways—then the reciprocity is always assumed to involve the loss of a quantity by *A*, which *B* gains; exchange goes one way from *A* to *B*. It is this unidirectionality of the exchange that is meant to constitute the direction of causation.

3. Mechanistic Accounts

The core idea of mechanistic accounts is that causation is the activities of compound parts of organised wholes that produce changes in either whole and/or parts. It is in fact a requirement that causes and effects *must* be connected by mechanisms (Glennan 2017: 145). There are some disagreements to be found between accounts, say, between the 'mechanism first' approach favoured by Stuart Glennan in earlier works (2009) and the 'activities first' approach advocated by Machamer (2004) and Bogen (2008). However, I don't think it is of any consequence for this paper to confine the discussion to the most recent systematic presentation of the mechanistic philosophy found in Glennan's book *The New Mechanistic Philosophy* (2017). Indeed, the influences from Machamer and Bogen, as well as from William Bechtel and Adele Abrahamsen (2005), Carl Craver (2007), and Lindley Darden (2008), just to name a few are clear to see. I will refer to this view as 'NMP'.

NMP differs from the empiricist and reductive accounts of causation it is sometimes associated with, in rejecting the idea that causes and effects can be understood as Kim-style events. NMP stresses the *processual* nature of causation as a natural consequence of taking causal production to consist in the *exertion* of influence between particulars, and that this exertion must be understood as an *activity* that cannot fit in an instant (Glennan 2017: 177). To be more precise, the causally relevant sense of 'event' is of particulars (plural) doing something *to each other* since a cause is never the activity of just one particular but an interaction between parts of a mechanism. This removes the need to further distinguish between causal and non-causal events.

In more detail, Glennan suggests we should understand causation in terms of *constitutive*, *precipitating*, and *chained* production. Constitutive production is what happens as a result of interactions between parts of a mechanism, and which can only be understood in terms of the activities of the whole mechanism. Glennan uses the example of searing a steak to illustrate. When a steak is put on a hot skillet, energy is transferred to the steak (which is precipitating production) until the temperature of the steak rises above a certain limit (+140°C). At that point the constituents of the steak—various sugars and amino acids—react with each other in what is known as the *Maillard reaction* to produce a variety of molecules responsible for a range of pleasant flavours. So, the activities of the parts of the steak produce in those very parts a change to turn raw meat to seared delicacy.

Precipitating production is "the way in which one or more events produce another event […] by creating start-up conditions for a different mechanism" (2017: 182). The skillet transferring heat to the steak is an example of an event that creates the start-up conditions for the Maillard reaction in the steak.

Finally, *chained production* is a connection between events separated by a chain of intermediary precipitating events which in turn involve chains of constitutive production. We might say the connection between someone buying a steak and eating it, boils down to a series of precipitating events from grabbing the steak at the store, putting it in the carrier bag, carrying it home, putting it onto the skillet, searing it, placing it on the plate, and finally eating it.

There are certain points I puzzle over, regarding the distinctions between constitutive and precipitating production and their relationship to each other. Chained production, on the other hand, is straightforwardly understood as a connection between two non-adjacent events consisting of a chain of precipitating production, but precipitating production is not obviously constituted by constitutive production. Constitutive and precipitating phenomena, we are told, marks the difference between "mechanisms that produce phenomena (non-constitutive) and mechanisms that underlie phenomena (constitutive)" (Glennan 2017: 109). This could mean that the heat transmitted from skillet to steak produces the phenomenon of Maillard reaction, while the interactions between the constituents of the steak underlie the Maillard reaction. In that case, I would have guessed precipitating production would correspond to what Salmon calls 'production', i.e., an interaction between two distinct causal processes which produces a modification in them (skillet cools down, steak heats up), while constitutive production would correspond to the interactions that constitute the changes internally to each process when two processes interact; something Salmon does not have a separate name for other than simply 'modification'. However, Glennan says that constitutive production corresponds to Salmon's production and offers such examples as hammering a nail and a horse pulling a cart (Glennan 2017: 184). To my mind, hammering a nail is in the same category of phenomena as skillet searing a steak. In both cases we arguably have two separate and well-defined wholes interacting with each other—hammer and nail+board (skillet and steak)—which produces a change in each other but neither whole appear to underlie the changes in the other, not in the same way the constituents of the steak underlie the Maillard reaction. Similarly, the hammer is not an underlying component of the phenomenon of nail being driven into the board, but it is something that contributes to that change.

What I suspect we are seeing here, is the difficulty of applying a theoretical model to a many-layered and thoroughly complex reality. There is a problem in demarcating clearly between mechanisms on roughly the same level, and the many layers of mechanism within any one of them. Is hammer and nail and board separate mechanisms or are they all united by the person wielding the hammer and steadying the nail and board? And, underlying the amino-acids and fats that underlie the Maillard reaction are the atoms of the amino acids and fats whose interactions also must be constitutive of the wholes they are parts.

NMP and transmission accounts are similar in that NMP admits (or can admit) that transmission of conserved quantities is one of the ways we can understand what goes on in causal interactions, as for instance when energy flows from skillet to steak. NMP could even admit that in the domain of particle interactions, transmission of conserved quantities may possibly be all we need to understand what is going on. However, NMP claims to offer a more general model that can be applied to other scientific disciplines that do not operate with conserved quantities. Another difference between transmission and NMP accounts is that the former treats a mechanism as a nexus between distinct entities, whereas the latter treats a mechanism as a complex system. I take these differences to be large enough to treat transmission and mechanistic accounts as distinct theories despite the similarities.

Finally, like transmission accounts, NMP tend to characterise the influence exerted between distinct objects in precipitating production as unidirectional (from skillet to steak). However, the assumption of unidirectionality is much less pronounced in NMP, especially when it comes to constitutive production where it is not at all clear that interactions are assumed to be unidirectional.

4. Powers-based Accounts

Powers-based accounts (for instance, Martin 1997; Ellis 2001; Molnar 2003; Mumford and Anjum 2011; Heil 2012; Marmorodo 2017) are more heterogenous compared to transmission and mechanistic accounts, mostly because of a disagreement about the nature of powers. *Dispositional essentialists* accept the categorical/dispositional distinction—a distinction firmly rooted in the empiricist tradition—and thus say there is a difference between the properties that determine the qualitative state of the object at any given time, i.e., its *qualities*, and the properties that determine what the object is liable to do if certain conditions arise, i.e., its *powers/dispositions* (for instance, see Ellis 2001). Proponents of the *identity theory* reject the distinction and insist that the very same properties that determine the objects qualitative state also determine its ability to affect and be affected; they are *powerful qualities* (for instance, see Ingthorsson 2013).

Dispositional essentialists (for instance, Ellis 2001; Molnar 2003) tend to think of causation as a matter of pure potentialities (or unmanifested powers) being triggered by stimuli to transition from potentiality to actuality, and thereby manifesting some qualitative state; the *manifestation.* Identity theorists (Martin 1997; Heil 2012; Ingthorsson 2013) tend to think of causation as a matter of two powerful qualities mutually modifying their bearers from one powerful qualitative state to another powerful qualitative state; something called a *mutual manifestation*. There is a greater heterogeneity within each view, but for the purposes of this discussion I see no reason to delve into that. The two conceptions presented here, of how something is manifested, are so close that dispositional essentialists have found it easy to identify their manifestations with mutual manifestations, under the assumption that their manifestations are also a joint product of power and stimuli.

Now, most powers-based accounts, regardless of the differences mentioned, accept something equivalent to the distinction between *active* and *passive* powers that we find in Aristotle, the Stoics, the Scholastics, and the natural philosophy of the Early Enlightenment, i.e., between the ability to exert an influence on other objects/powers (active) vs. ability to change in response to an external influence (passive). Indeed, they will represent the direction of causation as the direction of the influence flowing from the object with the active power to the object with the passive power. A dispositional essentialist will say that a ball at rest can potentially move (it has the power to move while it is not actually moving), and that this potentiality will become an actuality when influenced by a ball in motion colliding with the ball at rest. The resulting motion will be a mutual manifestation of the active and passive potencies. Most identity theorist will say that what we perceive as a ball at rest is really a ball with an actual and determinate momentum, p, which is understood as a state of motion. If such a ball, a, with momentum p_a, collides with another ball, b, with momentum p_b, the two balls will mutually modify each other in accordance with the laws of motion, resulting in a transition of a from p_a to p_{a*}, and of b from p_b to p_{b*}. However, despite differences in the manner which dispositional essentialists and identity theorists characterise the collision between billiard balls, they will agree that one of the balls exerts an influence on the other, while the latter receives the influence. Like transmission accounts, most powers-based accounts characterise causal influence as unidirectional.

It is now time to introduce the powerful particulars view, but it is best to first point out what I think is problematic in the views already presented, because the powerful particulars view is best understood as an attempt to overcome those problems.

5. Influence: Unidirectional or Reciprocal?

Transmission, mechanistic, and powers-based accounts agree that causation is the exertion of influence of something on something else but disagree on the nature of the 'something' that exerts influence and about the nature of the influence being exerted. Most importantly, they all tend to treat influence as *unidirectional*. This is an overgeneralisation, as already hinted at when presenting NMP. Indeed, it will emerge that I am an identity theorist rejecting the unidirectional influence. But it is true enough for purposes of presentation. I'll qualify it later with respect to variations within each family of views.

In taking influence to be unidirectional the accounts under scrutiny are implicitly endorsing an understanding of influence that has been a part of the causal realist tradition since Aristotle wrote that "whenever the potential active and the potentially affected items are associated in conditions propitious to the potentiality, the former must of necessity act and the latter must of necessity be affected" (1998: 9, 5; 264). Indeed, we see a similar agreement about the unidirectionality of influence in the Aristotelian, atomistic, Stoic, Scholastic, and natural philosophy of the early modern period as the one we see today. They all share certain core notions, that together make up what I have elsewhere called *the standard view* (for further details, see Ingthorsson 2021: 45*ff*). The idea is that a new state is produced when an already existing entity, or complex of entities, changes due to an influence external to that entity, one without which the change would not have occurred, and the new state never exist. The external influence comes from an entity possessing powers that allow it to influence other entities (active), and that the entity upon which it acts possesses the power to receive the influence and change in some particular way (passive).

It is true that if we limit the discussion to the kind of examples that philosophers typically consider, the unidirectionality of influence and the distinction between active and passive entities appear to make good sense. Philosophers talk about billiard balls in motion acting on a ball at rest (Hume 1748: § 36); lead ball dropped onto a pillow produces a hollow (Kant 1787: A203); a locomotive pulls a truck (Taylor 1973: 35); a baseball is hit by a bat to fly across the pitch (Salmon 1980: 50). However, as Mario Bunge first pointed out (1959: ch. 6), it should be recognised as a serious problem that modern science categorically rejects the reality of unidirectional action even in the apparently asymmetric cases that philosophers typically consider. It is instead insisted that all influence comes in the form of *reciprocal action*, or 'interaction' as the term is defined in classical physics.

To be sure, classical physics is in many ways an outmoded framework, but we need to be careful in dismissing every component. While the first and second laws of motion are known to fail in extreme situations, the core idea expressed in the third law is still believed to hold good even in quantum and relativistic physics (albeit adapted to fit more generally to conservation laws), and it is the third law that is the basis for the rejection of unidirectional action. The core idea of the third law can be generalised in the following way:

> Whenever any object whatsoever exerts any influence whatsoever on any other object whatsoever, the latter exerts at the same time an influence of the same magnitude on the former, but in opposite direction.

While classical mechanics offer a fairly accessible treatment of the notion of reciprocal action, a full understanding of the concept still requires something of a shift of perspective. The notion is most clearly expressed in the third law of motion which says that the force by which object 1 acts on object 2 is equal to the oppositely directed force by which object 2 acts on object 1 ($F_{1on2} = -F_{2on1}$). However, because the prevailing understanding that influence is unidirectional has been such an integral part of the popular and philosophical understanding of causation for so long, the third law has always been widely misunderstood. Indeed, Newton's unfortunate decision to use the words action and reaction when explaining the law in plain English contributed to that misunderstanding. Let me explain.

The correct understanding of reciprocal action is not that of a phenomenon *composed* of two different kinds of actions, of which one (the action) gives rise to or provokes the other (the reaction)—as two tennis-players returning each other's strokes—but of a *single* phenomenon of mutual influence occurring simultaneously between two objects such as when bat collides with baseball. In Bunge's words "physical action and reaction are, then, two aspects of a single phenomenon of reciprocal action" (1959:153). Or, as physicists Resnick, Halliday, and Krane put it: "Any single force is only one aspect of a mutual interaction between two bodies" (2002: 83). In reciprocal actions, neither side has priority, and the terms 'action' and 'reaction' can be arbitrarily assigned to either.

However, researchers are not always interested in both sides of an interaction equally, and therefore often focus on the effect that matters to them, neglecting other outcomes. The point is stated beautifully by James Clerk Maxwell:

> The mutual action between two portions of matter receives different names according to the aspect under which it is studied, and this aspect depends on the extent of the material system which forms the subject of our attention. If we take into account the whole phenomenon of the action between the two portions of matter, we call it Stress […] But if […] we confine our attention to one of the portions of matter, we see, as it were, only one side of the transaction—namely, that which affects the portion of matter under our consideration—and we call this aspect of the phenomenon, with reference to its effect, an External Force acting on that portion of matter. The other aspect of the stress is called the Reaction on the other portion of matter (Maxwell 1877: 26–7)

The lesson to be learnt is that even if it is recognised that interactions are perfectly reciprocal, they are often treated, for the sake of convenience, as if they were instances of unidirectional flow of influence from one portion of matter to the other.

However, even if it is a part of the amassed scientific knowledge that interactions are reciprocal, misunderstandings are pretty common even among professional physicists. At least if we are to believe physics educators Steinberg, Brown, and Clement (1990) as well as Hellingman (1992). The main misconception they identify is the following:

> The 'inappropriate conception' (as Steinberg et al call it) here at stake is the same as in Newton's time: action and reaction are conceived as separate agents instead of as two sides of one interaction. Since they are conceived this way and since there are always two interacting bodies in producing a force, the suggestion of an action belonging to one body and a reaction belonging to another is virtually inescapable (Hellingman 1992: 112)

Anyone teaching classical mechanics faces the challenge of how to prevent students from falling under the spell of the 'virtually inescapable suggestion'.

But what are the philosophically interesting consequences that follow from the established fact that interactions really are reciprocal and not unidirectional? Like Bunge, I think we must accept that the polarization of interacting entities into an Agent that exerts an influence and a Patient that suffers a change upon receiving an influence "is ontologically inadequate" (Bunge 1959: 170–1). The reciprocity of interactions shows that there are no strictly passive entities—those who only receive influence but do not themselves influence other things—nor are there entities who influence other things without being themselves equally affected. Furthermore, since the mutual exertion of influence occurs simultaneously in equal magnitude in opposite direction between two objects, there is no way to give priority to one or the other as chiefly responsible for the outcome or even for initiating the interaction. It is because of this that the terms action and reaction are considered arbitrary and why "we are free to consider either of them as the force or the counterforce" (Hertz 1956: 185). If one comes first, we would not be free to do this. Ultimately, the conclusion is that in so far as any causal account assumes unidirectionality of influence, it is based on an empirical falsity in light of the current understanding of physical interactions.

There are two worries I expect many readers will find difficult to let go of. One arises from the fact that physical interactions are indeed so often described and understood in terms of transmission of quantities even in physics. Is physics both right and wrong, but in different ways, about how we should understand interactions? The other arises from the widespread use of the terms 'mutual' and 'reciprocal' in the literature about powers-based accounts. How exactly is reciprocal action as defined by physics different from mutual manifestations of disposition-partners as defined by powers-based accounts?

To answer the first question, then yes, physicists often apply the strategy Maxwell describes, notably to treat interactions as if only one side of the transaction mattered, such as when describing interactions in terms of transmission of conserved quantities. The inadequacy of this kind of characterization comes out clearest in the fact that while they often function fine for the intents and purposes of a given occasion they fail to generalize: they don't work for symmetrical interactions. When we turn our attention to the whole phenomenon, the transmission account only half-explains. As Maxwell points out, this is a choice of convenience we can safely make when the other side of the 'transaction' doesn't matter to us.

Consider a cue ball striking the black eight at a slight angle, pushing the black eight into a corner pocket while the cue ball continues in a new direction to end in position for the next shot. It seems easy to explain this by the transition of a quantity from the cue ball to the other, but then we are neglecting the details about what makes the cue ball change its trajectory in the way it did. Can we explain its change of trajectory merely by its loss of a quantity without bringing in any notion of an influence being exerted by the black eight on the cue ball? I at least find it difficult to entirely replace talk of forces with talk of transmitted quantities. You would have to make it so that the loss of the momentum *p* by the cue ball not only takes the role of the force usually meant to act on the black eight but also of the force presumed to act on the cue ball to change its trajectory. This is difficult to do even in asymmetrical interactions, but even harder in symmetrical interactions, say, when two cue balls move in opposite directions with the same velocity and collide head on. The outcome cannot be explained only in terms of something being lost by A which is gained by B. At least it must involve both A and B losing and gaining something.

An interesting example to consider is the common practice of explaining particle interactions by the transference of a virtual particle from one to the other, represented by vertices in Feynman diagrams. This is a way of conceptualizing what happens as if something carries a quantity from A to B. But when we consider that the term 'virtual particle' is not generally believed to denote actual particles in motion (Jaeger 2021), then the imagery of a transmission of something from one thing to the other appears to be misleading. Very simplified, a virtual particle is a technical notion used to signify the presence of certain theoretically calculated quantities considered to mediate interactions between real particles, and which obtain within time limits that are too narrow for anything to be observed. The idea is that the calculated quantities are of the kind that particles may have but it just isn't clear there are any, wherefore there is talk of a 'virtual' particle. The quantities kind of add up to something that might belong to a particle but may better be said to belong to a quantum field. However, if we were to assume there is a particle that actually moves between the two colliding particles, then in light of the reciprocity of interactions stipulated by the third law, then we would have to ask in what direction the particle is moving—from particle A to B, or vice versa—and whether the assumption of it moving one way, or the other, would really explain the whole interaction or only one side of the 'transaction'. Again, it is easier to think of asymmetrical interactions in terms of transmission of a virtual particle from A to B, than it is to think of symmetrical interactions in that way.

Turning now to the question of whether powers-based accounts already operate with an understanding of reciprocity when using terms like mutual manifestation of reciprocal disposition-partners. I think clearly not. According to physics, interactions are reciprocal in the sense that any two interacting things simultaneously influence each other and to the same magnitude. Powers-based accounts describe powers as reciprocal even when it is assumed that one thing influences and the other merely receives the influence. For them, to say that powers are 'reciprocal' is only an admission of the fact that two powers are always involved, and that both are considered equally important for the production of the outcome, but the outcome is typically only the change produced in the passive recipient. An example would be that the power of fire to heat and the ability of a hand to be heated both contribute to the end result of a hot hand. This is the traditional way of distinguishing between an active and passive powers as well as between Agents and Patients. The point is that if we accept what physics takes to be an established fact—that all interacting objects influence each other—it is no longer possible to argue that the distinctions between active vs. passive powers and Agent vs. Patient are based on the unidirectionality of causal influence. This realisation is the starting point of the powerful particulars view.

6. Powerful Particulars Account

I have now presented the core ideas of what must surely count as the main causal realist accounts in recent times, the most serious objection to them all, and some worries specific to each view. It is time to add my preferred account of causation in terms of reciprocal action between powerful particulars into the mix. It first appeared in the paper 'Causal Production as Interaction' (Ingthorsson 2002) but again in more developed form in a monograph (Ingthorsson 2021). The basic idea is that a single conception of causal interactions can account—in one and the same way—for (a) *changes* produced by interactions between previously unconnected entities, whether simple or compound, (b) *composition* of compound unities, (c) *persistence* of such wholes over time, and (d) that changes *internal* to a whole are just as causal as the changes resulting from interactions between wholes. It does so in a manner that easily relates to the theories and findings of the empirical sciences, in two ways. First, it is transparent how the philosophical account and the theories and data of the sciences offer complementary accounts on different levels; philosophy providing a general model applicable to all the specific phenomena, different scientific disciplines providing the details about each particular phenomenon. Second, it is clear how future developments in the sciences could falsify the general philosophical account. For instance, if physics ever finds truly asymmetrical interactions that violate the third law, my view is falsified.

Accounts of causation in terms of interactions between powerful particulars is no novelty. They are found in Aristotle, the Stoics, the atomists, the Scholastics, and the natural philosophers of the early enlightenment (for further discussion, see Ingthorsson 2021: ch. 3.7). In the 20th Century we find Dorothy Emmet (1985) and Ingvar Johansson (1989) defending this view. However, with the exception of Johansson's 'action by mixture' view, these earlier versions either assumed the unidirectionality of influence or construed the reciprocity of interactions more in the form of mutual manifestations and therefore are subject to the same kind of criticism levelled at powers-based accounts. More recently—indeed, while making final revisions of this paper—Andrew Newman has independently offered an account of the constitution of objects in terms of interactions between their constituent parts, one that appeals to our scientific understanding of how particles form atoms (Newman 2022). Newman doesn't address water dissolving salt, but his understanding of the collision of billiard balls (2022: 307–8) agrees exactly with the analysis I offer below, as well as with my earlier analysis of a brick hitting a window (Ingthorsson 2002: §3).

My powerful particulars view aims to offer a unified explanation of what Glennan calls precipitative and constitutive production (2017: ch. 7), what Craver and Darden think is needed to 'maintain' certain structures (2013: 65–6), what explains the production and destruction of the kind of entities that Salmon calls 'causal processes' as well as the unified wholes that NMP would denote as 'mechanisms'. Bear in mind that I only describe my view in enough detail to make comparisons between views meaningful (for full detail, see Ingthorsson 2021).

The powerful particulars view bears very little resemblance to transmission accounts, although it is meant to fit well to the scientific conception of the world as ultimately constituted by particles carrying properties of the kind described as conserved quantities. It resembles powers-based accounts, some more than others, to the degree that I think the best way to explain the behaviour of the entities that are parts of organized wholes (both why they constitute wholes and how they and the wholes behave in interactions) is to attribute powers to them. However, I prefer the conception of powerful qualities and I like powerful qualities to correspond to the natural properties that the sciences think we have good reason to think are real; I don't think powers come in addition to the natural properties, as something an object has 'in virtue of' its other qualities (Ingthorsson 2013). This still leaves place for plenty of emergent properties, i.e., properties of wholes that cannot be reduced to the sum of the properties of the parts.

I also prefer to think that it is the particulars that bear the properties that exert influence on each other, and which change, because I cannot make much sense of the idea that powers act on other powers to change the powers. Do we say a force makes a quality of velocity speed up from 10 to 16 km/hr? No, we say a force exerted between bodies makes a body accelerate from 10 to 16 km/hr. My reasons for preferring a 'particulars first' ontology are closely connected to my views on time, change, and identity, but this is not the place to elaborate (for details, see Ingthorsson 2002, 2016).

I am not convinced about the legitimacy of the distinction between active and passive powers or between agents and patients, at least not as they are drawn today. That is, as power to affect (active) vs. power to be affected (passive), whose possession determines whether an object is active (an agent) or passive (a patient). I might be persuaded to think there is a real distinction between active and passive powers of a single object, but not that they determine that some objects are active while others are passive. The only distinction between agents and patients that I am prepared to admit at present is the one between intentional agents and inanimate objects. If there is anything that can truly initiate an action, it would be an intentional agent.

I think my view resembles NMP the most, and if its proponents could be persuaded to think that the concept of powers need not be of unobservable mysterious potencies but simply of powerful qualities such as spin, charge, momentum, valency, etcetera—the kind of properties that are ubiquitous in the ontology of all the natural sciences—then I think they could warm up to a powerful particulars view. In the end, the resemblance is mostly that we seem to identify the same kind of phenomena as belonging to the class of relevant causal phenomena of which we need to develop a unified view, and this I believe is a consequence of the shared idea that the theories and findings of the empirical sciences is an important input, and that an agreement between the philosophical and scientific images of the world is desirable.

It is relevant to note that the powerful particulars arose from the realisation that for more than 333 years the philosophical implications of one of the most significant results of classical physics has either not registered as relevant in the minds of philosophers of causation or been misunderstood and dismissed. Mario Bunge (1959) is a notable exception. I am talking about the result that there are no unidirectional actions, only reciprocal interactions. To me it appeared foolish to challenge the validity of this result, because it would require me to show that our current physics is fundamentally wrong. It seemed more fruitful to explore the consequences of accepting the result as true (as already noted, I treat it as a provisional and falsifiable truth). Bunge had already argued convincingly that the reciprocity of interactions demonstrates that the prevailing idea must be wrong that causes are active objects or events that exert influence on some passive recipient, and that effects are the changes suffered by the recipient. This idea can only be considered an approximation (Bunge (1959: 151*ff*) calls it the 'causal approximation') which in many or most cases might be good enough given our particular explanatory interests, pretty much in the way Maxwell describes. However, Bunge comes to the conclusion that an account of causation based on reciprocal action is equal to a reduction of the asymmetric and productive relation of causation, to a non-productive and symmetric functional relationship of the kind Russell favours (1912). I disagree with this conclusion. Yes, interactions are symmetrical in the sense that two interacting objects influence each other simultaneously and to the same magnitude, but that influence produces a succession of states between which there is an asymmetric relation of one-sided existential dependence: one of producer to product.

Let me repeat that the kind of reciprocity we can see traces of in NMP and powers-based accounts is not drawn from the realisation that all interactions are reciprocal. It rather boils down to a realisation that even the scientific explanations that only pay attention to one side of an interaction—in a so-called 'causal approximation'—cannot entirely ignore the role of the perceived 'passive' recipient. To fully embrace the reciprocity of interactions as expressed by the third law, one need also embrace the idea that interactions are not composed of two actions belonging to two entities but is a single phenomenon of reciprocal action.

In the end, despite undoubtedly being a conceptual revolutionary, I think Bunge was not fully able to embrace the conclusion of his own argument (for details see Ingthorsson 2021: ch. 4.5). He insists that interactions are a single phenomenon of mutual action, but then rejects interaction as a basic principle for causation because it would involve the error of singling out the 'action' as cause and 'reaction' as effect and therefore characterise the relationship between cause and effect as symmetric, functional, and non-productive. I am still puzzled why he did not realise that his own argument suggested instead that we must accept that neither action nor reaction can count as cause, but that the interaction as a whole—being a single unified phenomenon—must count as the cause to whatever changes it brings about. On this view, it is not the baseball hitting a window that causes a breaking; that is only half the story. The interaction causes not just a breaking but a change in the state of the ball such that we end up with a ball at rest in a pile of broken glass. Similarly, the mutual action between billiard balls, regardless of their initial state of motion, causes quantitatively proportional changes in the state of both balls. It takes an equal amount of work to stop a ball as to move a ball.

It is often objected that the assumed reciprocity of interactions fails to account for the asymmetry exhibited by many interactions. The window breaks, but the ball does not. The standard explanation to asymmetries of this kind is simply that since the initial states of the two things are different, then one and the same influence exerted on them will lead to different changes. If you hit a window with a hammer it will break. If you hit a rubber ball in exactly the same way, it will not break. Instead, the hammer may bounce back to hit you in the face. The observed asymmetry is perfectly compatible with the reciprocity of interactions.

I will admit though that *if* a case is to be made for a distinction between active and passive powers and agents and patients—one that still respects the reciprocity of interactions—is to argue that in many cases the consequences on either side of the interaction must be considered in some sense graver for one than the other, even though they are otherwise proportional. Davis Kuykendall (*forthcoming*) argues just that, developing an earlier suggestion from Anna Marmodoro (2017). For instance, he suggests that enzymes, as biological catalysts, speed up chemical reactions without themselves being destroyed by the reaction (Kuykendall *forthcoming*: § 4.1). He argues in a similar vein, in §4.2 of the paper, that the interaction between H_2O and NaCl results in the destruction of NaCl but not H_2O. My answer is, first, that asymmetries of that kind do not violate the reciprocity of interactions; they can be explained by appeal to the differences in the initial states of the two interacting things. Second, that to find a handful of asymmetrical examples does not justify the conclusion that causal interactions are generally asymmetrical. It can at best show that sometimes interactions bear signs of some kind of asymmetry, but not of the kind that contradicts the third law of motion.

According to the powerful particulars view, causation is best described in terms of reciprocal action, and that the particular ways that entities can affect each other is determined by their intrinsic and powerful qualities; qualities whose determinate nature is best discovered by science. Indeed, science has found out that there are four different kinds of fundamental interactions, some of which are attractive while others are repulsive, some stronger than others, and they operate at different distances. We are talking about the four fundamental forces of nature. However, for all of them it is true—despite differences—that they exert their force reciprocally between interacting entities.

If causation is reciprocal action which sometimes comes in the form of attraction and sometimes repulsion, the lesson I draw is that philosophy has only been looking at a subclass of causal phenomena, notably where previously unconnected entities suddenly interact, often with very disruptive consequences; brick breaking windows, balls in motion disturbing balls at rest, etcetera. Philosophy has focused on this subclass because of the assumption that causation involves a unidirectional influence of something on something else and therefore essentially involves an *external* compulsion. Also, of course, because the interactions taking place between the component parts of objects could not be observed until quite recently. From the assumption that causation involves an external compulsion, it follows that anything happening inside an object, especially if it does not produce visible changes on the surface, is judged to be non-causal. Consider that the definition of spontaneous change is of a change that happens in the absence of an external compulsion. Even with the knowledge that only compound entities decay and knowing that compounds are held together by interactions between the component parts (which, on the hypothesis being considered, are causal), the idea that causation involves external compulsion has made it difficult to think of decay as a causal process. Hence it is called spontaneous decay. However, if causation is the phenomenon of reciprocal action, and such interactions occur within compound entities, then interactions between the parts of compound objects also are causal. Even if such interactions do not always produce visible disruption and change but instead appear only to maintain the inner structure of the entity, this is no reason to conclude that constitution and persistence are non-causal. Indeed, considering our current knowledge of the physical constitution of compounds it is misleading to say that they *maintain* an inner structure, at least if you interpret 'maintain' as simply remaining the same. As far as I know there are no interactions that do not involve some form of continuous change, even though many of them also preserve certain structures. All known compounds, beginning with protons and neutrons as compounds of quarks, are thoroughly dynamic entities in the sense that while they may stay the same structurally, they are still continuously changing on a more fine-grained level. Indeed, that change is what gives them their stability, which is why it makes sense to say that continuous reciprocal actions between the parts of a whole can *produce* stability, since that stability is dynamic. Just consider any atom of your choice. They are composed of a nucleus of protons and neutrons 'encircled' by a cloud of electrons. Every component part of that atom, as well as the atom as a whole, is continually changing. It follows that everything made up of such atoms is continually changing too. Consequently, such wholes are processes, if by process we mean any entity for which change is essential.

I hope I have now said enough to make it possible to compare the four different accounts with respect to how well they handle two different concrete cases: (i) collision between billiard balls, and (ii) water dissolving salt.

7. Case Study I: Collisions Between Billiard Balls

7.1 Transmission Accounts

The inadequacies of the transmission account when it comes to explain collisions between billiard balls has already been drafted in some detail above, so I'll be brief but still add a couple of details. I said that transmission accounts appear to make intuitive sense in cases like 'ball in motion acts on ball at rest', but not when two identical billiard balls moving at the same speed in opposite directions collide head on; the account doesn't generalise to fit all the cases.

Furthermore, it is relevant to point out that it is only under the assumption that the balls are perfectly rigid and friction against the table is ignored, that transmission is only between the balls. In the real world of billiard—considering only symmetrical collisions—the balls take away from the interaction an equal share of conserved quantities, but some energy is dissipated as heat and as sound. We now have at least three directions in which conserved quantities flow. Which of these directions represents the direction of causation? To say it is the direction that matters most to the player is to decide to only look at one side of the story for anthropocentric reasons. My suggestion is that a deeper account of how interacting particulars exert an influence on each other is needed, one which explains why the conserved quantities are distributed/changed the way they are when particulars interact. A mere description of the actual exchange of quantities doesn't answer that question. If you look at the full range of cases, there is no single direction in which quantities flow between entities.

Traditionally, the concept of 'force' has served as the explanation of how interacting particulars influence each other. However, in the friction between reductionistic empiricism and anti-reductionist rationalism, the concept of force has been just as controversial as the notion of power. Empiricists like Hertz (1956) and Mach (1919) wanted to get rid of the concept. To them the notion of force was an unnecessary postulate based on a redundant inference from observed changes to some imagined invisible cause to those changes. They instead wanted to describe interactions between material entities merely in terms of the changes in the state of motion that they can be observed to suffer. Accordingly, the second law of motion should not really be understood as saying that in any interaction there is this special thing, a force, that is *proportional in magnitude* to the object's mass times the acceleration it suffers, but as stating an *identity* of force and change in state of motion. Really, the third law of motion

($F_{1on2} = –F_{2on1}$) can then just as well be expressed by a reconstructed two-way second law; m_1 x a_1 = m_2 x a_2. In plain English, when two material systems interact, the observed change in the state of motion of the first (m_1 x a_1) is always equal in magnitude to the observed change in the state of motion of the second (m_2 x a_2). But, having effectively removed the notion of force from the equation, we no longer have an explanation as to what it is that causes the changes in the two material systems; we just have a description. Why should we accept a reduction of this kind? The empiricist answer is that we otherwise must appeal to the mysterious notion of force.

Contemporary proponents of transmission should not have any qualms about the concept of force, in so far as they seek to ground their view in the notions already in use in the natural sciences. Appeals to forces is ubiquitous in physics. However, to accept it, transmission theorists have to accept that there is a more fundamental feature to physical interactions than the transmission of conserved quantities, notably something (force/influence) that makes the quantities be transmitted in a particular way. Transmission accounts avoid commitment to the reality of forces and therefore end up being merely descriptive and in fact empirically inadequate when we consider the full range of interactions.

7.2 Mechanistic Accounts

The worries I have about mechanical explanations of billiard balls colliding, are, first, that it isn’t clear to me whether they would be treated as reciprocal (and therefore more like constitutive production) or unidirectional (and therefore more like precipitative production). The difficulty is partly to decide whether the change suffered by a ball in motion would be considered of ‘lesser’ importance than the change suffered by the ball at rest, say, because it matters more to the player that a ball at rest goes into a pocket. However, for an experienced player it is equally important to down a ball as it is to place the cue ball in position for the next shot. They must consider the effects on both balls equally. Is it the novice or experienced player that is best equipped to decide which side of the transaction matters? To my mind, the criteria for judging the direction of interactions in terms of ‘importance’ seem clearly anthropocentric and will be a case of deciding on the basis of the interest of individual players which side of the interaction they favour. At the very least, assuming the third law is valid, whether any change is ‘lesser’ in importance must be wholly unrelated to the purely physical magnitude of any change. We must then still accept the reciprocity of interactions but, like Kuykendall (*forthcoming*) and Marmodoro (2017), may perhaps look elsewhere for objective criteria for treating interactions as asymmetric. Kuykendall and Marmodo´s suggestion is that if *A* interacts with *B* with the result that *B* breaks but not *A*, then we have a kind of ‘directedness’. Perhaps not one that could be accounted for in terms of quantities, but still be objectively real. However, the handful of examples that might plausibly be considered asymmetric in this sense, would only show that some interactions are directed; we are short of a generalisable account of causation. More about that when discussing water dissolving salt.

Second, I am uncertain of how to apply the notion of mechanism to collisions between billiard balls, especially if we are to follow the idea that every interaction is an activity between parts of a mechanism. How exactly are the billiard balls part of a mechanism? I am worried that the senses in which they are part of a mechanism is either anthropocentric or turns out to include the entire universe. To be sure, the balls and table are designed to allow the balls to roll in certain restricted ways, for the purposes of the game. We also have the cue and the players, but where do we draw the lines for this particular mechanism? Is it only the table, cue, balls, and players, or is it also the room, the building, the planet, and solar system? There doesn't seem to be any particular physical bond between balls, table, cue, and players that organise them in any particular way in which they are not equally connected to the floor on which the table is standing, or the planet on which the building stands. The molecules of each individual ball are however clearly connected to each other in a way the ball is not connected to the table. But such bonds do not obtain between the balls, or between them and table. To be sure, the balls, table, cue, and players form a unity in our minds, but that would again introduce an unwanted anthropocentric feature. If proponents of NMP want the theory to be a contender among mind-independent theories of productive causation they must provide an account of how billiard balls are parts of a mechanism that doesn't rely on human cognition.

The question perhaps ultimately is whether it is unnecessarily confining to require that every causal interaction be between parts of an already existing organised whole. Can we not talk about interactions between previously unconnected entities? If the case of billiard balls colliding is not persuasive enough, consider the random collision between two pieces of rock drifting aimlessly in space. How can we understand their interaction as an activity of the parts of an organised whole? And if we can't, is their collision not causal?

An alternative is to ask whether previously unconnected entities may achieve some kind of unity when the interaction starts, such that they become parts of an organised whole during the interaction even if they were not so connected before. I have suggested that this is in fact what happens in interactions (Ingthorsson 2021: ch. 4). An interaction, on the realist stance taken here, is a substantial connection between two or more entities. This is clearest in cases when two entities attract each other to form a compound, but the same is the case in repulsion. Indeed, once interactions are accepted as the basic mechanism of causation, we have already accepted that interacting entities are parts of an organised whole. In any interaction—even between previously unrelated entities—the entities achieve a substantial connection in virtue of the forces they exert on each other (attractive or repulsive) and then form a unity of parts acting on each other, albeit the unity sometimes is very short lived. Indeed, the general understanding of interactions as a unified phenomenon rather than a composite of two separate actions implicitly support this understanding. Accordingly, two colliding billiard balls become a mechanism on contact and continue to be one as long as they are in contact. Indeed, on this view there really is no difference between precipitating and constitutive production, other than this: precipitating production is an interaction between previously unconnected entities while constitutive production is an interaction between already connected parts of a whole. Indeed, I have suggested that reciprocal action allows us to understand not just the changes that happen inside a steak when it is being seared as being causal, but also the material constitution of the steak during periods when the interactions between its constituent parts only serve to maintain the steak, for instance, while it is being carried home from the butchers (Ingthorsson 2021: Ch. 6).

In sum, my worries about NMP revolves around an uncertainty in the application of the model to particular cases, an uncertainty that arises partly because it hasn't been designed to take an explicit stance on the problems I raise here and elsewhere concerning the ago old view that exertion of influence is unidirectional (Ingthorsson 2002; 2021), and partly because NMP does not have an account of what it is for entities to constitute a mechanism that seems to cover all putative cases of causal interactions. As far as I can see, if NMP were to embrace my account of causation as reciprocal interaction as the fundamental feature of causation, both these problems would be solved.

7.3 Powers-based accounts

Powers-based accounts are more heterogenous than transmission and mechanistic accounts and therefore more difficult to present and criticise all in one go. One general worry already mentioned is that they tend to explain collisions, like transmission accounts, in terms of an exertion of influence of one ball on another which receives the influence, which is in conflict with the result that there are no unidirectional actions. Then there are worries connected to the specific conceptions of powers.

Dispositional essentialists characterise the powers of the balls in terms of potencies, i.e., properties that are not instantiated by the object until the change is being manifested in a collision. Accordingly, the ball at rest can potentially move, and the moving ball can potentially make another ball move. My worry is how to reconcile it with the scientific picture. According to science, the billiard balls are solid and have a shape that together allow them to roll. More importantly, they have at any given time a quantifiable and directed momentum, which is represented as a property that the balls have before the collision, and which is what supposedly makes the balls able to influence each other. None of these properties that figure in the scientific explanation are pure potencies but actual and occurrent properties of the balls. If there are any pure potencies in that picture, it would have to be something in addition to these determinate and fully realised properties. The dispositional essentialist must postulate that in addition to, and independent from a given balls actual momentum, it has a potency to change its momentum p_1 to another momentum p_2 by going continuously through the intermediaries. But on this view momentum is itself an inert property, but perhaps one from which emerges a power to change momentum.

My objection to dispositional essentialism is not that it is incoherent but needlessly complicated and does not easily link up with the scientific image, which is an ambition for many dispositional essentialists (for instance, Ellis 2001). Why not just accept, since you have accepted the reality of properties like momentum, that it is momentum that is responsible for both resistance to change (inertia) and ability to exert forces on other balls? Why insist that in addition to having momentum an object has a particular power for every kind of visible consequence of interactions between objects with momentum (power to make other balls move, power to change one's own state of motion, power to make a hollow in a pillow, power to pull trains, power to break windows), especially when you can't avoid appealing to momentum when explaining all these various consequences? Do we want to say that the ball broke the window because it had the power to break windows? These conceptual quirks are a worry in addition to the main problem, that powers-based accounts do not take into account the established fact that whenever any object whatsoever acts on any other object whatsoever, the latter always acts on the first in the same way to the same magnitude and at the same time.

The identity theory of powers identifies an objects power to change its own state of motion, as well as of other objects, with momentum. Hence it does not add to the set of natural properties defined by the sciences an infinite set of powers or dispositions corresponding to every distinguishable kind of behaviour. However, with the notable exception of myself (Ingthorsson 2002 and 2021) and John Heil (2012), proponents of the identity theory have not generally taken to describing interactions in any other way than in terms of mutual manifestations of reciprocal disposition partners where passive and active powers jointly contribute to a change in the object with the passive powers. I outlined what I think is wrong with that view in §5.

It bears to mention that both C. B. Martin (1993) and Mumford and Anjum (2018) are sceptical to the distinctions between active/passive and agent/patient and so do not characterise mutual manifestations in those terms. But they do not base this scepticism on the fact that whenever any object whatsoever acts on any other object whatsoever, the latter always acts on the first in the same way to the same magnitude and at the same time. In conclusion, powers-based accounts are by and large incompatible with the reciprocity of interactions, although this is not true of one or two alternatives. However, those who are compatible with the reciprocity of interactions are so for completely different reasons than the one's I outline in §5.

7.4 Reciprocal action between powerful particulars

According to the powerful particulars view, colliding balls each have their own directed momentum p_1 and p_2. Momentum is at the same time the power to resist changes in the state of motion and to change the state of motion of other balls. On contact the balls exert an equal and oppositely directed force on each other, that mutual exertion of influence being the cause to a change in their respective momenta from p_1 to p_{1*} and p_2 to p_{2*}. To repeat: the cause is the interaction as a whole and the effect is the sum total of changes suffered by the interacting entities. This formula applies to any collision regardless of initial state of any ball.

Note that there is an unresolved issue about the exact understanding of the nature of force. I lean to the understanding that forces are not properties of the balls either prior to collision or during the collision. To be exact, I do not consider them properties in virtue of which the balls exert influence; forces do not really push or pull, objects do. I understand the terms 'force', in this particular case, to denote the magnitude of the influence that balls with momentum exercise on contact. Similar understanding can apply to the force objects exercise in virtue of various other powers, e.g., those related to charge, spin, etc. In this sense the concept of force is an abstraction, not the least in light of the fact that forces only arise in interactions, and the fact that interactions are not considered to be composed of two separate entities, the action and reaction. However, by saying the concept is an abstraction I am not saying it does not relate to a real phenomenon. The real phenomenon is reciprocal action, the mutual exertion of influence between two entities, and the abstraction is the conceptualisation of reciprocal action as having two sides because it affects two (or more) entities. My understanding of forces comes close to that of Johansson (1989: 167–8) and Massin (2009), in that they treat them not as properties of an object but as something real that essentially holds symmetrically between objects, but I am uneasy about understanding reciprocal action as a relation. If it is a relation it is a very special relation since it is an efficient relation and so more like an activity or process. Note that Massin, like Bunge, takes mutual forces to be non-causal relations because he thinks of them as symmetrical and therefore cannot be productive, and yet he thinks they are relata of production. My view also has affinity with Jessica Wilson's view that forces are aspects of the objects that exert them and therefore not something in addition to the object and its properties, although I am uneasy with her description of the objects as non-causal entities (Wilson 2007). As I mentioned earlier, Newman has recently offered an analysis of the collision of billiard balls that coincides with mine (Newman 2022: 307–8).

8. Case Study II: Water Dissolving Salt

The scientific explanation of water dissolving salt appeals, first, to the properties of H_2O and NaCl molecules. Molecules of H_2O are covalent dipole compounds (have negatively and positively charged poles). NaCl is a nonpolar ionic compound. When molecules of H_2O and NaCl come into contact there will be attraction between H_2O and either the Na or Cl in NaCl, depending on the spatial orientation of H_2O (negative pole attracted to positive ion, positive pole to negative ion). The covalent bond between O and H in water is stronger than the ionic bond between Na and Cl, which is why the tug of war between the various compounds (Na and Cl also attracting each other) ends in the breaking of the ionic bond of NaCl but not the covalent bonds in H_2O. The dissolution continues as long as individual water molecules can interact with individual NaCl molecules. Important to note that the mutual attractions between the parts of each compound are all considered to be reciprocal.

8.1 Transmission accounts

Transmission accounts are bound to explain the dissolution of salt in water in terms of transmitted quantities, but I have been unable to find in the philosophical literature any attempt to do this. Transmission accounts typically chose examples from the domain of thermodynamics such as the heating or cooling of water (for instance, see Fair 1979; Kistler 1998). I myself find it difficult to see an explanation of water dissolving salt solely in terms of transmission of conserved quantities.

The standard explanation of why and how water dissolves salt—the one given above—appeals to electrical charges and electrostatic interaction. One can however come across thermodynamic accounts of solubility, some of which give the impression of *explaining* solubility rather than just *describing* the thermodynamic aspects of dissolution. They claim to explain what happens in terms of systems striving for equilibrium. However, in so far as they only appeal to least-energy principles or say that NaCl breaks because that requires less energy than H_2O breaking, they turn out to be half-explanations. The reason that the breaking of NaCl is the most energy-efficient outcome is because in water covalent bonds are stronger than ionic bonds, wherefore it takes more energy to break the covalent bonds. But the relative strength of those bonds is not decided by transmission of conserved quantities.

In the end it seems difficult to explain attractions of any kind *solely* in terms of transmission of conserved quantities or a strive towards equilibrium. Indeed, as Tracy Lupher argues, it is even difficult explain any kind of static interactions in terms of transmitted quantities (Lupher 2009). I think his result is really the same as the more general conclusion I have reached that transmission accounts fail for symmetric interactions, whether dynamic or static.

The criticism here is that even in the domains of physics and chemistry that operate with conserved quantities, transmission accounts can at best only be applied with some plausibility to asymmetric interactions, but then only as approximations.

8.2 Mechanistic accounts

Mechanistic accounts can easily be applied to explain the dissolution of salt by water. All the interactions taking place are either between the parts of organised wholes, or between organised wholes that consequently morph into other kinds of organised wholes where interactions continuously preserve the whole. My only complaint is that NMP does not really address the question of unidirectionality vs. reciprocality of the interactions, wherefore it appears NMP can treat interactions sometimes as reciprocal when a given scientific explanation clearly tells us so, and sometimes as unidirectional when that is suggested by the scientific account. That raises the question whether the deciding factor of whether any given interaction is treated as unidirectional or reciprocal hinges on whether the scientific explanation for that particular interaction is considering the whole phenomenon or is only concerned to explain one side of the transaction. This ambiguity derives, arguably, from the fact that its proponents have not been aware of the problems I have raised here and elsewhere, and so have not taken any measures to respond to it.

8.3 Powers-based accounts

Powers-based accounts explain water dissolving salt in two different ways, depending on whether they relate only to the manifest or also to the scientific image. According to the former approach, the power to dissolve salt is attributed to water as a body of matter, and to salt the power to be dissolved. Accordingly, when salt is put in water the two powers mutually manifest the dissolution of the salt. This kind of explanation is open to two objections: (i) that it presents causation as a phenomenon involving unidirectional influence of the kind the natural sciences say does not exist, and (ii) that it simply does not even address the fact that the scientific explanation of the phenomenon ties the ability to dissolve not to water as a body of matter, but to the properties of individual H_2O and NaCl molecules. That raises the worry that these accounts at best relate to the way we ordinarily think about water dissolving salt, but not to what really happens.

The second approach takes the scientific explanation as its starting point and says that covalent dipole H_2O molecules have the power to break the ionic bond in NaCl molecules (Marmodoro 2017; Kuykendall *forthcoming*). When such particles interact, the power to break and the power to be broken manifests the breaking of NaCl. This explanation is also open to the objection that it assumes interactions are unidirectional and is therefore in conflict with the third law. However, it does *ground* the powers of water and salt on the physical properties of the molecules, and therefore cannot be said to ignore the scientific explanation. On the other hand, the account does not *identify* the powers with the physical properties recognised by the sciences. The assumption is that in addition to H_2O having one slightly positively charged end, and one slightly negatively charged end, which gives the molecule the ability to interact electrostatically with other charged molecules (mutually attract or repel), it also has the more specific power to 'break NaCl'. As far as I know, chemistry does not postulate any such specific property in addition to the already mentioned physical properties, but it acknowledges that for NaCl to break is a known consequence of electrostatic interaction between H_2O and NaCl. Importantly, it is clear that the properties of H_2O alone are not enough to ground its power to break NaCl. It can only be said to have that power with respect to the specific properties of the NaCl, notably that it is an ionic compound. Indeed, power-based accounts rarely explicitly address the forces of push and pull operating between the molecules, but only on the consequences of the forces exerted—NaCl breaks—and then say that the strength of the covalent bond, and/or the strength of the ion-polar electrostatic interaction between water and salt, gives H_2O the power to break.

I should address explicitly that Anna Marmodoro (2017) and Davis Kuykendall (*forthcoming*) complain that my account of causation in terms of reciprocal action "fails to capture the directionality of the causal process, which is underpinned by the different 'actions' of salt on water and water on salt respectively" (Marmodoro 2017: §7). On the one hand, it is true that I say that molecules exert an equal and oppositely directed force on each other, and that the interaction is in that sense symmetrical. On the other hand, I say that this symmetry allows for very different outcomes for each interacting entity in so far as they are different from the outset. One and the same kind of influence can result in different types of changes in each interacting entity. In other words, we can talk about reciprocal actions as always being symmetric in terms of magnitude of influence and yet distinguish between symmetric and asymmetric interactions when talking about (i) interactions between two similar entities that produce similar changes in both, and (ii) interactions between two dissimilar entities that produce different changes in each. My account therefore perfectly well explains why NaCl but not H_2O is destroyed in the interaction, and apparently without having to appeal to the kind of directionality that Marmodoro and Kuykendall are talking about. Indeed, I argue we should not appeal to it if we are to arrive at an account that applies to the full range of interactions and in a manner consistent with the third law. My argument doesn't show that there is no way to consider one side of the transaction as 'graver', only that previous ways of motivating that conclusion conflicts with established scientific facts.

I can in turn complain that Marmodoro and Kuykendall's accounts of water dissolving salt bear some signs of the kind of half-explanations I have mentioned before. To be fair, Kuykendall recognises that the breaking of NaCl is not the only outcome. Another outcome is the production of two different kinds of formations, which significantly alter the properties of the resulting liquid. Several molecules of H_2O will surround each Na^+ and Cl^- ion because of strong ion-dipole interactions between them to form what is called a 'hydration shell'. We now really have a liquid that no longer is made up only of H_2O molecules connected by hydrogen bonds, but a mixture of molecules of which some are connected by hydrogen bonds but others by ion-dipole bonds. All of this, I argue, can be explained by reciprocal actions between the parts in the liquid. The point is that Kuykendall's explanation, detailed as it is, is not complete in relation to the whole phenomenon since it focuses almost entirely on the 'water dissolves salt' aspect, rather than 'water and salt merge to form saline'.

Mumford and Anjum indeed take the position that we shouldn't think in terms of water dissolving salt but instead that salt and water together produce saline, and similarly that sugar and water produce a sweet solution (2011: 123). However, while much more fully recognizing the reciprocity of interactions it is not because they accept the kind of reciprocity I have been arguing for here, or even are aware of the problem stemming from the rejection of unidirectional action, and their account still retains a notion of 'agency', i.e., of something being the pivotal trigger to a change that upsets a state in equilibrium, say, a match being struck in a place with oxygen and flammable material. Indeed, they don't think so called 'countervailing' powers should be included in the notion of cause (2011: 33–4). So, while we often come to similar conclusions, our accounts are decidedly different.

8.4 Powerful particular account

According to the powerful particulars view, the dissolution of salt by water can be explained fully in terms of reciprocal interactions between the component parts of water and salt. The mutual electrostatic attraction between H_2O molecules on the one hand, and the components of NaCl on the other, results in separation of NaCl into Na^+ and Cl^- ions and the formation of hydration shells around each ion. The result is saline. Not only does the powerful particulars view agree in this way to the scientific understanding of 'water dissolves salt' but also of the formation of H_2O and NaCl molecules respectively, as well as of the constellations of Na^+ and Cl^- ions surrounded by a hydration shell. It does so without postulating a distinct second order class of dispositions/powers. It is all down to a question of who is the winner in various tugs of war between different entities, tugs of war whose strength is determined ultimately by the different strengths of the fundamental interactions on which they are based. Indeed, reciprocal interaction is a viable candidate for being what can give rise to the kind of organised wholes that mechanistic philosophers call 'mechanisms'.

9. Conclusion

In sum, I have argued that transmission, mechanistic and powers-based accounts do not offer a generalizable explanation of physical phenomena, mainly because they either assume that all influence is unidirectional or open to it being both unidirectional and/or reciprocal. Transmission accounts appear unable to account for mutual attractive influence and so are unable to account for the creation of the kind of organised wholes that would fit the description of mechanisms. Indeed, mechanistic accounts also struggle to explain how previously unconnected entities could become parts of a mechanism, and so struggle to explain how interactions between such entities would count as causal. I have suggested that my account of causation in terms of reciprocal action would resolve that problem.

It bears to mention that there is one salient feature of the world that I haven't worked out how, or whether, the powerful particulars account could explain, and which may perhaps be the main reason Bunge, Marmodoro and Kuykendall are persuaded that all interactions cannot really be reciprocal and that my account fails to account for some kind of causal directionality present in the world. I am talking of the kind of feedback loops that we find in biological systems. Or, really, self-organising structures whose parts act on each other in a manner that suggest linear progression in a certain direction. The Krebs cycle is a good example. Every particular interaction that takes place in the cycle appear to be reciprocal and yet the process as a whole is directed in such a way that it repeats the same pattern again and again. Indeed, this would apply to the explanation of the role of enzymes as catalysts, as Kuykendall mentions (forthcoming), to maintain a directed process. An explanation of that kind of direction is needed, but I doubt it will come in the form of overthrowing the third law of motion and vindicate the active/passive or Agent/Patient distinctions. One possibility is that this kind of directionality can be grounded partly on the directionality of the relation between the successive states of any organised whole of reciprocally acting parts, A—the one-sided existential dependence between producer and product—but which will have to be constrained somehow by the structure of a larger organised whole B of which A is a part. To work out the viability of that idea is work for the future.

Acknowledgements

I am very grateful to Yafeng Shan, Ingvar Johansson, and an anonymous referee, whose comments and suggestions decidedly helped to improve on the initial draft of this paper.

References

Aristotle (1998), *Metaphysics*, H. Lawson-Tancred (trans.), London: Penguin.

Bechtel, W. and Adele Abrahamsen (2005). 'Explanation: A Mechanistic Alternative', *Studies in History and Philosophy of the Biological and Biomedical Sciences* 36: 421–41.

Bogen, J. (2008). 'Causally Productive Activities', *Studies in History and Philosophy of Science*, 39: 112–23.

Bunge, M. (1959). *Causality and Modern Science*, New York: Dover. Rpt. 1979.

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

Craver, C. F. and Lindley Darden (2013). *In Search of Mechanisms: Discoveries across the Life Sciences*, Chicago: University of Chicago Press.

Darden, Lindley (2008). 'Thinking Again About Mechanisms', *Philosophy of Science* 75 (5): 958–69.

Dowe, P. (1992). 'Wesley Salmon's Process Theory of Causality and the Conserved Quantity Theory', *Philosophy of Science* 59: 195–216.

Dowe, P. (2009). 'Causal Process Theories', in Helen Beebee, C. Hitchcock & P. Menzies (eds.), *The Oxford Handbook of Causation*, Oxford: Oxford University Press.

Ellis, B. (2001). *Scientific Essentialism*, Cambridge: Cambridge University Press.

Emmet, Dorothy (1985). *The Effectiveness of Causes*, Albany: SUNY Press.

Fair, D. (1979). 'Causation and the Flow of Energy', *Erkenntnis* 14: 219–50.

Galavotti, Maria Carla (2022). 'Wesley Salmon', in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, available online: https://plato.stanford.edu/archives/fall2022/entries/wesley-salmon/.

Glennan, S.S. (2009). 'Productivity, Relevance and Natural Selection', *Biology and Philosophy* 24: 325–39.

Glennan, S.S. (2017). *The New Mechanical Philosophy*, Oxford: Oxford University Press.

Heil, J. (2012). *The Universe as We Find It*, Oxford: Oxford University Press.

Hellingman, C. (1992). 'Newton's Third Law Revisited, *Physics Education* 27(2): 112–5.

Hertz, H. (1956). *The Principles of Mechanics*, New York: Dover.

Hume, D. (1748). *Enquiry Concerning Human Understanding*, Harvard Classics, Vol. 37, Boston: MA: P.F. Collier & Son, 1910.

Ingthorsson, R.D. (2002). 'Causal Production as Interaction', *Metaphysica* 3: 87–119.

Ingthorsson, R.D. (2013). 'Properties: Qualities, Powers, or Both?', *Dialectica* 67 (1): 55–80.

Ingthorsson, R.D. (2016). *McTaggart's Paradox*, New York: Routledge.

Ingthorsson, R. D. (2021). *A Powerful Particulars View of Causation*, New York: Routledge. *Open Access*: https://doi.org/10.4324/9781003094241.

Jaeger, G. (2021). 'Exchange Forces in Particle Physics', *Foundations of Physics* 51:13 https://doi.org/10.1007/s10701-021-00425-0

Johansson, I. (1989). *Ontological Investigations: an Inquiry into the Categories of Nature, Man and Society*, New York: Routledge. Rept. 2004 by Ontos Verlag. *Open Access*: https://www.degruyter.com/downloadpdf/title/303559.

Kant I. (1787). *Critique of Pure Reason*, N. Kemp Smith (trans.), New York: St. Martin's Press, 1965.

Kistler, M. (1998). 'Reducing Causality to Transmission', *Erkenntnis* 48: 1–24.

Kitcher, P. (1989). 'Explanatory Unification and the Causal Structure of the World', in P. Kitcher & W.C. Salmon (eds.), *Scientific Explanation*, Minnesota Studies in the Philosophy of Science, vol. 13. Minneapolis: University of Minnesota Press: 410–506.

Kuykendall, D. (*forthcoming*). 'In Defense of the Agent and Patient Distinction: The Case from Molecular Biology and Chemistry', *British Journal for the Philosophy of Science*: https://www.journals.uchicago.edu/doi/abs/10.1086/715470.

Lupher, T. (2009). 'A physical critique of physical causation', *Synthese* 167: 67–80.

Mach, E. (1919). *The Science of Mechanics: A Critical and Historical Account of its Development*, T. McCormack (transl.), Chicago: The Open Court Publishing Company.

Machamer, P. (2004). 'Activities and Causation: The Metaphysics and Epistemology of Mechanisms', *International Studies in the Philosophy of Science* 18(1): 27–39.

Marmodoro, Anna (2017). 'Aristotelian Powers at Work: Reciprocity Without Symmetry in Causation', in J. Jacobs (ed.), *Causal Powers*, Oxford: Oxford University Press, 57–76.

Martin, C. B. (1997). 'On the Need for Properties: the Road to Pythagoreanism and Back', *Synthese* 112: 193–231.

Martin, C. B. (1993). 'Power for Realists', in K. Cambell, J. Bacon and L. Reinhardt (eds), *Ontology, Causality, and Mind: Essays on the Philosophy of D. M. Armstrong*, Cambridge: Cambridge University Press: 175–83.

Massin, O. (2009). 'The Metaphysics of Forces', *Dialectica*, 63(4): 555–89.

Maxwell, J.C. (1877). *Matter and Motion*, Dover, New York.

Molnar, G. (2003). *Powers: A Study in Metaphysics*, Oxford: Oxford University Press.

Mumford, S. & Anjum, Rani (2011). *Getting Causes from Powers*. Oxford: Oxford University Press.

Mumford, S. & Anjum, Rani (2018). 'Powers and Potentiality', in K. Engelhard and M. Quante (eds), *Handbook of Potentiality*, Dordrecht: Springer: 261–78.

Newman, A. (2022). 'A causal ontology of objects, causal relations, and various kinds of action', *Synthese* 200(4): article 308. https://doi.org/10.1007/s11229-022-03752-5

Resnick, R., Halliday, D., & Krane, K. (2002). *Physics* (5th ed.), New York: John Wiley & Sons.

Russell, B. (1912). 'On the Notion of Cause', *Proceedings of the Aristotelian Society*, New Series 13: 1–26.

Salmon W. (1980). 'Causality: Production and Propagation', in Sosa, E. and Tooley, M. (eds.), *Causation*, Oxford University Press, Oxford, 154–71, 1993.

Salmon, W. (1984). *Scientific Explanation and the Causal Structure of the World*, New Jersey: Princeton University Press.

Steinberg M. S., Brown D. E., and Clement J. (1990). 'Genius is not immune to persistent misconceptions' *Int. J. Sci. Educ.* 12: 265–73.

Taylor, R. (1973). *Action and Purpose*, New York: Humanities Press.

Wilson, Jessica (2007). 'Newtonian Forces', *Brit. J. Phil. Sci.* 58: 173–205.