

DISCUSSION

Top-down causation and social structures

Dave Elder-Vass*

Department of Social Sciences, Loughborough University, Loughborough, Leicestershire LE11 3TU, UK

Top-down causation has been implicit in many sociological accounts of social structure and its influence on social events, but the social sciences have struggled to provide a coherent account of top-down causation itself. This paper summarizes a critical realist view of causation and emergence, shows how it supports a plausible account of top-down causation and then applies this account to the social world. The argument is illustrated by an examination of the concept of a *norm circle*, a kind of social entity that, it is argued, is causally responsible for the influence of normative social institutions. Nevertheless, social entities are structured rather differently from ordinary material ones, with the result that the compositional level structure of reality implicit in the concept of top-down causation has some limitations in the social world. The paper closes by considering what might be involved in examining how top-down causation can be shown to be at work in the social domain.

Keywords: downward causation; emergence; norm circles; social ontology; social structure

1. TOP-DOWN CAUSATION AND SOCIAL STRUCTURES

The question of *top-down*, or *downward*, causation has always been central to debates over the concept of social structure in the social sciences, though these terms themselves have rarely been used. Instead, this issue is known as the problem of structure and agency: how can we reconcile claims for the causal effectiveness of social structure (the top level, in this case) with our belief that individual humans (the lower level, in this case) have the capacity of agency, the capacity to have a causal influence of our own on the world? Traditionally, this debate has been presented as one between methodological collectivists or sociological holists, on the top-down side, and methodological individualists on the other. The holists, according to the usual presentation, argue that social events, including at least some aspects of individual behaviour, are fully determined by social structure. For holists, our sense of individual agency is merely an illusion, itself produced by structural forces. The individualists argue that social events are produced entirely by human individuals in their relations with each other, and hence that it is the idea that social structure is causally effective that is an illusion, or an ‘actor’s fiction’.

This is, however, a somewhat caricatured way of presenting the debate, and in practice many of the more

recent contributions to it (such as those of Giddens, Bourdieu and Archer) have offered competing accounts of how *both* structure and agency may be significant ([1–3], [4], ch. 1). In advocating an emergentist account of top-down causation by social entities, this paper is firmly aligned with this tradition, and particularly with Archer’s version of it: although its focus is on the ways in which social entities may influence individuals and events, it also recognizes that human individuals do have emergent causal powers of their own, which co-determine social events. Where, for example, the structural pressures on an individual are in conflict with each other, the individual may have to choose between multiple possible courses of action that are endorsed by different structural forces ([4], ch. 8).

In recent years, the concept of emergence has become increasingly popular in thinking about social structure. Perhaps surprisingly, however, the concept of top-down causation has often been neglected or even rejected outright by advocates of social emergentism. This is potentially problematic if, as Kim has argued, emergentism logically implies downward causation, and the viability of emergence theories therefore depends upon reconciling them with a theory of downward causation ([5], p. 3, [6], p. 121). This paper will argue that emergence and top-down causation can be reconciled, and that both play important roles in the social world. This argument entails a degree of integration of the social with the domains of the natural sciences, both in the sense that we may be able to apply similar ontological frameworks to both, and in the sense that there

*d.elder-vass@lboro.ac.uk

One contribution of 15 to a Theme Issue ‘Top-down causation’.

is an ontological dependence of people—who I take to be the building blocks of higher social structures—on the biological (and thus physical) structures of which they are composed.

But ontological relations are sensitive to material particularities, and this paper will identify one such particularity that places limits on the applicability of top-down causation to the social. Because of the nature of the relations between human individuals that constitute them into social entities, the metaphor of levels breaks down in relations between social structures above the level of the human individual. Top-down causation, in consequence, does not operate directly between different social entities, but top-down causation remains important, as interaction between those entities depends upon their downward effects on the humans that are their parts.

The paper begins with a brief survey of attitudes to top-down causation among sociological emergentists, then moves on to develop a substantive account of the phenomenon, beginning with a relatively abstract discussion of emergence and top-down causation. The next section moves to a more specific discussion of the social domain, offering an account of how social entities such as organizations can have emergent causal powers, and the implications of this for top-down causation between social entities and their human parts. Finally, it turns to the question of relations between social entities, and the need to apply a rather different understanding of levels, and thus top-down causation, to this class of cases.

2. TOP-DOWN CAUSATION IN SOCIAL ONTOLOGY

In recent years, numerous scholars have suggested that emergence may be significant for understanding ontological questions in the sociological field. There is even a journal dedicated to pursuing this programme (*E:CO—Emergence: Complexity and Organization*, formerly known as *Emergence*). There is no space for a thorough survey of the field here, but a brief examination of some key contributions will reveal the rather reserved attitude to top-down causation that has been characteristic of much of this literature. This section will discuss the work of the sociological theorist Margaret Archer, the educational psychologist R. Keith Sawyer and the institutional economist Geoffrey Hodgson, each of whom, in rather different ways, has made a significant contribution to the study of emergence in social ontology.

Archer, who has aligned herself with the critical realist tradition of social philosophy (as I do), relies heavily on the concept of emergence to justify her social ontology. For Archer, social events are produced by the interaction of personal, cultural and structural emergent properties, and it is essential for an adequate understanding of the social world that we recognize, for example, that structural emergent properties (such as, for example, educational systems or the division of labour in economic production) pre-exist any given case of social action and exert an influence on it, while being reproduced (or transformed) as a

consequence of these social actions [1,7]. Archer therefore aligns herself with a long-standing tradition of sociological thinking by asserting that social structure can have a downward influence on the acts of human individuals, while recognizing that those acts in turn act back on social structure, reproducing and/or transforming it. This produces what Archer calls the *morphogenetic cycle*, in which structures influence actions, which then in turn reproduce or transform those very structures. However, Archer's focus is very much on the application of her system to sociological questions and although she does devote some attention to the nature of emergence ([1], p. 9, [8], p. 475, [9], p. 87), she says little about the concept of top-down causation as such. For Archer, the pre-existence of social structures (as the outcome of earlier cycles of action) is enough to justify the claim that they can influence individuals ([1], p. 76, [10], p. 19) (an argument that has been criticized by a number of authors ([11], p. 227, [12,13], p. 211, [14], pp. 83–84). What is less clear is exactly what form these social structures take and just how the part–whole relationship characteristic of emergent systems is structured.

Sawyer, by contrast, bases his account of social emergence, which he labels *non-reductive individualism*, on Fodor's functionalist philosophy of mind [14,15]. Sawyer combines the claim that only individuals, and not social entities, are real (i.e. causally effective) in the social world with the denial of methodological individualism, on the grounds that social *properties* are irreducible to properties of individuals ([16], p. 541). This is a move that separates explanatory considerations from causal powers. His argument implies that we can invoke social (structural) properties in explanations, but in any particular case it is always the human individuals that are doing the real causal work. This difficult balancing act is achieved by invoking the concept of *multiple realizability*. Multiple cases of the same type of social structure, it is argued, may be based on diverse combinations of low-level properties, and yet have a similar causal impact. Religious organizations, for example, tend to produce solidarity between their members, even though they are highly diverse in the beliefs they endorse and the ways in which they are organized. In each case, it is the low-level properties (the human individuals) that are taken to do the real causal work, but because these contributions vary from case to case, they do not explain the similarity of the causal impact of the higher level property across its multiple configurations. We are therefore entitled to construct explanations that invoke the higher level property to account for these similarities of causal impact (social laws), while still insisting that it is the lower level that is really causally effective. From this perspective, we could arguably construct top-down *explanations* of social events, in which social laws are used to explain individual behaviour, but top-down *causation* as such would be ruled out, since in any given case, it would always be a particular set of lower level properties that really possessed the causal power.

Perhaps the clearest endorsement of top-down causation in the social science literature can be found in the work of Hodgson. Discussing normative institutions,

for example, he argues that ‘institutions have the power to mould the dispositions and behaviours of agents in fundamental ways’ ([17], p. 107), acting on ‘individual habits of thought and action’ in a process that he calls ‘reconstitutive downward causation’ ([17], p. 108). In some respects, Hodgson’s framework is similar to Archer’s (and indeed many others in the social sciences) by virtue of seeing a kind of cyclical interaction between social structures influencing human action and that human action in turn shaping social structures. He is more explicit than Archer, however, about the potential for social structures to exert a downward causal effect on the actual nature of individuals—their deeply ingrained habits or dispositions—rather than simply providing a recognizable environment to which the individual consciously responds. Habits, he says, change with experience, and hence enduring social institutions ‘reconstitute “downwards” the preferences of the agent’ ([17], p. 109). And he is more explicit about top-down causation as a process, invoking Sperry, for example, to stress its compatibility with lower level causal processes ([17], p. 109). Regrettably, however, Hodgson [18] has recently retreated somewhat from this position. Although he claims that the substance of his argument is unaltered, he now prefers the term ‘reconstitutive downward effects’, invoking Craver & Bechtel [19] to suggest that ‘causal relations are exclusively intra-level’ and thus that downward relations cannot be causal.

The social sciences, one might suggest, are having some difficulty in coming to terms with the top-down causation that has always been implicit in the concept of social structure.

3. CAUSALITY AND EMERGENCE

One prerequisite of making sense of top-down causation is an understanding of causality in general. This paper builds on the account of causality and emergence developed by the critical realist philosopher Roy Bhaskar [20] and which I have expanded elsewhere ([4], ch. 2 and 3). For Bhaskar, events are caused by the interacting causal powers of things (I shall say *entities*), and these causal powers are emergent properties of the entities concerned, produced by generative mechanisms in which the parts of the entity interact in a process that depends on the relations they have with each other in an entity of this kind. Let me expand just a little on this argument using a partial explanation of a simple non-social example. A particularly simple example is used in order to isolate the ontological arguments as clearly as possible. The event to be explained is the light starting to shine when a person turns on a torch.

In this case, causation results from an interaction between the causal powers of two entities: the person and the torch. The powers of the torch include the power to shine a light in a focused beam when it is switched on, and the powers of the person include the power to press a switch. The actual event of the torch starting to shine is thus produced by an interaction between these two powers.

In both the case of the torch and that of the person (and indeed in entities in general), each causal power

depends on the entity’s composition (parts) and structure (the relations in which its parts stand to each other). Such causal powers are held by all tokens of the relevant type, because anything with the composition and structure characteristic of this type of entity will have the potential for its parts to interact in the way necessary to produce the effect. Everything with the characteristic composition and structure of a torch, for example, will at least potentially have the power to shine a light. Such powers, however, are not necessarily realized on all occasions. In this example, the switch might not be pressed, or when some other non-conducting object is inserted into the torch it might interrupt the electrical circuit.

Each of these causal powers is an emergent property of the entity concerned, produced by a generative mechanism ([20], p. 14, [21], p. 21). An emergent property, as I use the term, and ignoring one or two complexities, is a property of an entity that would not be held by the parts if they were not organized into such an entity (I take this to be similar to Wimsatt’s view of emergence as non-aggregativity [22]). A generative mechanism is a process in which the parts of an entity interact to produce a causal power. Thus, for example, the parts of a torch can interact in such a way that it will shine a light in a particular direction. The underlying generative mechanism is a process—in this case, an electrical flow through a circuit and thus the filament of the bulb, producing light, which is directed by a curved mirror in a certain direction. The process in turn depends upon the composition and structure of the entity possessing the power—the torch. The parts of the torch could not have this effect in the world unless they were organized into the form of a torch, so this is an emergent causal power of the torch and not a causal power of the parts.

One consequence of this way of thinking is that causality must be disconnected from strong conceptions of empirical regularity (which have tended to be identified with causality in the social sciences). Particular causal powers might produce partial regularities, but there can always be exceptions, cases when the power is present but the usual event does not occur owing to the intervention of other powers. And some causal powers might lie dormant on most occasions, only occasionally being stimulated by a fortuitous combination of other factors.

In the terms suggested by Achim Stephan, this is a *weak* rather than a *strong* conception of emergence: weakly emergent properties are collective properties that can potentially be explained in scientific terms, for example by explaining the mechanisms that generate them, whereas strongly emergent properties are collective properties that cannot be explained in terms of the contributions of the lower level parts and the relations between them [23,24]. Strong emergence is somewhat problematic; not only is its scientific unexplainability likely to be unattractive to readers of this journal, but also it has seemed unlikely to some observers that there actually *are* any strongly emergent properties ([25], p. 18). Weak emergence of the kind advocated here, on the other hand, is sometimes thought to lead to thoroughly reductionist conclusions ([24], p. 79, [26], p. 563, [27], pp. 486–488). However,

as a number of authors have pointed out, to explain a higher level property does not entail explaining it *away*; on the contrary, such explanations demonstrate that the higher level property is consistent with our understanding of other levels ([28,29], p. 181, [30], p. 112). The fact that the power of the torch to shine a light is explainable in terms of the interactions of its parts does not mean that those parts would possess such a power if they were not organized into the form of a torch. This is still a causal power of the torch and not a causal power that would be possessed by the battery, bulb, mirror, casing, etc., organized in any other form.

This is a theory of so-called weak emergence, then, that is nevertheless *strong enough* to resist the kind of reductionism that is incompatible with science, and *strong enough*, I will argue, to give us a meaningful and plausible account of top-down causation.

4. LEVELS AND TOP-DOWN CAUSATION

If one prerequisite of making sense of top-down causation is a theory of causation, another is an interpretation of the term *top-down*. One such interpretation is that there are ontological *levels* of reality, and that top-down causation occurs when an entity or property or event at a higher level has a causal impact on one at a lower level [31]. Accounts of such level structures are generally based on the ontological dependence of the higher on the lower levels. Perhaps the least problematic version of ontological dependence is that entities of a higher level are composed of entities of a lower level: the parts of a whole are at a lower level than the whole. We may argue, for example, that human individuals are at a higher ontological level than molecules, as they are composed of molecules. Hence, an experiment in chemistry would be a case of this kind of top-down causation, if the human scientists create the conditions in which a particular molecular reaction occurs.

Under the account of causality given above, this version of top-down causation is entirely unproblematic. The power to bring certain chemicals together in the appropriate conditions is an emergent causal power of the scientist, and it is a power that interacts with the powers of various molecules, containers, etc., to bring about the chemical reaction. In such cases, the scientist has a downward causal impact on the reaction. There is no reason why we should follow Craver, Bechtel and Hodgson in thinking of this as merely an effect that is inferior in some way to any other kind of causation. The power of the scientist to set up an experiment is an emergent human power: the molecules that make her up could not do this without being organized into the form of a human being. In just the same way, the power of the molecules involved in the reaction is an emergent power of the molecules concerned: the atoms composing the molecules would not be able to react in this way if they were not the parts of such molecules (the classic example of this is the various powers of water: if, for example, we spray water on a fire, we will get rather a different reaction than if we spray hydrogen and oxygen on it). In other words, the

ontological structure that gives the scientist the power to co-cause the experimental outcome is, at a certain level of abstraction, exactly the same as the ontological structure that gives the molecules concerned the power to co-cause that outcome.

A second, and at first sight more demanding, interpretation of the *top-down* in *top-down causation* is that *top-down* causation occurs only when an entity has a causal impact on its *own* parts. We need to alter our example only slightly to show that much the same considerations apply in this case as in the first interpretation. If our scientist was to conduct her experiment on herself, perhaps by placing a drop of some chemical on her skin and observing/measuring the chemical reaction that ensued, we would have just such a case. The scientist's power to co-cause the chemical reaction remains an emergent power of the scientist as a whole, while the molecules that take part in the reaction, including those that form part of the scientist, also co-cause the reaction. Causation is exerted both by the scientist and by the molecules, and the scientist has a downward causal effect on some of the parts of herself. Events, in other words, may be co-determined by the causal powers of both higher and lower level entities, even when the lower level entities concerned are parts of the higher level entities concerned, and even where the events that result are changes in those same lower level entities and thus also in the higher level entity itself.

This is not to deny that there is a sense in which we may think of wholes as exerting their causal effects *through* their parts, and this brings us to a third interpretation of top-down causation. To return to my earlier example, when a person presses the switch that turns on the torch, she may do so using her finger. Those of a reductionist streak may be tempted to argue that it is the finger that presses the button and not the whole person as such, perhaps thereby satisfying Craver & Bechtel's requirement for purely intra-level causation. And of course, it is true that it is the finger that presses the button, in a certain material sense. But the issue here is causality, and it is quite clear that a human finger that was not part of a functionally whole human would generally be incapable of pressing a switch. (As I have implicitly done throughout, I ignore here those counter-cases in which the same result is achieved as a result of the intervention of some *other* higher level force, such as a robot using a severed finger. In such cases, the result would still depend on a higher level force, although a different one.) A severed finger would not have the causal power required; and so when the person's actual finger does press the switch, we can conclude that this pressing was produced by a causal power of the whole human, acting *through* the finger. The pressing by the finger is produced by the causal power of the whole human. Although the finger's action is itself produced by a series of other lower level events, such as the firing of various neurons and the transmission of signals along various nerves, these too are produced by the causal power of the whole human. If there was not a functionally whole human, there would be no set of connections between the various parts implicated in the process, and no possibility of the relevant neurons firing to produce the movement

of the finger. There is, in other words, a sense in which *whenever* a higher level entity has a causal impact, it achieves this impact through the effects it has on its own parts. In a sense, we might say, *all* causation implies top-down causation, except causation by entities at the lowest (*fundamental* or *basic*) possible level, if there are such entities.

Even neglecting this point, however, top-down causation, on this account, works in just the same way as any other type of causation. In all cases, it is the presence of the level of organization provided by the entity that possesses the power as an emergent property which is required to generate the causal power. In all cases, the operation of the higher level power depends on the causal powers of the parts, but it is only when they are organized in the form of the requisite type of whole that they can have these powers.

5. THE CAUSAL POWER OF SOCIAL ENTITIES

In one sense, the application of this argument to the social case is thoroughly straightforward: the same abstract principles should (and, I argue, do) apply to social entities as to any others as far as causal powers are concerned. Several considerations, however, stand in the way of such a straightforward extension to the argument. The first, to be addressed in this section, is that social scientists are unaccustomed to thinking of the social world as being populated by entities with material (e.g. human) parts. Most conceptions of social structure, for example, do not see social structures as things composed of people, but as some kind of social property that arises from relations or interactions between people without taking anything like an entity form. Another alternative, advocated in Giddens's structuration theory, is that social structure exists only as a property of human individuals, in the form of mental properties such as rules [3,32]. It is thus somewhat novel to argue that social structures are essentially causal powers of social *entities* (although not unique—for example, [33]). This section will seek to justify this view, using two kinds of examples: first, organizations, and second, the entities I have called *norm circles*.

Generally, I argue, the causal effects that social scientists attribute to *social structure* are produced by the causal powers of social entities. Typically, these entities are groups of people (though some may also include other types of parts; a university, for example, may include a campus and buildings as well as staff and students). There may be different types of social entities, all composed primarily of people, but with those people related by different types of structural relation, each type of structural relation corresponding to a different type of entity (just as atoms of carbon, oxygen and hydrogen may be combined in many different ways to produce different types of molecular entity). Any given causal power of a social entity will depend on a generative mechanism: a process of interaction between the people who are its members, a process that in turn depends upon the type of structural relations that exist between them in an entity of this type.

Perhaps the most obvious type of such a social entity is organizations. An organization consists of a group of people, in a structured set of relationships to each other. Those relationships are often characterized as the roles of the individuals concerned. As a result of its members being committed to interact in the ways specified in their roles, the organization has the capacity to have a causal impact on the world that its members would not have if they were not parts of the organization concerned. Thus, for example, an orchestra has the causal power to produce harmonious music because its parts—the players (and arguably their instruments)—inter-relate in the ways specified in their roles (violin player, pianist, conductor, etc.) and indeed further specified in those more specific sets of relational instructions that we call musical scores. If they were not organized into an orchestra, the players and instruments would not have the causal power to produce harmonious music.

This paper, however, will focus on a different kind of social entity: what I have called *norm circles*. Norm circles, I argue, are the type of social entity that is causally responsible for what sociologists refer to as normative social institutions or social practices. A *social practice* is a recognizable pattern of behaviour or action that occurs repeatedly in a social space, ranging from something as simple as standardized forms of greeting to something as complex as the practices that surround (and indeed create) what we think of as *property*. Such practices are of profound importance in our social world. Not only simple day-to-day practices like queuing, but also phenomena like money, religion, language, culture and indeed organizations depend on normatively standardized practices.

Sociologists often invoke the concept of a *normative social institution* as a kind of explanation of the regularity of such practices. What is generally accepted is that such institutions are driven by normative pressures (*socialization*, for example) that encourage people to conform to the practice and may penalize those who do not. What is not generally accepted, and indeed has been the focus of over a century of debate, is just what form social institutions take that gives them the capacity to exercise such an influence.

The hypothesis examined in my work is that social practices are produced by the causal power of social entities that I call norm circles ([4], ch. 6). A norm circle is the group of people that are committed to endorsing and enforcing a specific norm, a specific standard of observable behaviour. The relation between them that gives them the collective capability to influence behaviour—a greater influence than an unconnected group of individuals would have—is the sense of shared commitment they have to supporting the norm. The members of a norm circle may be unaware of the full extent of the group, and they may not even think of it as a group, but they are generally aware when they act in support of a norm that they are not simply expressing a purely idiosyncratic personal attachment to a particular standard of behaviour. Rather, they are aware that when they do so, they are endorsing a standard that others also endorse, and often do so with the expectation that others would

support and approve of their action. The individual, in other words, has a sense, however vague and minimal, that she is acting on behalf of something wider than herself when she acts in support of a norm, and that sense increases the likelihood that she will act in its support, by comparison with the isolated individual with a purely personal attachment to the standard of behaviour concerned.

This sense, in turn, is a product of the same process that tends to encourage conformity to the norm—the generative mechanism that underpins the power of a norm circle to increase such conformity. The heart of this process is repeated exposure of individuals to acts of endorsement and enforcement of the norm concerned. If, for example, I repeatedly see people criticizing those who try to jump queues, I will start to understand the norm of queuing, and to believe that I face an environment in which I will be sanctioned negatively if I fail to observe it. I will, in other words, develop beliefs about my normative environment which will tend to lead me to conform to the norm of queuing in the future, as a result of the actions of members of the norm circle for queuing. Here, then, social structure—the norm circles that produce the normative environment—is exerting a top-down influence on individual action, representing the first half of Archer’s morphogenetic cycle. Similar effects can be produced without us forming conscious beliefs; most of us, for example, understand and implement the norms prevalent in our social space regarding how close one should stand to someone when talking to them, even though these norms are rarely stated explicitly and are mostly endorsed and enforced rather subtly by non-verbal signals.

Norm circles, then, operate through individuals. On the one hand, it is the actions of individual members of the norm circle acting in support of a norm that signal the normative environment to other individuals (individual actions, then, reproduce and/or transform these social structures, as in the second half of Archer’s morphogenetic cycle). On the other hand, those pressures do not lead directly and mechanically to norm conformity but rather influence the stored beliefs and dispositions of the affected individuals, which then in turn influence their subsequent behaviour. Nevertheless, I argue that the resulting increase in the tendency of those affected individuals to conform to the norm is causally influenced by the norm circle, and not just the individuals. Earlier we saw that although a person who switches on a torch must do so using her finger, the finger could not do so except as a result of the causal influence of the whole individual, and so the act of switching it on is produced by a causal power of the whole individual and not just of the torch. In a similar way, the norm circle can only influence us through its individual members, but those individual members would not influence us in that way, or at least not as strongly and as often as they do, if they were not part of a wider norm circle, and so their act of influencing us is produced by a causal power of the norm circle and not just of the individual.

There is much more to be said about the theory of norm circles. For example, norm circles in contemporary societies are diversely intersectional—different

norms are supported by different but profusely overlapping groups of people. This in turn makes normative change more likely than in more homogeneous societies, as individuals are open to the influence of competing norm circles and may move between them. One implication is that the theory of norm circles is not merely concerned with the reproduction of a stable normative environment. It seeks to explain how normative influences contribute to the production of social actions, which may, in the morphogenetic cycle described by Archer, contribute to the reproduction of the normative environment, but there are many reasons why norms may be transformed rather than reproduced in some social situations. And while there is a link here to those who think of norms in strongly moral terms, it is not necessarily the case that those who conform to norms, or even endorse them, adopt them in such a deeply committed way—conformity and even endorsement may be motivated by much more instrumental reasons. Those readers interested in these and other complexities are encouraged to look at the discussion of them elsewhere ([4], ch. 6).

6. TOP-DOWN CAUSATION BY SOCIAL ENTITIES

A further complexity is implicit in the theory of causation advocated by this paper: any given social event may be influenced not only by one norm circle, but also by multiple, perhaps even conflicting, norm circles, as well as by other social entities such as organizations. And where these structural pressures on an individual are in conflict with each other, the individual retains the capacity to make choices between them ([4], ch. 8). The causal powers of the individual concerned, in other words, may also contribute to causing the subsequent event. This is, therefore, a framework that allows events to be co-determined by both social structure and individual agency.

With this qualification in place, we may summarize the ways in which organizations and norm circles exert the three types of top-down causal power discussed earlier. First, the capacity to have causal effects on entities at a lower level that are not parts of the whole exercising the power. This is unproblematic. Organizations commonly have causal impacts on non-members and the events in which they participate. For example, governments compel us to pay taxes, commercial enterprises sell us things and political parties persuade us to vote for them. And as we have already seen, norm circles influence non-members in ways that lead them to adopt the practices the norm circles support.

Second, the capacity to influence their own members and events in which they participate, in much the same way as they influence non-members. Again, there are straightforward parallels to the ordinary material cases. Civil servants, for example, have to pay taxes too, and members of parties vote, and in both cases the organizations concerned may influence their own members in this respect. And norm circles continue to influence their members as well as non-members to conform with the norm the norm circle supports.

Third, as we have already seen for the case of norm circles, social entities, like ordinary material ones, act through their parts, and when they act they do so by downwardly causing the behaviour of the part that they act through. Thus, for example, when a retail company sells an item to a customer in a shop, it will generally do so through a sales assistant or checkout operator who is employed by the company. In one sense, it is the employee who sells the item to the customer, but they do so on behalf of the company, and when they do so, this behaviour is at least partly co-determined by the causal influence of the company on its employee. If the employee was not employed by the company, she would not believe that she could sell the item on its behalf, nor would she feel under an obligation to do so. The employee's act is determined directly by her current set of beliefs and dispositions, but those include a number of such beliefs that have been produced by her membership in the company, such as the belief that she is entitled to act on behalf of the company in some respects, the belief that she is expected to sell to customers under certain circumstances and the belief that she faces a normative environment in which failure to do so may be penalized. This, of course, introduces a further complexity: the roles that individuals are expected to fulfil within organizations are essentially bundles of norms, and organizations operate partly by constructing or at least shaping norm circles to support those norms. Such interdependencies between different types of social entity raise ontological issues that are not generally found in the study of ordinary material entities, issues that require some caution in the use of the concept of top-down causation.

7. INTERLEAVING ONTOLOGICAL STRUCTURES

Put briefly, the problem is this: above the level of the human individual, social structures are not nested neatly in level structures like those of ordinary material entities, such as the person composed of cells, each of which is composed of molecules, each of which is composed of atoms, and so on. To understand how this can be, we need to make explicit an important feature of social entities: that there is a fundamental difference between the kinds of relations that structure social entities, and those that structure ordinary material entities. While ordinary material entities are structured by strongly spatially constrained relations, social entities are not. Organizations and norm circles depend not only on spatial but also on *intentional* relations between their members. In other words, they depend upon relations and commitments to interact in certain ways that are represented in the mental properties of their members: in their beliefs and dispositions.

This has a number of consequences. One is that a certain temporal disarticulation between causes and effects becomes possible. The influence of the whole on the part does not depend on synchronic spatial relations with the other parts of the whole, but instead the whole may have an impact on a member's beliefs at one time, which affects her behaviour some time later.

The consequence that is most significant for the argument of the present paper, however, is that the members of a social entity are not constrained to be members of only *one* such entity. The spatial constraints on parthood of ordinary material entities mean that they can usually only be a part of one entity at the next highest level of organization at a time. An atom, for example, can generally only be part of one molecule at a time. But a human individual can simultaneously be a member of an unlimited number of organizations, norm circles and other social entities, because these are not spatially structured entities, and an individual is capable of simultaneously possessing the many different mental properties required to be a member of each of them ([4], ch. 9).

This in turn affects the ways in which different social entities can interact. As we have seen, for example, organizations are structured by roles, which are effectively bundles of norms, and these norms like any other are causally produced by norm circles. But norm circles are not parts of the organizations that depend on them. The relationship is much more subtle than this. Consider the case of the conductor of our orchestra. There are certain norms about what conductors should do that are widespread in the classical music community, and thus supported by norm circles that are far wider than the membership of our particular orchestra. But our particular orchestra may also have developed norms about other things the conductor should do—organizing concert dates, perhaps—that would not generally be recognized as part of every conductor's role. The organization, we could say, is *selective* about which norms are to be bundled into the role. It may depend on wider norm circles for the enforcement of some of them—not only the musicians in the orchestra but also audiences, for example, may react badly if the conductor fails to keep time adequately. But it may develop internal norm circles, perhaps consisting just of the director of the orchestra or the board, to enforce others such as the responsibility to book concert dates. The orchestra as an organization exerts influence on the conductor at least partly by shaping her normative environment, including the norm circles to which the members of the organization belong.

If this is a challenging case, consider a simpler example: an organization can influence another organization by influencing its own members, when these are also members of the second organization. In the recent election of a new leader of the British Labour Party, for example, some trade unions encouraged their members who were also members of the party to vote for a particular candidate. Or another: organizations can influence normative social institutions by endorsing certain practices. In the early development of industrial capitalism, for example, manufacturing firms encouraged the development of punctuality among former agricultural workers who had previously been able to work to their own schedules, thus contributing to the widespread endorsement of punctuality norms in modern societies.

In each of these cases, one social entity exerts a causal influence on another by influencing its own members, who are also (or who also become) members of the

second entity. These are not nested, but rather *interleaving* ontological structures, and this brings us to a limit of the concept of top-down causation. The concept rests on the assumption that there are distinct ontological levels, based essentially on the principle of composition. But the intentional nature of the relations that produces social entities from human parts permits a kind of branching of the compositional structure that is not possible for most ordinary physical structures of composition. Thus, for example, although organizations depend ontologically on norm circles, norm circles are not parts of organizations. Instead, norm circles and organizations are interleaving structures composed of overlapping sets of parts—the orchestra discussed above, for example, includes people who are members of the norm circles for what a conductor should do but not all members of the orchestra need belong to all of these norm circles, and not all members of these norm circles will usually be part of the orchestra. The question of top-down causation therefore does not apply to the direct relation between them. Instead, we have a more interesting process of inter-influencing between mutually overlapping entities ([34], pp. 106–107).

Nevertheless, this inter-influencing happens through a process in which each structure exerts a downward causal influence on a lower level which is also the lower level of the other structures affected. The intra-level influence thus appears to operate through downward influences on the shared lower level. Top-down causation continues to be important, but in a somewhat different way than we generally find in the ‘natural’ sciences. This is not to say, however, that there could be no other cases. Such phenomena might be possible in any kind of system whose structure depends on stored states and which is capable of storing multiple such states. Computer systems, for example, might also support such ontological structures, and it is perhaps conceivable that such structures might also be one of the sources of complexity in neurological/mental structures.

8. EMPIRICAL TESTING

How might we test claims for top-down causation in the social domain? In one sense, the problem seems trivial. We are all constantly exposed to cases in which we are causally influenced by social entities—for example, the organizations we work for, governments, banks and voluntary organizations in which we participate. If these are higher level entities and they influence our behaviour, then we have evidence of top-down causation. Furthermore, although they have not been discussed here, there may be many other types of social entity composed of people (and perhaps of other things, including organizations themselves), which may exert downward causal influence on us. Perhaps the most striking case would be market systems, which seem to have a profound causal influence on phenomena like employment and unemployment, economic growth, inflation and technological innovation.

In a different sense, however, the problem seems to be more challenging. Can we really provide empirical proof of ontological claims that we make about the

empirical phenomena? By contrast with the kinds of argument that have been advanced in this paper, for example, methodological individualists deny that organizations have causal power on the grounds that what seems to be their causal power is really just the causal power of the individual members who act in any given case. While we can advance abstract arguments against such claims, it is nevertheless the case that the empirical facts are apparently compatible with *both* of these positions. Any empirical fact that the emergentist interprets as evidence of causal power of a social entity can be re-interpreted by the individualist as evidence of the causal power of the particular individuals involved in the events concerned. Thus, for example, when the emergentist say that a bank has the causal power to make a loan to a person, the individualist can respond that it is really the individuals that make up the bank that have this causal power.

When ontological theories clash in this way, the best tests we have are not directly empirical. Instead, they are coherence tests—(i) how coherent is the claim being made with other claims that are generally accepted, and (ii) how coherent is it with the rest of the claims being made in the argument. For example, individualists often deny causal power to social structure on the basis of the reductionist argument that we can always explain the effects of the structure in terms of causal contributions of the individuals concerned and the relations between them, but they consistently fail to apply this argument recursively to human individuals. If they were to do so, they would have to argue that we can always explain the effects of an individual in terms of the causal contributions of their cells and the relations between them, the effects of cells in terms of the causal contributions of the molecules that are their parts and the relations between them, and so on. They could not be individualists but would have to successively become cell-ists (!), molecule-ists, atom-ists, and so on. Causality, on this account, would seem to drain away—either to the lowest physical level, if there is such a level, or entirely, if there is not [35,36]. In neither case are humans causally effective, hence we can assert that the emergentist argument advanced here is more coherent than the individualist one. It must also, of course, be consistent with empirical evidence, but when multiple theories are equally consistent with the evidence we must turn to other criteria of theory choice. I would suggest that such other criteria are both in principle necessary and in practice widely used.

9. CONCLUSION

This paper has advanced an account of causality and emergence within which top-down causation is justified on just the same basis as any other kind of causation. All causation, at least at the ontological levels to which we currently have cognitive access, is inter-entity causation, and all such causation depends on the levels of organization possessed by the entities concerned, irrespective of their compositional level. This paper has applied this view to social structures and their relation to human individuals, stressing

that social structural powers are powers of social entities, composed primarily of people. Both people and social entities (and indeed ordinary material entities) possess emergent causal powers, which interact in the causation of social events. Nevertheless, social entities differ from ordinary material entities in at least one interesting and significant way: they are structured not only by spatial but also by intentional relations. One consequence is that above the level of the human individual, social entities do not form nested compositional structures, as ordinary material objects do, but instead are interleaved. Top-down causation ceases to be a suitable model in accounting for the relations between such entities, although they may influence each other by the top-down effects that each produces in their shared parts.

REFERENCES

- 1 Archer, M. S. 1995 *Realist social theory: the morphogenetic approach*. Cambridge, UK: Cambridge University Press.
- 2 Bourdieu, P. 1990 *The logic of practice*. Cambridge, MA: Polity.
- 3 Giddens, A. 1984 *The constitution of society*. Cambridge, MA: Polity.
- 4 Elder-Vass, D. 2010 *The causal power of social structures*. Cambridge, UK: Cambridge University Press.
- 5 Humphreys, P. 1997 How properties emerge. *Phil. Sci.* **64**, 1–17. (doi:10.1086/392533)
- 6 Kim, J. 1992 ‘Downward causation’ in emergentism and nonreductive physicalism. In *Emergence or reduction?* (eds A. Beckermann, H. Flohr & J. Kim), pp. 119–138. Berlin, Germany: de Gruyter.
- 7 Archer, M. S. 1996 *Culture and agency*, 2nd edn. Cambridge, UK: Cambridge University Press.
- 8 Archer, M. S. 1982 Morphogenesis versus structuration. *Br. J. Sociol.* **33**, 455–483. (doi:10.2307/589357)
- 9 Archer, M. S. 1990 Archer replies to outhwaite. In *Anthony Giddens: consensus and controversy* (eds J. Clark, C. Modgil & S. Modgil), pp. 86–88. London, UK: Falmer.
- 10 Archer, M. S. 1979 *Social origins of educational systems*. London, UK: Sage.
- 11 Domingues, J. M. 2000 Social integration, system integration and collective subjectivity. *Sociology* **34**, 225–241. (doi:10.1177/S0038038500000158)
- 12 Elder-Vass, D. 2007 For emergence: refining Archer’s account of social structure. *J. Theory Soc. Behav.* **37**, 25–44. (doi:10.1111/j.1468-5914.2007.00325.x)
- 13 King, A. 1999 Against structure: a critique of morphogenetic social theory. *Sociol. Rev.* **47**, 199–227. (doi:10.1111/1467-954X.00170)
- 14 Sawyer, R. K. 2005 *Social emergence*. Cambridge, UK: Cambridge University Press.
- 15 Fodor, J. A. 1974 Special sciences (or: the disunity of science as a working hypothesis). *Synthese* **28**, 97–115. (doi:10.1007/BF00485230)
- 16 Sawyer, R. K. 2002 Nonreductive individualism part I: supervenience and wild disjunction. *Phil. Soc. Sci.* **32**, 537–559. (doi:10.1177/004839302237836)
- 17 Hodgson, G. M. 2007 Institutions and individuals: interaction and evolution. *Organ. Stud.* **28**, 95–116. (doi:10.1177/0170840607067832)
- 18 Hodgson, G. M. 2011. Downward causation: some second thoughts. See <http://www.geoffrey-hodgson.info/downward-causation.htm> (accessed 7 June 2011).
- 19 Craver, C. & Bechtel, W. 2007 Top-down causation without top-down causes. *Biol. Phil.* **22**, 547–563. (doi:10.1007/s10539-006-9028-8)
- 20 Bhaskar, R. 1975 *A realist theory of science*, 1st edn. Leeds: Leeds Books.
- 21 Bunge, M. 1999 *The sociology–philosophy connection*. New Brunswick, NJ: Transaction.
- 22 Wimsatt, W. C. 2000 Emergence as non-aggregativity and the biases of reductionisms. *Found. Sci.* **5**, 269–297. (doi:10.1023/A:1011342202830)
- 23 Bedau, M. A. 1997 Weak emergence. *Phil. Perspect.* **11**, 375–399.
- 24 Stephan, A. 2002 Emergentism, irreducibility, and downward causation. *Grazer Phil. Stud.* **65**, 77–94. (doi:10.1177/S0038038500000158)
- 25 Kim, J. 1999 Making sense of emergence. *Phil. Stud.* **95**, 3–36. (doi:10.1023/A:1004563122154)
- 26 Sawyer, R. K. 2001 Emergence in sociology. *Am. J. Sociol.* **107**, 551–585. (doi:10.1086/338780)
- 27 Stephan, A. 2006 The dual role of ‘emergence’ in the philosophy of mind and in cognitive science. *Synthese* **151**, 485–498. (doi:10.1007/s11229-006-9019-y)
- 28 Kitcher, P. 1998 1953 and All that: a tale of two sciences. In *Philosophy of science: the central issues* (eds M. Curd & J. A. Cover), pp. 971–1003. New York, NY: W W Norton.
- 29 Meyer, T. C. 2000 Physicalism and downward causation in psychology and the special sciences. *Inquiry* **43**, 181–202. (doi:10.1080/002017400407744)
- 30 Gell-Mann, M. 1995 *The quark and the jaguar*. London, UK: Abacus.
- 31 Ellis, G. 2006 On the nature of emergent reality. In *The re-emergence of emergence* (eds P. Clayton & P. S. Davies), pp. 79–109. Oxford, UK: Oxford University Press.
- 32 Stones, R. 2005 *Structuration theory*. Basingstoke: Palgrave Macmillan.
- 33 Sheehy, P. 2006 *The reality of social groups*. Aldershot: Ashgate.
- 34 Wan, P. Y.-z. 2011 *Reframing the social*. Aldershot: Ashgate.
- 35 Block, N. 2003 Do causal powers drain away? *Phil. Phenomenol. Res.* **LXVII**, 133–150. (doi:10.1111/j.1933-1592.2003.tb00029.x)
- 36 Kim, J. 2003 Blocking causal drainage and other maintenance chores with mental causation. *Phil. Phenomenol. Res.* **LXVII**, 151–176. (doi:10.1111/j.1933-1592.2003.tb00030.x)