# EMERGENCE AND THE REALIST ACCOUNT OF CAUSE

BY

DAVE ELDER-VASS

[This is an electronic version of a paper published in the *Journal of Critical Realism*. Please cite as:

Elder-Vass, D. (2005) 'Emergence and the Realist Account of Cause'. *Journal of Critical Realism* 4:2, 315-38.]

*Abstract*. The concept of emergence is routinely invoked in critical realist theory, but rarely examined. Yet emergence is fundamental to the realist account of cause. This paper aims to improve critical realism's understanding of emergence by discussing, first, what emergence is and how it works; second, the need for a compositional account of emergence; and third, the implications of emergence for causation. It goes on to argue that the theory of emergence leads to the recognition of certain hitherto neglected similarities between real causal powers and actual causation.

*Key words*: cause, critical realism, emergence, reductionism, causal powers, emergent properties

## 1. *Introduction*

The concept of emergence is routinely invoked in critical realist theory, and indeed it is foundational to critical realism's account of the world.[1] Bhaskar, for example, has written: 'It is only if social phenomena are genuinely *emergent* that realist *explanations* in the human sciences are justified […] Emancipation depends upon explanation depends upon emergence.'[2] Yet few critical realists have examined the nature of emergence itself, while those few who have done so have been far from consistent in their approach.

The objective of this paper is to take some steps towards a more coherent approach to emergence as a general phenomenon, and thus to strengthen the usage of the concept in the critical realist tradition. In doing so, it will engage critically but constructively with the work of some leading critical realists. It will also draw on the work of other

[1] I would like to thank Jason Edwards and two anonymous referees for their valuable comments on earlier drafts of this paper.

[2] Roy Bhaskar, *Scientific Realism and Human Emancipation*, London: Verso, 1986, pp. 103-4.

thinkers who have discussed emergence from a variety of disciplinary and philosophical standpoints, including recent philosophy of the mind, systems thinking, and complexity theory. Although all these traditions have useful and interesting things to say about emergence, I suggest that none of them has yet developed an adequate understanding of emergence and how it works. This leads to a degree of confusion about causal explanations, and a further objective of this paper is to help dispel some of this confusion by developing the theory of emergence in directions that show its implications for causation.

The paper begins by defining emergence and the terms required for its analysis, then identifying the various elements that interact to produce emergent properties. One of these, I argue, is the *compositional* basis of emergence. This has been disputed within critical realism, notably by Andrew Collier, and the paper will defend the compositional approach against his argument. It then moves on to examine the implications of emergence for our understanding of causation, and in particular the relation of emergence to real causal powers and actual causation. It concludes that emergence is indeed foundational to a realist understanding of causation, but that we might have to modify some parts of the critical realist ontology if the arguments of this paper are accepted. In particular, the distinction between real causal powers and actual causation might need to be complemented by a recognition of some important underlying similarities between the two.

## 2. *Defining Emergence*

Put at its simplest, emergence is operating when a whole has properties or powers that are not possessed by its parts.

However, a number of clarifications and expansions of this claim are required. First, we must distinguish between synchronic and diachronic conceptions of emergence. Lay usages of 'emergence' generally refer to diachronic emergence, which denotes the first appearance or initial development of some new phenomenon. While this is of some interest, and is certainly related to the synchronic conception, I shall generally *not* use the word 'emergence' in this sense. Instead I focus on the synchronic sense, which is concerned with the relationship between the properties and powers of a whole and its parts at any single instant in time.

Secondly, we must clarify what is meant by 'wholes' and 'parts'. Both 'wholes' and 'parts' in the basic definition above are 'entities', and the terms 'whole' and 'part' therefore describe roles played by particular entities in particular cases (an entity that is a whole in one context can be a part in another). Entities are to be identified with objects or things, although this does not mean that they are necessarily *material* things – examples include atoms, molecules, cells, trees, human individuals, theories, business corporations, and armies.[3] Any entity (except perhaps the most fundamental material particles, if there are such things) consists of a set of parts that is in some

[3] As one referee has pointed out, social entities raise further difficulties, such as the problem of unambiguously identifying their boundaries. I discuss this and other aspects of the emergence of social entities in Dave Elder-Vass, 'The emergence of social structure and the question of naturalism', paper presented at the BSA Annual Conference, York, 2005.

way structured, such that the relations between the parts are more than merely aggregative. There may therefore be collections of parts that do not form entities, such as relatively arbitrary constructs like 'all the rice in China'.[4] I follow Laszlo in calling such unstructured collections of parts 'heaps'.[5] Now, strictly speaking, all collections of parts, however arbitrary, have relations between them – all material entities, for example, exercise a gravitational force on each other – so to be strictly accurate, we must define heaps as collections of parts that lack *significant* structure; I shall return to the meaning of 'significant' in this context below. Furthermore, an entity must have the quality of persistence, in the sense that it must sustain its existence over a significant period of time. To summarise, an entity may be defined as *a persistent whole formed from a set of parts, the whole being significantly structured by the relations between these parts*.

Thirdly, we must clarify what is meant by a 'property or power'. A property is some intrinsic aspect of an entity that can have a causal impact on the world (I use 'intrinsic' in order to exclude purely formal relations with other entities, such as 'larger than *x*', from the definition of 'properties.')[6] 'Properties' and 'powers' may therefore be regarded as synonyms.

Emergence occurs when an entity possesses one or more 'emergent properties'. An 'emergent property' is one that is not possessed by any of the parts of the entity individually, nor when they are aggregated, without a structuring set of relations between them. Perhaps the commonest illustration of emergence in the literature is the example of water, which has been used to illustrate this point as far back as John Stuart Mill.[7] The properties of water are clearly very different from those of its components oxygen and hydrogen when these are not combined with each other in the specific form that constitutes water. One cannot, for example, 'put out a fire with oxygen and hydrogen'.[8] Hence water has emergent properties. Another illustration is provided by colour: 'The collective structure of bulk matter reflects light at certain preferred wavelengths; those determine the color. Color is an emergent phenomenon; it only makes sense for bulk matter'.[9] Molecules – the parts of bulk matter – simply do not have the property of colour; hence this property emerges from their structured combination into larger wholes. These particular examples rest on the ability of the entities concerned to interact with other external entities (fire and light), or in other words, these powers can only be exercised when and because there are other entities with corresponding liabilities, but this is not the case for all powers.[10] Stars, for example, have a power to emit radiation over a variety of wavelengths, including

[4] Andrew Collier, *Scientific Realism and Socialist Thought*, Hemel Hempstead: Harvester Wheatsheaf, 1989, p. 193.

[5] Ervin Laszlo, *The Systems View of the World*, Oxford: Blackwell, 1972, p. 28.

[6] See Sayer on 'formal relations': Andrew Sayer, *Method in Social Science*, London: Routledge, 1992, p. 88.

[7] John Stuart Mill, *A System of Logic*, 8th ed., London: Longmans, 1900, p. 243.

[8] Kevin Mihata, 'The persistence of "emergence"', in *Chaos, Complexity, and Sociology*, eds Raymond A. Eve, Sara Horsfall, and Mary E. Lee, Thousand Oaks, Calif.: Sage, 1997, pp. 30-8, p. 31. Sayer uses the same example in Sayer, *Method in Social Science*, p. 119.

[9] Jack Cohen and Ian Stewart, *The Collapse of Chaos*, New York , London: Penguin, 1995, p. 232.

[10] Liabilities are described in Rom Harré and Edward H. Madden, *Causal Powers : A Theory of Natural Necessity*, Oxford: Blackwell, 1975, p. 89.

visible light, and this power does not significantly depend upon the liabilities of other entities.

Emergent properties may be contrasted with 'resultant properties' – these are properties of wholes that *are* possessed by its parts in isolation, or in an unstructured aggregation. The classic example of a resultant property is mass – the mass of a molecule, for example, is the sum of the mass of its constituent atoms.[11] A property that is resultant at one level may be emergent at a lower level. And it is entirely possible – indeed it is normal – for entities to possess both emergent and resultant properties.

It is less obvious whether it is possible for an entity to possess *only* resultant properties. Ultimately this is a definitional question, since the answer rests upon whether we choose to define 'entity' to include a persistent collection of parts that does not have emergent properties as a whole – say, for example, a fence with a post box attached to it. This brings us back to the difference between a significant and an insignificant structural relation which came up in the earlier discussion of entities and heaps. In order to distinguish between the two, we had to define entities as possessing *significant* relations between the parts. If we make it the criterion of significance for a relation that it leads to the whole possessing a causal power not possessed by the parts without it, then the 'fence and post box' is a heap and not an entity. In this paper I accept this strategy, and so it now follows by definition that all entities have emergent properties or powers.

This does not mean, however, that entities can be dispensed with in favour of an account in terms of properties or powers alone.[12] Properties are not free-floating phenomena; they always occur as the effects of a particular configuration of lower-level parts.[13] Colour, for example, can not exist except as an effect of a coloured thing. Now, admittedly, if a whole is an organised set of parts, and each part is itself an organised set of parts, then unless there is some lowest-level thing which is not 'just' organisation, then any entity can ultimately be decomposed into a set of relations between relations. Nevertheless, an entity remains a *real* and persistent set of relations between relations, with causal powers that are irreducible to any of its lower-level decompositions (see the discussion of reduction below). This real set of relations is different from the properties that depend upon it, and any attempt to eliminate the entities from this picture obscures the nature of emergence.

It is often useful in discussing emergence, as for example in the previous paragraph, to use the concept of higher and lower levels, and I shall follow convention here in treating wholes as higher-level entities than their parts or components, and vice-versa. The term 'higher-level entity', however, is more general than the term 'whole', since

[11] For a particularly clear account of what it means for a property to be resultant, or aggregative, see W. C. Wimsatt, 'Emergence as non-aggregativity and the biases of reductionisms', *Foundations of Science* 5, 2000, pp. 269-97.

[12] A tendency for philosophers of mind to talk in terms of properties alone has been one of the reasons for their failure to develop an adequate understanding of emergence.

[13] Bhaskar himself clearly takes this view, e.g. 'Most things are complex objects, in virtue of which they possess an ensemble of tendencies, liabilities and powers' (Roy Bhaskar, *A Realist Theory of Science*, 2nd ed., Hassocks: Harvester Press, 1978, p. 51).

as well as being at a higher level than its own parts, any given whole is at a higher level than all other entities of a similar type to its parts or their parts, and so on recursively. Thus, a water molecule is a higher-level entity than an atom of any kind, and also a higher-level entity than an electron, a proton, a quark, and so on. Similarly, any given part will be considered to be at a lower level than all entities of a similar type to any whole(s) it belongs to, and so on. (In places Roy Bhaskar seems to use higher and lower level in the opposite sense to that used here,[14] but otherwise the usage I adopt here seems to be universal.)[15]

## 3. *Structure*

This section discusses the question: 'How is emergence possible?' If we accept that emergent wholes have properties or powers that their parts do not have in isolation, then where do those properties come from?

Following a great many other theorists of emergence, I argue that they are a consequence of the organisation of the parts, of the maintenance of a stable set of relations between the parts that constitute them into a particular kind of whole. As Buckley, for example, writes:

> When we say that 'the whole is more than the sum of its parts', we are pointing to the fact of *organization* as the 'more than' which imparts to the aggregate of parts characteristics that are not only different from, but often *not found in* the components alone; and the 'sum of the parts' is taken to mean their unorganized aggregation.[16]

A variety of other writers have recognised the central role of organisation in emergence. Here we can go at least as far back as Lloyd Morgan: 'At each ascending step there is a new entity in virtue of some new kind of relation, or set of relations, within it.'[17] More recently Emmeche and his colleagues have written: 'What is "more" about the whole is a specific series of spatial and morphological relationships between the parts.'[18] Many other examples could be cited.[19] The crucial role played by the *organisation* of the parts has been recognised in all the major literatures on emergence.

---

[14] E.g. ibid., Ch. 3.

[15] The 'level' metaphor, however, can sometimes be less than helpful. As Collier points out, a 'tree' metaphor may be more useful, although even this oversimplifies the structure of emergence: Collier, *Scientific Realism*, p. 45. It is also important to be very clear that it is entities that emerge from other entities, and not levels as a whole that emerge from other levels. Levels are a secondary concept, being merely collections of entities that share a common type of parts. Occasionally Collier seems to imply that it is the emergence of levels that is primary, e.g. Collier, ibid., p. 102.

[16] Walter Buckley, *Society: A Complex Adaptive System*, Australia: Gordon and Breach, 1998, p. 36.

[17] C. Lloyd Morgan, *Emergent Evolution*, London: Williams and Norgate, 1923, p. 64.

[18] C. Emmeche, S. Koppe, and F. Stjernfelt, 'Explaining emergence: towards an ontology of levels', *Journal for General Philosophy of Science* vol. 28, no. 1, 1997, p. 106.

[19] Similar points are made, for example, in all of the following: Margaret Archer, 'Morphogenesis versus structuration', *British Journal of Sociology* vol. 33, no. 4, 1982, p. 475; Paul Cilliers, *Complexity and Postmodernism*, Routledge, 1998, p. 43; B. Cunningham, 'The re-emergence of "emergence"', *Philosophy of Science* vol. 68, no. 3, 2001, p. S68; R. K. Sawyer, 'Emergence in sociology', *American Journal of Sociology* vol. 107, no. 3, 2001, p. 560; R. W. Sperry, 'Macro- versus micro-determinism', *Philosophy of Science* vol. 53, no. 2, 1986, p. 266.

My argument, then, is that it is because a higher-level entity is composed of a *particular stable organisation* or configuration of lower-level entities that it may be able to exert causal influence in its own right. This does *not* mean that an emergent property of a higher level entity is *caused* by its parts or by their powers in the usual sense of the term; emergence is a synchronic relationship between a whole and its parts, whereas cause is a diachronic relation in which the powers of a group of entities at one moment *causally* determine the events which follow at the next. The point of emergence is that it is the way that a set of entities is related to each other at a *given* point of time that determines the joint effect they have on the world at that moment. Emergence, then, is a synchronic relation amongst the parts of an entity that gives the entity as a whole the ability to have a particular (diachronic) causal impact. The relation between a whole and its parts is thus a relation of composition, and not of causation.

Although this relation is not causal, we can nevertheless often explain how the relation between the parts produces the overall effect. Indeed, this is the *mechanism* that sustains the emergent property concerned. Returning to the case of water, for example, it is possible to explain why water has the property of being liquid at certain temperatures, why it has the property of being solid (ice) at others, and why its solid form is less dense than its liquid form (unlike most other materials), purely as a result of the properties of hydrogen and oxygen atoms and the sorts of bonds that form between them.[20]

Now, this might seem to be a reductionist explanation of these properties. There are, however, a number of varieties of reducibility or reduction, and it is important to distinguish between them. Most significantly, we must distinguish between 'eliminative' and 'explanatory' reduction.

Eliminative reductions, as described most famously by Ernest Nagel, occur when a higher-level theory is shown to be nothing more than a summation or a special case of a lower-level theory, with the result that the higher-level theory can be dispensed with entirely.[21] Eliminative reductionism argues, in effect, that certain higher-level properties and events can be fully explained by properties at the lower level. This in turn, when combined with the dictum that to be real is to be causally effective, can be taken to imply that the higher-level properties and entities, having no causal impact other than that of their parts, have no real existence, at least as entities in their own right.[22] Thus eliminative reductionists deny both the causal effectiveness of the higher-level entities and their properties, and the need for (or value of) any science conducted in terms of these higher-level entities.

Emergentism is compatible with the eliminative reduction of *some* properties. Indeed this is precisely the significance of 'resultant' properties. The resultant properties of

[20] See, for example, John R. Gribbin and Mary Gribbin, *Almost Everyone's Guide to Science*, London: Phoenix, 1999, pp. 84-7.

[21] Ernest Nagel, 'Issues in the logic of reductive explanations', in *Philosophy of Science: The Central Issues*, eds Martin Curd and J. A. Cover, New York: W W Norton, 1998.

[22] J. Kim, 'The non-reductivist's troubles with mental causation', in *Mental Causation*, eds John Heil and Alfred R. Mele, Oxford: Clarendon, 1993, p. 202.

an entity are not dependent upon the relations between its parts, since they are already possessed by the parts themselves, and the resultant property of the higher-level entity is merely the result of the aggregation of the equivalent properties of its lower-level parts. It is therefore possible to explain the property purely in terms of the parts, neglecting any relations between them that are constitutive of the higher-level entity.

However, it is impossible in principle to provide an eliminatively reduced explanation of a genuinely emergent property, since the mechanism underpinning an emergent property depends upon the relations between the higher-level entity's parts. Any attempt at a truly eliminative reduction of an emergent property, then, will suffer from a loss of relevant structure. It cannot succeed without invoking a particular *configuration* of lower-level entities – a particular set of parts and the relations between them – as the relevant causal factor. But if it does so, it has effectively reintroduced the higher-level entity into the analysis, since that set of parts and relations is simply the higher-level entity under a different description.

It should be clear that this argument in no way denies that we may be able to explain the relationship between higher and lower levels. It is not the attempt to explain higher-level properties that is reductionism's fault; it is the belief that explanation entails elimination. As Collier says, 'it is also important to avoid the temptation of thinking that a mechanism which explains another explains it away, so that the higher-level mechanism drops out of the scientific account.'[23] What this form of reduction *does* seek to do is to explain how the higher level can be emergent, by examining its parts and their relations to each other.[24] This is what we may call *explanatory* reduction.[25] The use of the term 'reduction' at all in this context is perhaps misleading, given the eliminative connotations it often seems to carry, but so many approaches to reduction take this form that it would seem idiosyncratic to refuse to use the term.[26]

One important consequence of this approach is that rather than eliminating higher-level theories, explanatory reductions do precisely the opposite: they provide extra justification for them by demonstrating that they are well founded in the theory of the lower level, that they are consistent with other accepted bodies of theory, and indeed that they extend their explanatory power.[27] In Gell-Mann's words, they are not eliminated but 'cemented'.[28]

---

[23] Andrew Collier, *Critical Realism*, London: Verso, 1994, pp. 109-10.

[24] This argument does *not* entail, however, that all emergent properties can necessarily be reduced, even in explanatory terms. Some reductions may be impossible because of the sheer complexity of performing them, and the concepts of multiple realizability and wild disjuncture, developed in the philosophy of mind, seem to provide cases where even explanatory reductions are impossible. See Jerry A. Fodor, 'Special sciences,, or: The disunity of science as a working hypothesis', *Synthese* 28, 1974, pp. 97-115; Sawyer, 'Emergence in sociology'.

[25] Bhaskar uses 'explanatory reduction' in a similar sense in *A Realist Theory of Science*, p. 181.

[26] Excellent arguments for non-eliminative reduction can be found in Fodor, 'Special sciences', p. 107.; Murray Gell-Mann, *The Quark and the Jaguar*, London: Abacus, 1995, p. 112; Donald T. Campbell, '"Downward causation" in hierarchically organised biological systems', in *Studies in the Philosophy of Biology*, eds F. J. Ayala and T. Dobzhansky, London: Macmillan, 1974, pp. 179-86.

[27] Philip Kitcher, '1953 and all that: a tale of two sciences', in *Philosophy of Science: The Central Issues*, eds Martin Curd and J. A. Cover, New York: W W Norton, 1998, pp. 971-1003; C. Meyering,

Now, the particular causal influences that any particular type of entity may exert, and the way in which the presence of its parts in the required relations produce these higher-level effects, are a matter for the particular science of the case – we cannot go any further at the philosophical level in explaining why particular cases of emergence work.

We *can* go further, however, in identifying another general prerequisite for emergence. As the existence of the whole is inseparable from the continuing presence of the required parts in the required arrangement, then emergence itself depends upon the set of causes that maintain a set of such parts in just such an arrangement. Bhaskar has commented on the dual aspect of emergence as a synchronic and diachronic relation; but the diachronic aspect of his account seems to relate to the original creation of the new level of reality. While this original creation is clearly necessary, the maintenance of the particular entities that constitute that new level is equally important. There is not only a set of causes that brings the entity about – what I will call, after Buckley, its *morphogenetic causes* – but also a further (possibly overlapping) set that maintains its continuing existence – its *morphostatic causes*.[29] It is these latter causes that are responsible for the *stability* of the organisation that constitutes the higher-level entity from the composing entities.[30]

The role of these morphostatic explanations of continuity of structure is critical to emergence. Any number of implausible combinations of lower-level entities may be brought about by a vast range of morphogenetic causes over the course of time, but it is only those combinations that have continuity of structure that persist. At any time, it is possible that a more powerful morphogenetic cause may overcome the morphostatic causes for any given entity. At this point, the emergence of the higher-level entity is dissolved.[31] It is the ability of morphostatic causes to resist such effects that sustains the existence of higher-level entities and hence any emergent properties they may have. But this continuing existence is always contingent on the outcome of the ongoing tension between these different types of cause.[32]

The argument in this section, then, implies that a full explanation of any emergent property must identify five elements, which I label the five 'pillars' of emergence: (a)

---

'Physicalism and downward causation in psychology and the special science', *Inquiry: An Interdisciplinary Journal of Philosophy* vol. 43, no. 2, 2000, p. 181.

[28] Gell-Mann, *The Quark and the Jaguar*, p. 112.

[29] Archer, 'Morphogenesis', p. 480, n8; Buckley, *Society*, p. 53. Morphostasis seems to be similar to Spinoza's concept of 'conatus', as expounded by Collier in *Scientific Realism*, p. 86.

[30] Some critical realists have suggested that a prerequisite of emergence is the existence of 'internal relations' (in the sense of 'necessary' relations) between the parts, e.g. Margaret Archer, *Realist Social Theory: The Morphogenetic Approach*, Cambridge: Cambridge UP, 1995, p. 173. Sometimes it seems to be implied that the stability of the internal organisation of an entity arises from the existence of such relations. The whole conception of internal relations as 'necessary relations', however, is problematic, and beyond the scope of this paper, although I hope to address it elsewhere.

[31] Bunge refers to this as 'submergence': Mario Bunge, *Emergence and Convergence*, Toronto: University of Toronto Press, 2003, pp. 31-3.

[32] There is a useful role in the explanation of morphostasis for concepts like 'negative feedback' from cybernetics and 'strange attractors' from complexity theory.

the parts of the entity that possesses the property; (b) the relations between these parts that constitute them into this entity; (c) the mechanisms by which these parts and relations produce the property – which we attempt to explain in an explanatory reduction; (d) the morphogenetic causes that bring the entity into existence; and (e) the morphostatic causes that sustain the entity's existence.[33]

## 4. *Composition*

Clearly the composition of entities by their parts is central to this conception of emergence, as it has been from the earliest versions of the concept, developed primarily in the late nineteenth and early twentieth centuries. McLaughlin, in describing the best-established of these earlier approaches, tells us that

> According to British Emergentism, there is a hierarchy of levels of organizational complexity of material particles that includes, in ascending order, the strictly physical, the chemical, the biological, and the psychological level. There are certain kinds of material substances specific to each level. And the kinds of each level are wholly composed of kinds of lower levels, ultimately of kinds of elementary material particles.[34]

Most emergentists have continued to take the view, as I do, that the concept of emergence is inherently compositional. By this I mean that any higher-level entity (and its emergent properties) is dependent upon a collection of lower-level entities in the sense that (a) they are the necessary component parts of the higher-level entity; (b) the emergent property is dependent upon (but not eliminatively reducible to) the properties of these parts; and (c) the emergent property, in the sense of a power or tendency, is not dependent upon the properties of other entities that are not such parts (although it may be so dependent for its realisation).

At times critical realists have adopted a compositional definition of emergence, yet at other times they have seemed to deny such a view. Let us begin with some examples of the compositional approach. In *Scientific Realism*, Collier has written: 'Bhaskar is explicit about the reality of the complexity of complex entities, whose powers are "emergent" with respect to their components.'[35] In *Critical Realism* he writes:

> As against atomism and holism, Bhaskar's emergence theory allows us to conceive of real, irreducible wholes which are both composed of parts that are themselves real irreducible wholes, and are in turn parts of larger wholes, with each level of this hierarchy of composition having its own peculiar mechanisms and emergent powers.[36]

---

[33] These five elements are not intended to be taken as mutually exclusive ontological categories. The purpose of distinguishing them as different aspects of emergence is purely methodological. As one reviewer has pointed out, this means that the same real things could appear under more than one of these headings.

[34] Brian P. McLaughlin, 'The rise and fall of British emergentism', in *Emergence or Reduction?*, eds Ansgar Beckermann, Hans Flohr, and Jaegwon Kim, Berlin: de Gruyter, 1992, p. 50.

[35] Collier, *Scientific Realism*, p. 51. Despite my criticisms of Collier's views on composition below, this book remains one of the most valuable accounts of emergence in the critical realist literature.

[36] Collier, *Critical Realism*, p. 117.

Yet elsewhere Collier has denied a compositional account of emergence, and Bhaskar has sometimes seemed ambiguous on the question. Let me discuss Collier's argument first.

In *Critical Realism,* Collier argues 'that many (though not all) cases of rootedness-emergence relations are also relations of composition.'[37] The basis of his argument, however, is clearer in his earlier work:

> A level of mechanisms depends unilaterally for its existence on lower levels – that is what it means to call it 'higher' and to call them 'lower'. *One* way in which a mechanism may so depend, is that it is a feature of structurata that [they] are composed of structurata governed by the lower-level mechanisms. But that is not the only way it may so depend. Language is composed of signs, not of people, but it is dependent on there being people, and people being governed by certain (biological, etc.) mechanisms. So there will be more strata in the hierarchy of vertical causality (dependence and emergence) than in that of composition. […]
>
> I think it is true that the human body is a structuratum with different elements from those of the mind (i.e. on the one hand, cells; on the other, cathected intentional objects) and that, in the hierarchy of dependence and emergence (vertical causality), there are two intervening levels (society, language), since these depend on the bodily organism, and are conditions of the emergence of mind.[38]

Collier's argument here is somewhat unclear, particularly due to the linkage of dependence and emergence – it is not clear, for example, why dependence should be considered hierarchical (unless it is merely a synonym of emergence) and this causes difficulties, for example, with reciprocal dependence. Collier appears to be asserting that entities such as languages and minds do not emerge only from their parts, but rather that they emerge from a hierarchy of other entities, some of which are their parts and some of which are not. And the basis of this argument is that their existence (or perhaps that of their emergent powers) depends not only upon their parts but also upon the other things included in this hierarchy.

There are at least three possible interpretations of this argument, and we must consider each in turn.

First, Collier may be arguing that non-compositional lower levels may have been essential causes in the initial development – the diachronic emergence – of the higher level. Certainly it is true that the causal history of any given type of entity is likely to include the effects of a vast range of entity types that are not its parts. This, however, is completely irrelevant to the question of whether or not the entity currently possesses emergent powers and to how those powers emerge in the synchronic sense. If this is Collier's meaning, then he appears to be conflating diachronic and synchronic emergence in this account – or, to put it differently, conflating *morphogenesis* with the *mechanism* of emergence itself. The use of 'is' in 'language […] is dependent on there being people', however, suggests that this is not his intention.[39]

---

[37] Ibid., p. 116.

[38] Collier, *Scientific Realism*, p. 99. Note that 'structuratum' is synonymous with 'entity'. Also see Collier, *Critical Realism*, p. 133.

[39] However, it is arguable that this statement only makes sense as an account of diachronic emergence, since language could continue to exist in books, films, and the like even if people ceased to do so.

Secondly – and this seems the most likely interpretation – he may be arguing that non-compositional levels may be essential causes in sustaining the continuing existence of the higher level. Again, it is true that the continuing existence of any entity is likely to depend causally on a vast range of entities that are not its parts. Indeed, such dependence relations are far more widespread than Collier's treatment suggests. Many biological organisms, for example, cannot exist without a breathable atmosphere. Rivers and oceans cannot exist without the gravitational force of a planet. Many people in the contemporary world could not continue to exist without the set of technologies that we depend upon for our food and other material needs. But once again, this is completely irrelevant to the question of whether or not an entity currently possesses emergent powers and to how those powers emerge. It is, once again, only if 'emergence' is given something like the diachronic rather than the synchronic sense, that it makes any sense to say that an entity or property or power emerges from the entities that contribute to sustaining its existence but do not interact to provide its powers. On this reading, Collier conflates the *morphostatic* causes of an entity with the mechanism of emergence.

The third possible interpretation is that Collier *is* thinking here of the role of non-compositional levels in the mechanisms of emergence themselves – although his use of 'existence' in the first of the two quotations above seems to conflict with this. Here we could read his argument as an assertion that the emergent powers of a certain type of entity (*E*) can only appear when not only the parts of *E*, configured in the way characteristic of *E*, but also certain other entities (*F*), external to *E*, and in a particular set of relations to *E*, are also present. This is a different claim from the claim that these other entities are essential to the morphostasis of *E*, since the morphostatic relationship is a diachronic, causal one, whereas the relationship that I am now discussing is a synchronic one.

Such a claim is clearly true when the property or power in question is the ability of an *E* to affect the *F*s concerned – the power of water to put out a fire, for example, cannot be exercised unless there is a fire to put out. This sort of case is usually dealt with in critical realism via Bhaskar's argument that causal powers can exist unexercised if the conditions of their exercise are not met, and treating the presence of an *F* as such a condition – and also by treating such cases as an interaction between the causal powers of an *E* and the causal liabilities of an *F* (where a liability is merely a passive type of power – a power to be affected in a certain way). Both of these approaches imply that an *E* and an *F* each has a relevant causal power, and that in cases of actual causation the effect that follows from these causal powers will only be realised if both the contributing powers are present. On this account, the real *existence* of the relevant emergent property of *E* is *not* dependent on any external entity, but only its actual *exercise*, and hence this would seem to sustain a compositional account of emergence.

A more complex version of the argument results if we suggest that a property of *E* might depend on the synchronic presence of an *F* even though the exercise of the property does *not* affect the *F* directly. This is a problem only if there really are such cases (and where these cannot be resolved by identifying the existence of a third

entity, composed of an *E* and an *F*, that is the real possessor of the property).[40] The role of society with regard to the emergent powers of our mind does not seem to be such a case – social entities clearly affect our behaviour, but they do so through the intermediate step of contributing causally to our knowledge or beliefs. This represents a morphogenetic effect on the structure of our brains/minds and not a synchronic dependence of our mental powers upon social entities.[41] It seems entirely possible, therefore, to see mental powers as emerging simply from human bodies – as, for example, does Archer – or more specifically, as emerging from the configurations of neurons that underlie our knowledge and beliefs.[42]

It is certainly true that establishing the hierarchy of emergence mechanisms that underlie language and mind (if the latter is a useful concept at all) is a challenge that still requires further work, and Collier's argument is clearly motivated by an attempt to meet this challenge. It is also possible that the composition relations that underlie these entities may be unusually complex. However, there is no necessary reason to believe at the moment that the resolution of these problems will require a non-compositional account of emergence.

A second argument that appears at first sight to deny the compositional basis of emergence appears in Bhaskar's *Dialectic*, where he argues that emergence

> consists in the formation of one or other of two types of superstructure (only the first of which has generally been noted in the Marxist canon), namely, by the superimposition (Model A) or intraposition (Model B) of the emergent level *on* or *within* the pre-existing one – *superstructuration* or *intrastructuration* respectively.[43]

Although here he talks in diachronic terms of the formation of structures and thus in terms of morphogenesis, the argument clearly implies that these structures continue to exist and possess synchronic emergent powers. A similar argument is expressed in directly synchronic terms by Bunge:

> P is an *emergent* property of a thing b if and only if either b is a complex thing (system) no component of which possesses P, or b is an individual that possesses P by virtue of being a component of a system (i.e. b would not possess P if it were independent or isolated).[44]

The former case would seem to correspond to 'superstructuration', and the latter to 'intrastructuration'. Bunge's formulation, however, has the merit of making clearer that there is still a compositional basis to intrastructuration: in such cases, the properties of an entity are altered as a consequence of it having become part of a particular type of whole. Here, these new properties of the part are still a consequence of the composition of the whole by its parts; all that is different from the usual case of

---

[40] The 'third entity' point was made by one of *JCR*'s anonymous reviewers.

[41] This does not mean I am denying synchronic emergent powers to social entities: see Elder-Vass, 'The emergence of social structure'.

[42] Archer, *Realist Social Theory*, p. 102. See also John Searle, *The Rediscovery of the Mind*, Cambridge, MA: MIT Press, 1992.

[43] Roy Bhaskar, 'Critical realism and dialectic', in *Critical Realism: Essential Readings*, eds Margaret Archer et al., London: Routledge, 1998, p. 599.

[44] Mario Bunge, *Finding Philosophy in Social Science*, New Haven ; London: Yale University Press, 1996, p. 20.

emergence is that it appears to be the part that is exhibiting a different property rather than the whole. Bunge argues, for example, that atoms change their form when they become parts of a molecule, rather than simply being held together while retaining their previous form.[45] A more significant example for critical realists might be the case of human beings who become parts of organisations, and whose 'causal powers' are changed as a result of their adopting a role in the organisation.

Despite its advocacy by both Bhaskar and Bunge, I have some doubts about the concept of intrastructuration. What is at issue here is the question of whether a property of an entity that is the consequence of it being part of a larger whole is really a property of the part at all, or whether it is just a property of the whole that happens to be localised in some respect within the part. What is seen in Bunge's account as a property of an atom which has become part of a molecule, for example, might be better represented as being a property of the molecule itself. This makes no difference, of course, to the argument that such changes in properties can occur, nor to the argument that they remain consequences of the composition of the whole by its parts.

Neither Collier's nor Bhaskar's concerns, then, undermine the compositional account of emergence. And the compositional account also has a number of epistemological advantages over Collier's alternative. First, it maintains a clear relationship between emergence and its primary theoretical usage: to maintain the tenability of a stratified view of reality in the face of eliminative reductionist arguments. Second, it maintains a degree of simplicity and hence clarity to the concept of emergence that makes it easier for us to understand its theoretical role. And third, a more detailed compositional account of emergence makes it possible for us to understand how emergence works in practice, as discussed in the previous section. None of these arguments, of course, constitutes a definite proof of the compositional account of emergence. But I believe this section has offered good grounds, both ontological and epistemological, for holding the compositional view.

## 5. *Cause*

The philosophy of emergence is unavoidably interwoven with the philosophy of causation. In the critical realist account of cause, there are two key elements – the concept of 'real' causal powers and the combination of the causal powers of different entities to produce 'actual' causation. This section will relate each of these in turn to emergence, and ultimately relate the two in a mutually interdependent interaction. Furthermore, it will argue that a careful examination of the relationship between emergence and cause may lead to the conclusion that the 'real' and the 'actual' have more in common than might be expected from the standard critical realist accounts.[46]

Let me begin with 'real' causal powers: Bhaskar identifies these with 'relatively enduring structures and mechanisms' that are 'nothing other than the ways of acting

[45] Bunge, *Emergence and Convergence*, p. 12.

[46] Martyn Hammersley has also queried the existence of a sharp distinction between 'actuality and reality', although on different grounds than those discussed below, in a presentation to the BSA's Realism Study Group, February 2005.

of things';[47] or in other words, 'the generative mechanisms of nature exist as the causal powers of things'.[48] These things 'are complex objects, in virtue of which they possess an ensemble of tendencies, liabilities and powers'.[49] Although this formulation does not directly invoke the concept of emergence, the relationship with emergence is clear: the powers and properties of an object or entity can be ascribed to the organisation of its parts into a particular kind of complex whole. In other words, real causal powers are emergent properties. This is why Bhaskar argues that 'explanation depends upon emergence'.[50] And Collier makes the connection still clearer:

> As against atomism and holism, Bhaskar's emergence theory allows us to conceive of real, irreducible wholes which are both composed of parts that are themselves real irreducible wholes, and are in turn parts of larger wholes, with each level of this hierarchy of composition having its own peculiar mechanisms and emergent powers.[51]

The second element in the critical realist account of cause is the combination of the causal powers of different entities to produce 'actual' causation, in which actual events are produced by a complex interaction of the causal powers of the entities involved. These interacting powers may belong to entirely distinct entities, but it is also important to recognise, as Collier does in the piece cited above, that the various entities that are the parts of the distinct higher-level entities involved also have causal powers. Any given higher-level entity, then, can be seen as a 'pyramid' of successively lower-level parts, and the causal impact of the higher-level entity as a whole includes the causal impacts of those parts. At each level, the entities formed from the lower-level parts have causal powers in their own right by virtue of how those parts are organised. The total causal impact of a higher-level entity conceived of in these pyramidal terms, then, includes the impact of all its lower-level parts as well as the causal powers that are emergent at its highest level.[52]

Bhaskar addresses this question of the contribution of causes operating at different levels through a concept which he calls 'dual control', 'multiple control', or 'multiple determination'. In considering actual natural and social events, he argues, we must accept that different causal mechanisms and the interactions between them account for different aspects of the events concerned, and that no single law 'determines' the whole result:

> Laws leave the field of the ordinary phenomena of life at least partially open [...] To say that laws situate limits but do not dictate what happens within them does not mean that it is not possible to completely explain what happens within them. The question 'how is constraint without determination possible' is equivalent to the question how 'can a thing, event or process be controlled by several different kinds of principle at once?' To completely account for an

---

[47] Bhaskar, *A Realist Theory of Science*, p. 14.

[48] Ibid., p. 50; cf. Tony Lawson, *Economics and Reality*, London ; New York: Routledge, 1997, p. 21.

[49] Bhaskar, *A Realist Theory of Science*, p. 51.

[50] Bhaskar, *Scientific Realism and Human Emancipation*, p. 104.

[51] Collier, *Critical Realism*, p. 117.

[52] The relation between the powers of an entity and the powers of its parts is discussed at more length in the section on 'level abstracted' and 'downwardly inclusive' views of entities in Dave Elder-Vass, 'Re-examining Bhaskar's three ontological domains: the lessons from emergence', paper presented at the IACR annual conference, Cambridge, August 2004.

> event would be to describe all the different principles involved in its generation. A complete explanation in this sense is clearly a limit concept. In an historical explanation of an event, for example, we are not normally interested in (or capable of giving an account of) its physical structure.[53]

Bhaskar's argument does not relate *only* to the relations between causal powers at different levels of a given multi-layered entity; he is also concerned with the interaction of causal powers between entirely distinct entities, at whatever level they exist. But the same framework does apply equally well to the interaction between causal powers at different levels of the same entity, and Bhaskar makes the link to stratification explicit in a more recent work: 'Emergence makes possible the important phenomena of *dual* and *multiple control*'.[54]

We can transpose this argument into more clearly stratified terms: it is precisely because 'the [actual] ordinary phenomena of the world' are inherently multi-layered, that we need to bring to bear different (real) causal mechanisms, each of which emerges at a specific level, to explain different aspects of them. Thus explanation at each level, in the 'area of autonomy' left by the incomplete explanations at other levels, requires a 'putatively independent science' of that level.[55] And it is in combining all these level-specific explanations of the different levels of a particular event that we 'completely account for an event'. Although, of course, because we do not have viable sciences of every level, we can only produce incomplete subsets of the 'complete' multi-layered account, which is why such a complete account can be seen only as 'a limit concept'. And in practice, we will not be interested in such complete accounts: we may be perfectly happy to explain an event at a given level while ignoring its lower-level ramifications.

Actual causation, then, is a process in which the real emergent causal powers of a variety of entities interact to produce events. Now, as we have seen above, the sustained existence of any individual entity and hence of its emergent causal powers is the outcome of a set of interacting morphostatic causes. This interaction is itself a process of multiply determined actual causation. Thus, at the same time as actual causation is a process of combining multiple instantiations of real causal mechanisms, those instantiations of real causal powers are themselves the result of a process of actual causation. Once we take full account of emergence, then, we can see causation as a tightly interwoven interplay between 'real' causal mechanisms and 'actual' causation.

The workings of real powers and actual causation, however, also intersect in another interesting way. One way to understand this is to revisit the five pillars of emergence outlined earlier. Alert readers may have noticed that one of these five pillars is not a strict prerequisite for the possession of sui generis properties by a collection of entities taken as a 'whole'. While an entity, by my definition above, must be a *persistent* whole, there is no logical reason why a group of entities that forms a *fleeting* or temporary whole should not have causal powers *as a group* that are not possessed by any particular entities in the group. If emergent properties are the consequence of the

[53] Bhaskar, *A Realist Theory of Science*, pp. 110-11.

[54] Roy Bhaskar, *Plato Etc*, London: Verso, 1994, p. 75.

[55] Bhaskar, *A Realist Theory of Science*, p. 114.

existence of a particular set of entities organised in a particular way, then the presence of emergent powers need not depend on that configuration persisting for an extended period. Of course, this fleeting whole would possess those powers only for the few instants during which the particular set of parts and relations required to sustain those powers was in existence, but for this brief moment this whole would possess pseudo-emergent powers. Let me call these *fleeting emergent powers*. More mundanely, a collection of entities may interact causally in a strictly summative way, in which case we might argue that the implicit fleeting whole would possess fleeting resultant powers.

Now there is a clear analogy here between fleeting interactions between groups of entities and the process of actual causation. Indeed, the process of actual causation simply *is* a fleeting interaction between groups of entities and their causal powers. The outcome of that interaction may be determined by the simple addition of the causal powers involved (as in the classic Newtonian parallelogram of forces), or it may be determined by a more complex non-linear interaction between them. These two cases correspond exactly to the ideas of fleeting resultant powers and fleeting emergent powers.[56]

To put the same point in a different way, any given event is the outcome of the actual interaction between the real causal powers of those entities causally involved in it, and the net outcome of these interactions depends upon the (purely temporary) relations in which these entities stand to each other at the time. This is directly analogous to the generation of the causal powers of a particular type of entity, which is the outcome of the interaction between the causal powers of its parts. The primary difference is that in the first case, the relations between the entities concerned are contingent and temporary, whereas in the second, the same set of significant relations is maintained over time as a result of the operation of morphostatic causes that maintain the structural stability of the entity, and hence there is a level of consistency in these causal powers over time.

Real and actual causation both therefore appear to be consequences of the same generic type of structural relation: the (diachronic) causal consequences that flow from a given set of entities existing (synchronically) in a given set of relations to each other. Actual causation, then, depends upon four of the five pillars of emergence: (a) a set of parts; (b) the relations between them; (c) the mechanisms resulting from the combination of these parts in these relations; and (d) the morphogenetic causes that bring this configuration of parts into existence at the moment of causation. Only the fifth pillar of emergence – the morphostatic causes maintaining the existence of this set of parts in this set of relations – is absent.

Now it may appear that this account neglects three important differences between real causal powers and actual causation. First, the former represent a *potential* arising from the relations between the entities concerned, whereas the latter represents the *realisation* of the potential inherent in the relations between the entities concerned.

[56] There is therefore a sense in which at least some cases of actual causation are fleetingly emergent from the combination of real causal powers of the 'lower-level' entities involved. Here I should acknowledge Tobin Nellhaus's suggestion in a personal communication that Bhaskar's domains might be emergent from each other, although I'm not sure whether he intended it in quite this sense.

Secondly, real causal powers are repeatedly instantiated in many different contexts, whereas instances of actual causation may arise from unique configurations of causal powers. And thirdly, in the former case, any given operation of the causal powers of the entity will have to interact with the powers of other entities causally involved in the event concerned, whereas in the latter the set of causal powers is by definition a complete set of those involved in the event's explanation.

The first difference, however, is more apparent than real, since both real causal powers and actual causation can be described in terms of potentials or in terms of their realisation. Thus, the particular set of entities involved in a case of actual causation has the potential to cause just the sort of event that it does if they are brought into the requisite set of relations, whether or not that event has yet happened, in just the same way that the particular set of entities that are parts of an emergent entity has the 'causal powers' of the entity whether or not those causal powers are being exercised, and would have them, should an entity of that type come into being, whether or not such an entity currently exists. Both types of potential represent a *natural necessity* for these entities to have those particular effects if and when they are brought into that particular set of relations to each other.

The second difference is no more fundamental. In particular, it is entirely possible for some actual configurations of relevant causal powers to be replicated on many occasions, resulting in similar events each time. This is, after all, the normal form of the scientific experiments which formed the basis of Bhaskar's transcendental deduction of real causal powers and actual causation in the first place.

The third difference, however, is more significant, and underpins the essential usage of the real vs. actual distinction in critical realist theory: the use of real causal powers as building blocks in the construction of explanations of actual events. It is inconceivable that we could produce viable explanations of events in terms of the unique configurations of entities at every level of stratification involved in each case unless there was some way of analysing such situations into interacting component parts. Now this argument may seem unattractive to critical realists, since it proposes an epistemological reason for preferring an ontological distinction. But it is an epistemological reason that is itself ultimately grounded in an ontological distinction: it is because of the persistence of entities and their consistent re-occurrence that their causal powers can be differentiated from the ongoing stream of actual causation, and hence it is because of that persistence and re-occurrence that it is valid to use those powers as building blocks in explaining actual causation.

The distinction between a 'persistent' or recurring whole and a fleeting one, is of course a matter of degree (and hence, on this argument, so is the distinction between 'real' and 'actual' causal potential). Although most cases may well be fitted easily enough into one category or the other, there is at least one interesting and important class of intermediate cases. These are the cases where a particular configuration with causal properties occurs first as an isolated incident – an apparently fleeting combination of entities – but subsequently acquires a set of morphostatic causes and hence is transformed into a persistent whole, an entity with real emergent causal powers. Where the acquisition of morphostatic causes is itself a path-dependent outcome, there may be no obvious inevitability to this transition.

This would seem to be the case, for example, in the development of certain social institutions. When the first boy picked up a football and ran with it at Rugby School, for example, there was no institution of 'rugby football' that standardised this practice and made its reproduction likely. But today such an institution exists. What was a fleeting configuration of causal entities has been transformed by the creation of a set of supporting institutions into a social practice with real emergent causal powers – the game of rugby football. Many social institutions may have developed in this way – compare, for example, the origins of the insurance industry in a coffee house in London – although in many cases the first 'actual' interactions are lost to history and so we see only the fully developed 'real' form.

Such intermediate cases merely confirm the close resemblance between real and actual causation. As I have argued above, this distinction is still valuable, but the resemblances between the two may be more significant than would seem to be implied by Bhaskar's division of them into two distinct ontological domains.[57]

## 6. *Conclusion*

The theory of emergence remains central to the critical realist ontology. The accounts of emergence in the critical realist literature, however, are often sketchy, and where they are more substantial, some aspects seem questionable. This paper has sought to take a step towards a more rigorous and complete theory of emergence, and has identified some specific issues with previous accounts. In particular, it has argued strongly for a compositional account of emergence. Nevertheless, the account of emergence developed here is both consistent with, and an enhancement of, critical realism's account of causal powers, actual causation and multiple determination.

The paper has also argued that a more careful consideration of the nature of emergence casts new light on the relation between real causal powers and actual causation. In particular, the radical distinction drawn between these two by critical realists is to some extent undermined by a recognition of some significant resemblances between the two.

While there remain issues to be resolved in the theory of emergence, I argue that it should be seen as an important – and hitherto under-developed – component of the critical realist ontology.

[57] Bhaskar, *A Realist Theory of Science*, Ch. 1.