Response to Clayton: Taxonomy of the Types and Orders of Emergence

Walter B. Gulick

ABSTRACT Key Words: Michael Polanyi, Philip Clayton, emergence, causality, supervenience, self-organizing systems, part-whole relationships, properties, boundary conditions, epistemological and ontological hierarchies, complexity theory.

Inappropriately reductive or deterministic appropriations of science haunt Philip Clayton's otherwise instructive appropriation of Michael Polanyi's thought for theological and ethical reflection. The work at hand utilizes contemporary complexity theory to augment Polanyi's notions of emergence and hierarchy and to provide a vision within which moral responsibility and theological inquiry make sense. It sets forth types and orders of emergence that bypass untenable notions of causality, reducibility, and determinism.

Philip Clayton has done a great service for the Polanyi Society in indicating how Michael Polanyi's claims about evolutionary emergence – one of the most controversial aspects of Polanyi's philosophical position[1] – stand in relation to several of the most discussed topics in contemporary philosophy of science and theology: emergence, supervenience, and (in the background) complexity theory.[2] It is clear that several of Clayton's overarching interests are consistent with issues Polanyi felt passionately about. Both thinkers will not be satisfied with any theory that undermines the freedom and responsibility of the person. Any vision of responsible personhood must leave a role for (if not exhaustively explain) mental causation. Each takes science very seriously, but also recognizes the limits of scientific explanation. Each ponders questions of ultimacy and affirms the significance of dimensions or levels of being beyond *homo sapiens*.

In the first three sections of his article, Clayton separates those aspects of Polanyi's understanding of emergence that he believes have proven fruitful from those claims that he sees as problematic. I'll begin by assessing his assessment, providing my own critique of Polanyi in the process. Then he presents a spectrum of recent views concerning how the mind and the person can best be understood from a scientific perspective. He argues that emergence theory is more faithful to the insights of both science and personal experience than any variety of physicalism, substance dualism or monism. I concur, but the notion of emergence to which Clayton alludes seems sufficiently inchoate that it is hard to move beyond his general point to specific claims. Clayton's notion of emergence is still susceptible to being undermined by reductive and inappropriately deterministic scientific theories. In the hope of assisting Clayton and others using notions of emergence to gain greater conceptual clarity, I attempt to sort out and simplify the different sorts of phenomena sometimes labeled with the term "emergence." I distinguish between the types and the orders of emergence. This twofold taxonomy is consistent with Polanyi's hierarchical vision of things and demonstrates that his thought, appropriately qualified, continues to provide a heuristically powerful philosophical framework for both scientific and theological inquiry.

Clayton on Polanyi's Accomplishments

Clayton describes three aspects of Polanyi's thought that he thinks make important contributions to contemporary conversation about emergence and three that he thinks have proven to be erroneous. I find that two aspects of Polanyi's thought that Clayton appreciates lead to flawed analyses and require modification. My criticisms are actually directed more at Polanyi than at Clayton. I will initially attend to Clayton's second point of appreciation of Polanyi before I turn to his first point.

(2) *The "from-at" transition and "focal" attention.* First, a clarification: What Polanyi terms the "from-at" structure of consciousness in "Life's Irreducible Structure" (*KB* 225-239) is no different than the "from-to" structure of consciousness, his usual terminology. It is unfortunate that Clayton relies fairly heavily on this article, because I find it one of Polanyi's less coherent explications. Clayton quotes Polanyi as follows: "Mind is the meaning of certain bodily mechanisms; it is lost from view when we look *at* them focally" (*KB* 238). Polanyi will be misunderstood here unless it is recognized that he is speaking of how we know other minds. When one turns within, mind is not the (focal) meaning of bodily functions but the process whereby meanings are produced. Mind is at least potentially what the phenomenologists call appresented in all its activities, whether it is engaged in looking (retrospectively) at its own activities, perceiving the world, or reading a book. That is, one can never make the mind in its subjectivity a direct object of inquiry, but one can always ascribe its underlying activity to any state of consciousness.

The use of Polanyi's thought that Clayton goes on to announce also seems problematic. Clayton says, "In my recent work defending panentheism I have used a similar distinction between focal intentions and the body's autonomous functioning . . ." (9). However, a person's body does not function autonomously in the from-to structure. Rather the embodied subsidiaries are conjoined with the focal meaning in a dialectical dance of mutual dependence. Moreover, from a Polanyian perspective, our guiding intentions are not more than momentarily focal, but are usually dwelt in subsidiarily. They function as standards of achievement. Clayton's point, however, seems to be this: intentions and other ordinary mental activities occurring at a higher level than the body's many autonomic processes need not interfere with these lower level processes. Analogously, he seems to suggest, God can be seen as making cosmic decisions and taking actions in ways that do not disrupt autonomous human activity or violate natural law. This is an intriguing analogy, but not without its dilemmas. Isn't human freedom seriously compromised if one correlates autonomic processes with autonomous human activity? And in using this analogy is Clayton comfortable with apparently giving up the traditional Christian claim that God listens to people and interacts with the world?

From an emergence point of view, also, Clayton's analogy seems suspect. When one shifts attention from a given level to an emergent level, the laws and structures of the higher level are inevitably quite different than the laws and structures of the lower level. But Clayton's analogy patterns the divine mind on the human mind, which at least makes the analogy look suspiciously more like a sample of projected anthropomorphism than a common type of emergence. Clayton is quite aware of this sort of criticism, as Andy Sanders makes clear in his discussion of Clayton's *God and Contemporary Science* in this issue, but the questions raised by his analogy remain unresolved. It seems that the way Clayton appeals to Polanyi's from-to analysis of consciousness is not yet sufficiently thought through to see if it might be helpful in relation to the foregoing problems.

(1) *Active and passive boundary conditions.* Clayton, in discussing foreground and background interests, and active and passive constraints, is helpful in his exposition of Polanyi's thought. However, his claim that Polanyi recognizes two types of boundaries — natural processes and machines — is not quite accurate, but I think Polanyi is the primary cause of confusion. Polanyi does use natural processes and machines as examples of boundary conditions. He says these "useful restrictions of nature [involve] the imposing of *boundary conditions* on the laws of physics and chemistry" (*KB* 226). The problem lies with the ambiguous notion of boundary condition, which Polanyi generalizes unhelpfully beyond its original meaning in physics. Sometimes boundary conditions seem to be spatial in nature, sometimes inherent properties, sometimes functions, sometimes imposed operational principles. In what sort of instance would boundary conditions be imposed on the laws of physics and chemistry rather than on the properties of the object being analyzed? Clayton's exposition is clearer than Polanyi's discussion.

Polanyi gives the following examples of the two types of boundary condition:

> When a saucepan bounds a soup that we are cooking, we are interested in the soup; and, likewise, when we observe a reaction in a test-tube, we are studying the reaction, not the test-tube. The reverse is true of a game of chess. The strategy of the player imposes boundaries on the several moves, which follow the laws of chess, but our interest lies in the boundaries – that is, in the strategy, not in the several moves as exemplifications of the laws. (*KB* 226)

In context, it is clear that Polanyi equates the first two examples above with machine-type boundaries, where our interest lies "in the effects of the boundary conditions." But we aren't really interested in the fact that the physical form of the saucepan – the boundary imposed on the soup – shapes the soup in a certain way. Nor is our primary concern about any effect of the test-tube. Our interest in each case lies at a different level than what would ordinarily count as a boundary: with preparing lunch to eat, or with testing some materials for some purpose. To be fair to Polanyi, his intention no doubt was to refer to the shaping of a saucepan and a test tube for certain purposes even as one designs a machine for a certain purpose, but even so, the design shapes the properties of metal and glass; it doesn't bound the laws of physics and chemistry. In his example of the chess game, Polanyi uses a dynamic example (not a natural process) that also does not illustrate the imposing of boundary conditions on the laws of physics and chemistry but rather on the rules of a game. What is significant about chess is essentially mental: it involves the strategic manipulation of certain rule-bound pieces which have conventional physical shapes but may just as well be represented on paper or by name.

There is another feature of Polanyi's two-tier analysis of machines that I find to be erroneous. He asserts that the operational principles of a machine "can never account for the failure and ultimate breakdown of the machine" (*TD* 39). The physical-chemical structure of the machine is said to account for failure (*TD* 39 again, but see *PK* 330, where Polanyi more accurately states that "physics and chemistry are blind both to success and failure"). Well, sometimes the operational principle is at fault, as for instance when the machine is poorly designed, accounting for some but not all the conditions in which the machine must operate. Firestone engineers, for example, would love to blame the laws of physics and chemistry for the accidents caused by their faulty tire design. Operational principles should not be considered abstractly apart from the conditions of operation.

Clayton's Criticisms of Polanyi

I find Clayton's three criticisms of Polanyi, which overlap and are actually variations on the one theme of vitalism, to be right on target. Polanyi is not a vitalist in the sense that he postulates the existence of some life substance that is responsible for biological evolution. However, his misguided emphasis on guiding principles (*M* 169), his attraction to the ideas of Driesch (Allen, ed., *SEP* 295, *PK* 338), and his affirmation of creative agency (*TD* 46) shows he is sympathetic to non-substantive varieties of vitalism. This problematic tendency towards vitalism seems to have its origin in Polanyi's extrapolation from his analysis of machines to the realm of biology (thus see the title of section 2, *KB* 226: *Living Mechanisms are Classed with Machines*). Just as some intention or design must pre-exist a machine, so by analogy he tends to argue that the ordering principles of a living being must pre-exist the function being ordered. Thus Polanyi is led to the untenable claim that phylogenetic emergence is activated by "an ordering principle capable of producing operational principles which the system had not previously possessed" (*PK* 399). This shadowy ordering principle seems more like a first cousin to Platonic Ideas than a reliable foundation for scientific knowledge. To be sure, the demonstrations of bottom-up emergence now central to complexity theory had not been developed when Polanyi wrote *Personal Knowledge*. Complexity theory shows how emergent, stable patterns may spontaneously develop in large, multi-faceted systems; the existence of mysterious pre-existing principles need not be postulated. Polanyi's real concern, however, is denying that life can be reduced to the laws of physics and chemistry, and his inadequate ways of expressing this legitimate concern do not invalidate the basic structure of his vision.

TWO APPROACHES TO UNDERSTANDING EMERGENCE

Polanyi's usage of the notion of emergence within the evolutionary perspective of part IV of *Personal Knowledge* was innovative for its time. But in recent years, there has been an explosion of interest in emergence, and the term has been used in a variety of ways. One of the reasons that discussion of "emergence" quite understandably irritates many scientists is that the term is generally not clearly defined and in some discussions — for instance, theological discussions — has taken on an honorific coloring. In the hope of facilitating a more nuanced use of the term, I will offer two related taxonomies of emergence. First, I will set forth three distinct ways in which the term "emergence" has been used as a counter to reductive programs of analysis. That is, I'll delineate three *types* of emergence. Then, secondly, I will essentially follow Polanyi in distinguishing topologically three fundamental *orders* of emergence. Each order manifests a distinctive causal power.

Types of Emergence

It is characteristic of all usages of emergence to postulate a coming into existence of something different in kind from what existed before and not causally or logically reducible to the supporting conditions out of which it appeared. Here I have articulated what I will stipulate to be two essential characteristics of emergence: 1) an emergent entity or state operates according to higher level principles or rules different from the lower level rules of the entities or state out of which it emerged, and 2) there is what Heinz Pagels termed a causal decoupling between the emergent phenomenon and its antecedents[3] – in other words, the emergent cannot be causally reduced to its lower level support such that beginning from the lower level one might accurately predict what must follow from that initial set of conditions. Here, then are three ways in which an emergent entity can be claimed to conform to these defining characteristics.

1. Part-whole emergence. The first conception of emergence to be explored is homologous with Polanyi's notion of levels, is consistent with ecological notions of interdependence, and also manifests many characteristics of Gestalt theory, for it is grounded in part-whole analysis. It requires that one begin by starting from some entity or system that exhibits a functional coherence or stability.[4] That is, one must begin by identifying a whole rather than an aggregation. In its original context, the selected whole will be seen to have an inside set of parts, a descriptive unity, and an outside set of properties. The fundamental claim to be demonstrated is that the outside set of properties is not *causally* reducible to the inside set of parts. The properties are an emergent phenomenon.

The freedom of an interpreter to select a whole for a chosen purpose gives part-whole sorts of emergence a broad range of applicability. Locating wholes that lend themselves to fecund analysis is a matter of skill in knowing. While anything may be selected as a whole, consideration of arbitrary collections as wholes will likely not be fruitful because the components have no internal dependencies, systematic relationship, or naturally re-enforced unity. But when appropriate material wholes are selected, the outside will appear as a set of *properties* and /or law-abiding processes that come to expression out of interaction with its particular environment.[5] A richer vocabulary is utilized when investigating living things or their functioning systems; the outside may be seen as *characteristics, traits, behaviors,* etc., depending on the nature of the organic whole and its environment being investigated.

The chosen whole's inside is comprised of *components* and an organization or *structure*. The nature of the inside can be comprehended through some analytic procedure. In contrast, the whole's outside represents how the whole "appears" in the context of the other wholes and forces forming the initial whole's environment and existing on roughly the same scale (that is, in a neighborhood). The inside components influence but are not causally responsible for the outside properties.

Here a brief digression on the history of causal language seems in order. The successful identification of the laws of mechanics early in the history of modern science accorded prestige to experimental science where the number of variables to be studied was severely limited. The language of causality was used to describe how changes in those limited variables impacted the subject or system under study in the experiment. But while the language of causality is useful in a mechanistic context where variables are limited, it becomes misleading when applied to either the whole constructed from parts or the complex systems of the real world evolving in ways better understood ecologically than mechanistically. A subsequent variation on the mechanistic language of causality is the logical language of necessary and sufficient conditions. Both the mechanical language of cause and effect and the logical language of necessity and sufficiency contributed to positivistic dreams of reductionism, of simplifying our understanding of the empirical world to expressions of the laws of physics and chemistry. In the twentieth century that dream fractured from discoveries on many fronts, of which quantum theory and chaos/complexity theory are perhaps the best known.

If causality is an inappropriate notion to use with respect to how a whole is emergent in relation to its parts, what sort of explanation (beyond the vague "parts influence or participate in the whole") illuminates the relationship? I believe both epistemological and ontological insights can arise from considerations of how knowledge unfolds at different levels of interrelatedness. Epistemological insight begins with the observation that humans tend to focus on those significant functional wholes that illuminate their projects of the moment. They name them: "tree," "chair," "cloud," bravery," "party," "truck." These noun/wholes may then be clothed with adjectival qualities: "tall," "comfortable," "fluffy," "admirable," "political," "red." The adjectives tend to describe *properties* of the identified whole rather than parts (components organized in structures). Interest in parts is of a different order than daily commerce with things and events. Attention may be shifted from wholes to parts if there are problems to be solved or if scientific curiosity is aroused.

The significance of knowing a whole in terms of properties versus knowing it in terms of parts may be illustrated by comparing the recognition of a face with finding the right words to express a complex thought – a sort of poetic challenge.

> a) In recognizing the visage of an acquaintance one sees a familiar pattern – perceived image is correlated with mental image. In Polanyi's language, one tacitly indwells presented facial features and integrates them into the focal recognition of Jane or Joe as identified whole.
> b) However, in attempting to find the words to express adequately a still vague and hunch-like thought, one attends to the wholeness of the (still schematized and as yet not fully symbolized) thought and evokes words that are tried out to see if they say what one means. When one finds and organizes properly the right words, one feels a sense of satisfaction: the parts comprise the intended and previously mute whole.

In contrasting these examples, one sees that a face is known in terms of its visible *properties*, whereas articulating what one means to say involves construction out of organized *parts* (words structured grammatically). Two different epistemological procedures are involved, one leading to recognition of what is already known, and the other leading to discovery.

Polanyi's discussion of heuristic gradients grounded in tacit knowing and culminating in discovery or problem solving is illuminated by part-whole emergence in general[6] and the example of finding the right words in particular. In setting up a scientific program of research, a scientist attempts to identify an as yet incompletely understood whole in which the contributing parts and their structures and functions, only tacitly sensed, guide the scientist with a heuristic premonition that the way the whole functions is solvable. The more that the particulars of the whole are located and indwelt, the closer the investigator feels to comprehending through integration the meaningful whole that is the target of investigation. Once the indwelt subsidiaries merge coherently into a meaningful whole, the scientist experiences an "Aha!" Now attention can be shifted to making the parts explicit, just as the poet can attend to fine tuning the words once they begin fitting what the poet intends to say. The scientist can attend to the laws, material components, and arrangements that jointly constitute the no longer mysterious whole being investigated. Epistemology here harmonizes with ontology: just as the parts, laws, and arrangements are integrated by the knower to form the whole under investigation, so the relationship of the discovered parts to the whole is integrative rather than causal in nature.

The foregoing observations are but an application of Polanyi's comments, found for instance in *Science, Faith and Society*, about the skill required to identify a fruitful scientific problem and follow its tacitly sensed clues through to discovery.[7] It should be reiterated, vis-à-vis Clayton's threefold critique of Polanyi, that the final identification of wholes is a human activity made possible through symbols. While the identification and exploration of already existent wholes is a useful tool of discovery, contemporary complexity theory, consistent with evolutionary theory, does not postulate any pre-existing whole toward which a natural system must evolve.

A further ontological point, like the epistemological/ontological point just made, involves a change in perspective, but the shift is of fundamental importance. It involves the distinction between composition and property. The outside appearance of any whole is effectively determined by the nature of the environment within which it dwells. In an acidic environment, the alkaline quality of the whole will stand out. In an environment of pockets of food available to a bacterium as the whole, cellular properties of directional motility will not only stand out but be selected for. In a waiting room, the soft fabric and ergonomic qualities of a chair as the whole will stand out. In each of these cases, the qualities of the environment make a certain few of many potential properties stand out. When a human being is present in the environment, the particular interests and perceptual abilities of that person will make particular affordances (to use J. J. Gibson's term) of the whole stand out. In relation to what occurs in the environment, the properties that stand out determine what transpires; the composition of the whole is generally irrelevant. That is, causality in the situation is a product of the interaction of the evoked or out-standing properties of the wholes at a certain level, not a bottom-up production in which the parts create the whole. The relation of part to whole is not causal but emergent, distinguishing a self-sufficient yet supportive inside from an environmentally articulated outside.

2. Transformational emergence. Sometimes emergence seems to be used as a synonym for process or change in general. It is usually reserved for dramatic changes in some self-identical entity where the rules of the emergent entity are quite different from the rules governing its predecessor state. But because there are no clear criteria of what constitutes emergence, the tiniest of changes could be called emergent. That is, Whitehead's process philosophy could be regarded as a philosophy of emergence. Each newly concrescent entity has unique features different from its predecessor (at least in terms of its cosmic placement). In a strained sense of the term, it manifests new principles. But it must be stated right away that transformational emergence has the weakest claim of any of the three types to be called emergence. If the criterion of causal decoupling is insisted upon, it has no claim. Let me explain.

Whatever the extent of change, the focus of transformational emergence is upon an entity in which the initial and final stages of change are understood as existing at roughly the same scale. Understanding transformational emergence begins by attending to the significant properties of a whole. What happens to this whole in a neighborhood is traced over a period of time. The emergent whole interacting with other entities at the same ontological level may at some point cross a threshold into a dramatic new form of organization. Thus, for instance, the mass of a large collapsing star may become so dense that it collapses into itself, forming a black hole. Is the black hole, so different than other astral phenomena, an emergent entity? Because the transformation is of the same massed material and it occurs in sequence at the same ontological level, there should be a continuous thread of causality stitching together each step of the process of change. Thus, while the first criterion of emergence (markedly different operational principles or rules) may apply to this dramatic example of transformational emergence, the second criterion (causal decoupling) would not seem to apply. Moreover, if this change-related notion of emergence is to be even marginally useful, some broadly agreed upon criteria of what counts as sufficient change to be considered emergence needs to . . . emerge.

When one examines the interaction of a selected whole with other dissimilar wholes on the same scale in the neighborhood, one will often find a mere collection having no particular order. Yet over time (sometimes it may take eons) all entities are affected to some degree by their environment. The effect may take on at least three forms:

> a) The structures supporting the unity of the old whole may erode or decay because of long term interaction with water, exposure to heat, or other processes in the environment, effectively destroying the old whole as a unity. New wholes may form as byproducts of the change. Example: granite decays into clay and sand.

b) The whole may gradually – or dramatically — change as a result of external forces and environmental interactions. Because the evolved entity will not instantiate exactly all the same laws nor have just the same characteristics as it originally had, the new whole may be said to have emerged (in the weak sense) from the old whole. Example 1: the horse evolves from a small North American mammal into the magnificent beast it now is. Example 2: snow falls on a northerly mountain, is buried, and the pressure turns it into ice.

c) With skill one can sometimes identify a new stable pattern in which the whole one initially selected is captured and becomes a part contributing in some way to a new higher level whole, supporting thereby what Polanyi calls a "comprehensive entity."[8] The principles and rules governing the comprehensive whole will be different from the principles and rules governing the original whole, just as the principles and rules governing the original whole will be different from the principles and rules of its parts. The comprehensive entity may obtain a stability such that the original whole is always surrounded by the same other wholes and always involved in the same limited functions. In such a system, it will in effect lose its independent functioning. Causal efficacy will lie with the comprehensive entity. It would be a mistake to see the initial whole as having a bottom-up effect on the comprehensive entity. Transformation occurs in this instance, but the comprehensive entity that emerges is best analyzed in terms of either the first or third type of emergence. Example 1: copper ore is mined, smelted, and made into wiring for an electric motor. Example 2: snow falls on a northerly mountain, is buried and turns into ice, and becomes part of a glacier that severely erodes a mountain valley.

It can be seen from the two snow examples that what happens at one level may have an effect at another level. In fact, all things function at a variety of levels simultaneously. This view is consistent with Polanyi's hierarchical model of the world expressed, for instance, in his description of the making of bricks (*TD* 35) or the structure of a living being (*TD* 36-37). It is exemplified in the from-to structure of consciousness. But I do not believe this unfolding interpretation of emergence is burdened by the problems found in Polanyi's discussion of boundary conditions.

3. Self-organizing emergence. The third broadly accepted notion of emergence has risen to prominence out of complexity theory in general, and out of the study of self-organizing systems in particular. Self-organizing systems typically have the following components: material entities having properties (shape, charge, mass, etc.); forces relevant to the scale of the components under consideration (for humans, gravity and electromagnetic forces; for the nucleus of an atom, the strong and the weak force); specific energies (kinetic, chemical, etc.); and positive and negative feedback loops that facilitate growth, homeostasis or decay in relation to specific environments. It is an overall pattern of systemic functioning that may be said to be emergent. The emergent system operates according to rules different in type from the rules governing its various components, and there is a causal decoupling between the system as a whole and its contributing parts.

If part-whole emergence describes the structure of existent forms of emergence, self-organizing systems explain how emergent phenomena originally come into being. Self-organizing systems may arise, like transformational emergence, out of interactions at about the same scale, but what distinguishes them from examples of transformational emergence within a neighborhood is that the process of change issues in an entity or system that takes on a life, a dynamism, of its own. The continuity characteristic of a changing entity is broken, or perhaps better, is reconfigured (bifurcated) into a new sort of continuity. Different species of living beings are examples of self-organizing systems; their members receive feedback from their environments and regulate their processes so as to fit into their environments in ways that maintain their integrity. Indeed, they seek out and colonize environmental niches where they are best able to thrive.

The world is home to many complex systems balancing order and chaos through adaptive response systems: cities, ecosystems, economic systems, the "social insects" (ants, bees), the human immune system. Most frequently, self-organizing systems seem to rely on a shift in scale that allows patterns to emerge out of large numbers of entities – molecules, structural iterations, independent human decisions. Entities at the micro level have certain properties as individuals, but when they are taken en masse and interwoven with other neighboring entities, a macro set of properties and laws is observable in self-organizing systems. Of course, when some entities are mixed together, nothing stable emerges; the mixture remains chaotic rather than self-organizing. Complexity theory is concerned with those emergent patterns that not only are intelligible but display features that make large scale patterns useful in the macro world. In other words, the emergent pattern of the self-organizing system is selected for, survives, and carries out new functions totally different from the functions of its miniscule parts.[9]

This third type of emergence takes place within human culture as well as non-human nature. The organization of capitalism in Adam Smith's analysis would be a good example. At the micro level each person within market capitalism seeks to advance his or her own interests by finding the best investment opportunities. Money functions as the crucial feedback mechanism in this view; the cost of goods and their selling price provide a potential investor with information as to where the best opportunities to make a profit lie.[10] The well-known paradox is that although each individual seeks to maximize personal interests, the system as a whole benefits society as a whole because it provides consumers with goods and services previously in short supply.[11] Smith talks about these large-scale benefits in terms of an "invisible hand." In his economics writings, Polanyi brings the discussion of capitalism nearer to current understandings of self-organization by speaking of capitalism as an example of "spontaneous order."[12]

Emergence, Supervenience and Human Existence

The foregoing review of types of emergence has been conducted more at the level of a cosmological inquiry than with the focus on God and person that characterizes Clayton's essay. The cosmology that has resulted is supportive of Polanyi's hierarchical view of reality. Polanyi articulated a simpler version of part-whole emergence than the one presented here. He described self-organizing systems even if he did not use that terminology or focus on their broad implications. So although I agree with Clayton's threefold critique of Polanyi, in general I find Polanyi's thought confirmed by the recent work in emergence theory. But what difference does it make for speaking about responsible personhood? Has progress been made towards articulating the emergentist theory of mind that Clayton seeks?

With respect to scientific understanding of the mind and person, the relationship between brain and mind has been at the center of discussion. Clayton finds the notion of supervenience useful to describe the relationship. There are a number of different interpretations of supervenience, but they all tend to accept the following two claims: that mental properties are higher level (supervenient) phenomena different in kind than physical lower level properties, and that the higher level properties are dependent on the lower level properties in a way such that "two things indiscernible with respect to lower-level properties, are indiscernible with respect to their higher-level properties."[13] Jaegwon Kim, the most influential exponent of supervenience theory, adds what he calls the "dependence thesis" to the notion of supervenient invariance: "[I]t is an explicit affirmation of the *ontological primacy,* or *priority*, of the physical in relation to the mental, thereby opening the possibility of *explaining* the mental in terms of the physical."[14]

If my analysis of part-whole emergence was adequate, then certainly strong supervenience (by definition expressing the dependence thesis) must be rejected. Subvenient brain states support (as organized parts) but do not cause supervenient mental states (properties). What of the type-type correlation between brain state and mental state? Well, certainly I would want to claim that mental life is dependent on brain activity even if not reducible to it. So it would appear I would reject strong supervenience while accepting weak supervenience, thereby following Clayton. Not so. The problem with weak supervenience is that it really explains nothing. I accept the token-token correlation of strong supervenience, but I reject the ascription of causality that normally accompanies discussion of supervenience (and certainly is present in Kim). To explain my position more fully with respect to the causality involved in human thought, I will undertake two different sorts of analysis. First I will develop a useful analogy by exploring how water functions as a part and whole both as an emergent entity and a component of an emergent system. Then I will set forth three different orders of emergence that illuminate the uniqueness of human cognition.

Consider a molecule of water to be the beginning whole. It is composed of two hydrogen atoms bonded at a standard angle with one oxygen atom. The lower level properties of hydrogen and oxygen (the inside) do not cause the higher level properties of water (the outside given definition by an environment). Among the characteristics of water that may stand out as properties are its translucence, its polarity which dissolves many ionic compounds, its liquidity, its surface tension, its greatest density at a temperature above freezing, its relatively high boiling point, etc.

So much for a simplified version of water's part-whole emergence. I will now turn from a small scale analysis of a water molecule to a large scale analysis of water's participation in a system that exhibits some simple self-organizing characteristics: the ocean. The properties of water molecules, the overwhelmingly most numerous parts of an ocean, influence the ocean's dynamic structure. Oceans contain many dissolved salts because water molecules are strongly polarized. As a consequence, emergent but static properties of ocean water include a lower freezing point, a saltier taste, and a higher capacity for buoyancy than fresh water in a lake. Seawater's properties are emergent according to the first type discussed above. An ocean exists in a neighborhood that includes the sun's radiation, shorelines, an uneven bottom, and an air mass above it. Ocean water in tropical zones absorbs heat from the sun and becomes thereby less dense than the colder water of arctic regions. The colder water is moved by gravity to replace warmer water, and thus ocean currents come into being. Differences in heat and density also affect the air over the water and create winds that in turn create waves that in turn erode shores. Now, obviously, the properties of individual water molecules do not cause all the functions and behaviors of an ocean. To understand waves and currents, for instance, one needs to see the impact of forces at roughly the same macro-scale as the ocean.

An instructive analogy can be made using the relation between water molecules and oceanic erosion to illuminate the relation between brain neurons and a human state of mind or action. A cluster of water molecules in a large wave may actually crash against a cluster of rock molecules in a seaside cliff with enough force to dislodge bits of quartz, future grains of sand. Do we then say that the water molecules caused the rock to fracture? This would be to apply the notion of causality to the wrong level of influence. Much more explanatory power would come from stating that, say, the fury of a storm that was caused by a large atmospheric low coming from the north created the surf that caused erosion at high tide. In a certain way, the specific water molecules and rock molecules are innocent entities used by larger scale forces and conditions. The water molecules did not independently cause anything; their velocity was caused by wind and wave.

The analogy to human consciousness is no doubt obvious. A given thought or action may be seen as derived from the firing of a certain set of neurons as connected by linked filaments and also connected by neurotransmitters acting across certain synapses. But it makes no more sense to say that this one system of neurons among the brain's trillions of neurons and uncountable connections caused the thought or action than it makes to say a particular cluster of water molecules caused the rock to fracture. One more appropriately seeks to understand the thought or action by asking what large scale event caused or evoked that response, just as one more appropriately asks what caused the large waves than focuses on the wave's molecules in order to understand shoreline erosion. Yes, a particular set of neurons has a token-token relation to a particular thought, but each set is the supportive component of much larger systemic forces known within consciousness as intentions, motives, drives, mental models, and the like. It is this sort of point I believe Clayton is making when he opts for a dynamical systems approach to understanding conscious experience. There is no reason not to declare that these larger scale mental forces are causally implicated in the rise of certain thoughts and actions.

I do not want to leave the discussion of supervenience without noting how misleading the concept is for comprehending the mind-brain relationship. The very suggestion that identical lower level configurations of the brain must result in identical higher level mental experiences replaces empirical evidence with logical considerations in a continuing manifestation of the tendency to logicism characteristic of twentieth century analytic philosophy. Empirical studies have demonstrated that no two brains are alike because brains grow in ways that capture the experience and learning of individuals, and no two individuals have identical experiences. While a supporter of supervenience theory might protest that the thesis of identical correlations is only a thought experiment, it encourages a view that human brains are identical in much the way that different CD-ROMs of the same computer game are identical. It also encourages those construing consciousness as comparable to software to interpret the designed programs as being the cause of the virtual reality of the game experience. In contrast, I argue for a view that all the levels opened up by first and second order emergence (to be described shortly) are co-equal in reality even if the defining forces and sorts of material/energy vary from level to level. That is, quarks are no more real than molecules are no more real than cells are no more real than organs (to give a highly condensed summary of emergent ontological levels).

Orders of Emergence

If discussion of emergence is restricted to the three types I have distinguished (part-whole, transformational, and self-organizing emergence), then the world will be seen to be populated with all sorts of emergent phenomena (innumerable if the second, weak form of emergence, transformational emergence, is acknowledged as a true form of emergence). Terrence Deacon has taken the lead in sharpening discussion of emergence by distinguishing between three basic species or orders of emergence.[15] His criterion of classification is topological, that is, based on relationships that differ in *type of causality*. Deacon's categorization seeks to highlight significant junctures in the evolutionary development of the cosmos. In this respect it is compatible with Polanyi's even stronger bias toward evolutionary theory. Polanyi's interpretation is centered about explaining anthropogenesis (*PK* 386ff), an orientation compatible with Clayton's concern to deal with God and persons. Since my aim in this discussion is to critique helpfully Clayton's paper, my description of the orders of emergence is closer to Polanyi's account than to Deacon's. I distinguish three different orders by attending to the type of causality manifested in integrations or resolutions, moving toward autonomous causality by starting from mechanical causality, noting the emergence of signal responsiveness, and finally arriving at symbolic creativity.

First Order Emergence: Dynamo-Physical World. The first order of emergence describes processes at all levels of the material world. The complete realm of materiality with all its systems and subsystems is involved in the dynamo-physical world, but nowhere do self-replicating autonomous agents arise within this world. Rather first-order emergence concerns the developments and arrangements made possible by first, second and third types of emergence excluding the emergence of life. Events within the dynamo-physical world culminate in dumb and mute resolutions at both micro- and macroscopic levels, involving entities as divergent in scale as quarks and quasars. Typically resolutions emerge that occupy the lowest energy state possible in a given environment. In terms of classical Aristotelian notions, first-order emergence approximates formal and material causality.

Second-Order Emergence: Biological World. The second level of emergence arises when what Polanyi calls "active centres" (*PK* 336ff) come into being as living entities (primitive autonomous beings). Single-celled organisms introduce originally very primitive purposes into the world, telic behaviors which Polanyi claims can be distinguished from natural laws because they can succeed or fail (*PK* 331, 345; *SM* 60). As one ascends the evolutionary ladder, living things are coded with increasingly complex means of assessing and adjusting to their environments by finding food, avoiding enemies, and reproducing. Successful living can be perceived at an individual level – can the particular biological entity respond to the signals evident in its environment so as to survive? Can it learn to thrive? Success is more importantly measured at a species than an individual level, and here what is especially noteworthy is that all members of a species share in a family of genetically transmitted traits that allow the species to colonize a particular ecological niche and eventually either succeed or fail as a group. Living beings thereby transcend the mute resolution of forces characteristic of first-order emergence; they introduce a species of self-preserving, interest-centered causality into the world. In Aristotelian terms, final causality has entered the cosmos.

Third-Order Emergence: Human World. It seems highly presumptuous to claim for humans, puny in the vastness of the universe, one of three cosmic orders. But when one starts off intending to consider how human experience arises (anthropogenesis) and what, if anything, is unique about it, and if one uses as a criterion for distinguishing orders of emergence degrees of autonomous causality, then such a triumphalist conclusion seems unavoidable. Human consciousness, in its distinctiveness, is symbol-infused, and this allows for a selective purposefulness that transcends second-order emergence. Through indwelling symbols, a person is able to imagine what is not and what could be. Through symbols, humans are able to select between alternative visions for considered reasons, and this is the very essence of free choice. Symbols open up a considered past and alternative futures. They accumulate in memories and congeal as habits, and thereby they form cultures that shape human behavior. Thus humans engage in what Goodenough and Deacon[16] describe as niche construction: humans create the language and cultural worlds (Polanyi, following Teilhard, calls this the noosphere – *PK* 388) in which they live and have their being. Humans use the power of symbols to probe the material world and reshape it technologically in myriad ways. Humans comprehend and manipulate the blind processes of the first-order world as it emerges from mysterious, minute particles. Humans use symbols to create machines with operational principles comparable to the telic principles guiding biotic beings exhibiting second-order emergence. No other entities use symbols to manipulate the world as humans do. Thus symbol usage allows for self-generated causality, a kind of emergent activity not highlighted by any of the three types of emergence. Human causality is uniquely autonomous in the known universe; it introduces conscious efficient causality to flesh out Aristotle's three other notions of causality.

Clayton and Causality

In his section, "Emergence, Supervenience, and Personal Knowledge," Clayton glimpses a notion of responsible personhood again and again, but then his language (and its associated conceptuality) betrays him and he falls back into some quandary. I will close by offering two related samples of this sort of conceptual regression and by re-emphasizing how emergence theory might yet save the day.

After describing Van Gulick's notion of radical-kind emergence, Clayton states that it is "much more amenable to modern science to say that the emergence of new macro-properties is ultimately determined by the sum total of relations between the micro-properties of that system. To know the state of all the registers in your computer *just is* to know the state of the system, even though the *content* of the system may be a new emergent property such as a digitized image of the Mona Lisa" (14). In the first sentence, the problematic term is "determined." In the second sentence, one of Clayton's errors is to imply that a state of computer elements is equivalent to (or causally linked to) an identifiable image. Let me elaborate.

Complexity theory makes it evident that the behavior of self-organizing large scale systems can be ordered but not necessarily predictable or reductively deterministic. What theorists call strange attractors or other phenomena may influence self-organizing systems like the brain to produce non-deterministic order. But more to the point when considering human consciousness, thoughts supported by some system of neuronic connections do not interact with thoughts supported by some other system of neurons in a manner characteristic of first-order emergence. It is the character of the thoughts themselves that influences subsequent states of consciousness, not the state of their supporting neurons, much less the state of the neurons' supporting molecules. Recall the wave and water molecule example. At the level of consciousness, there are a number of factors influencing future thought, and they generally come to expression at the same scale so they may be integrated or rejected. Within the temporal thickness of consciousness, sense data, reflections, purposes, anxieties, memories, etc. jostle for dominance. To think that subvenient factors *determine* unfolding mental states reveals one is ignoring the part-whole distinction between supportive parts (lower level phenomena) and emergent properties of mental life. It is to apply a simple mechanistic model when a far more sophisticated approach is called for, an approach comprehended better through models of organic growth, evolution, and ecology than through billiard balls or F=ma.

Perhaps applying part-whole analysis to the second sentence I have just quoted from Clayton will prove helpful. If we take as our whole the digitized image having a certain pattern, the emergent property of the shape in this case would not be even conceivable on the strictly material plane. It takes a person to recognize the image's significance by identifying the shape as Mona Lisa — a same-scale recognition of the given shape as being isomorphic with a remembered image of Mona Lisa. To turn to an inside analysis, the array of pixels would be reliant upon a certain state of computer registers (and if one just looks at the inside of the perception of Mona Lisa considered as the whole, it would be supported by a certain set of neuronic firings). Again, the state of the computer registers creates a pattern of pixels but not an emergent image; only a human, symbol using ("Mona Lisa!") third-order emergent mind can do that. And that recognized image is fully consistent with Van Gulick's radical-kind emergence: it is a real feature of human experience not determined by the law-like activities of neurons but by the recognition of a similarity between a sensed image and a remembered image.

A second quotation from Clayton reveals how persistent is the lure of the language of determinism and causality for him, even when his interest is in human thought and action. "The major *non*-quantum accounts of how mental causation might occur appeal to non-linear dynamics, chaos theory, and the field of complexity studies. One problem is that all three of these approaches are still deterministic"(15). Again Clayton stumbles over an all too logically confined, level-ignoring notion of causality. Were he seriously to embrace emergence theory, especially third-order emergence when discussing human thought and action, he would not be faced with the obstacles to understanding his mechanistic conceptuality confronts him with. There are, of course, threads of continuity and connection observable in the three theories he cites. I say "of course," because the alternative would be a world of unintelligible chance and arbitrariness, and we do not experience such a random sort of world. But continuity and connection need not entail some anxiety producing determinism that proclaims first-order emergent systems have no independent integrity, second-order emergent systems are incompatible with purpose, and third-order emergent systems leave no room for human freedom. One of the lessons of emergence theory when it is set in a hierarchical view of the cosmos is that there are alternatives to rigid determinism and random chance. It is a lesson that Clayton's survey of the field helpfully prepares the way for. I await with great interest the theological vision Clayton might set before us if he were to use a more fully developed theory of emergence such as the one I have attempted.[17]

Endnotes

[1]Marjorie Grene is strongly critical of Polanyi's notion of evolutionary emergence. She states that when she had an opportunity to review evolutionary theory carefully, "I found Polanyi's argument (of Part IV of *Personal Knowledge*) even more shocking than I had originally thought it" ("Reply to Phil Mullins," in Randall E. Auxier and Lewis Edwin Hahn, eds., *The Philosophy of Marjorie Grene*. The Library of Living Philosophers, Vol. XXIX [Chicago and LaSalle, IL: Open Court, 2002], p. 61).

[2] Complexity theory stands for a perspective on a broad range of issues much more than it represents any generally agreed upon or precise set of theorems. Most fully identified with the work of the Santa Fe Institute, complexity theory emphasizes the unfinished, transitory nature of the universe in contrast to the Newtonian hope of discovering the limited set of defining laws that govern the universe. Emergence theory can be seen as a part of complexity theory. When thus considered, the orderly understanding of emergence sought in this response may seem quixotic. Santa Fe Institute economist Brian Arthur states, "What we're missing at the moment is any precise understanding of how complex systems operate" (quoted in M. Mitchell Waldrop, *Complexity* [New York: Simon & Schuster, 1992], p. 334). Robert May echoes this view. "It's becoming apparent that a theory of complexity in some sense of a great and transforming body of knowledge that applies to a whole range of cases may be untenable" (quoted in Georgina Ferry, "Sir Robert May: Complexity and Real World Problems," *Santa Fe Institute Bulletin* 16:1, 1.

[3] Heinz R. Pagels, *The Dreams of Reason: The Computer and the Rise of the Sciences of Complexity* (New York: Bantam Books, 1989), pp. 222-223.

[4] Percy Hammond provides a helpful overview of some of the issues confronting anybody wishing to make use of part and whole analysis. He reviews Peter Simons' book, *Parts: A Study in Ontology*, which concentrates much more on parts than wholes, and he judges Simons' mathematical approach to modeling to be at best only partly successful. See Hammond's "Parts and Wholes – Contrasting Epistemologies," *Tradition and*

Discovery XXVIII:3 (2001-2002), 20ff. My approach to identifying useful wholes is experimental in nature; one might use what Polanyi calls intuition to spot promising wholes and see if their analysis is fruitful. See Polanyi's "The Creative Imagination" in Marjorie Grene, ed., *Toward a Unity of Knowledge* (New York: International Universities Press, 1969), pp. 60-67 for Polanyi's discussion of intuition.

[5] Stuart Kauffman cites an example from Phil Anderson that nicely exemplifies the irreducibility of a property. "Gold is a yellow, malleable metal familiar to all of us. Nowhere in the quantum mechanical description of atomic gold are these macroscopic properties to be found. Moreover, there is no deductive way to arrive at these macroscopic collective properties from the underlying quantum mechanics of atoms of gold" (*Investigations* [New York: Oxford University Press, 2000], pp. 127-128). My reflections on the notion of properties began well before complexity theory came into being. I have long teased introductory philosophy classes with the query as to whether the chair we perceive or the chair formed by molecules is more real. As will become apparent, I've answered this variation on the appearance/reality (or primary/secondary quality) question with the claim that each is equally real, but differently real in terms of the requirements of unique ontological levels.

[6] Polanyi states, "I have identified the antecedents of problem-solving with the process of emergence" (*TD* 87).

[7] "There must be a sufficient foreknowledge of the whole solution to guide conjecture with reasonable probability in making the right choice at each consecutive stage. The process resembles the creation of a work of art which is firmly guided by a fundamental vision of the final whole, even though that whole can be definitely conceived only in terms of its yet undiscovered particulars" (*SFS* 32).

[8] See Phil Mullins' as yet unpublished manuscript, "The Comprehensive Entity as a Key Idea in Polanyi's Thought" (delivered at the Loyola Polanyi Conference in 2001) for a helpful description of how the notion of comprehensive entity comes to play a significant role in Polanyi's mature thought.

[9] Steven Johnson speaks of self-organizing systems as "complex adaptive systems that display emergent behavior. In these systems, agents residing on one scale start producing behavior that lies one scale above them: ants create colonies; urbanites create neighborhoods; simple pattern-recognition software learns how to recommend new books. The movement from low-level rules to higher-level sophistication is what we call emergence" (*Emergence: The Connected Lives of Ants, Brains, Cities, and Software* [New York: Scribner, 2001], p. 18).

[10] While in recent years a number of economists have begun understanding economic systems on the model of self-organizing systems, Jane Jacobs draws out the implications of this approach with panache and power I've not encountered elsewhere – see her *The Nature of Economies* (New York: Vintage Books, 2001).

[11] Johnson says a fundamental law of emergence is "the behavior of individual agents is less important than the overall system" (Ibid., p. 145).

[12] See Michael Polanyi, *The Logic of Liberty* (Indianapolis: Liberty Fund, 1998; reprint of University of Chicago Press edition, 1951), pp. 190-202, for a sustained discussion of spontaneous order. Struan Jacobs sets Polanyi's thought about spontaneous order in historical context in his "Michael Polanyi and Spontaneous Order, 1941-1951," *Tradition and Discovery* XXIV:2 (1997-98), 14-28.

[13] Dennis Bielfeldt, "The Peril and Promise of Supervenience for the Science-Theology Discussion," in Niels Henrik Gregersen, Willem B. Drees and Ulf Görman, eds., *The Human Person in Science and Theology* (Grand Rapids: Eerdmans, 2000), p. 129. I commend to Clayton John A Teske's "The Social Construction of the Human Spirit" in the same volume for a vision of the human spirit that would help flesh out Clayton's notion of personhood.

[14] Jaegwon Kim, *Philosophy of Mind* (Boulder, CO: Westview Press, 1998), p. 11.

[15] Deacon's argument is spelled out in erudite detail in "The hierarchic logic of emergence: Untangling

the interdependence of evolution and self-organization," a draft of a chapter to be published in B. Weber & D. Depew, eds., *Evolution and Learning: The Baldwin Effect Reconsidered* (Cambridge, MA: MIT Press, forthcoming.

[16] Ursula Goodenough and Terrence W. Deacon, "From Biology to Consciousness to Morality," article forthcoming in *Tradition and Discovery* and *Zygon.* This article, the Deacon article just mentioned, and Goodenough's *The Sacred Depths of Nature* (New York: Oxford University Press, 2000) have been important background resources for my thought about emergence. However, as mentioned, my delineation of the orders of emergence is parallel to Polanyi's account and differs slightly from Deacon's configuration. Roughly speaking, my first-order emergence combines Deacon's first and second-orders, my second-order is equivalent to his third-order, and my third-order underscores my conviction that the causal potency of symbolic thought is not done justice by seeing it as the complex product of Deacon's second and third-order emergence.

Notes on Contributors

Philip Clayton (pdclayton@mail.peoplepc.com) **is** Ingraham Professor at the Claremont School of Theology and Professor of Philosophy at the Claremont Graduate University . He was the featured speaker at the 2001 Polanyi Society annual meeting and his work is treated in this special issue of *TAD*. An extended biographical statement is included on page 7.

Walter Gulick (WGulick@msubillings.edu) is Professor of Philosophy, Humanities and Religious Studies at Montana State University, Billings. He became a member of the Polanyi Society at its founding in 1972 and is today the General Coordinator of the Polanyi Society and Book Review Editor for *TAD*. Many of his articles and reviews have appeared in issues of *TAD*.

Martinez J. Hewlett (hewlett@u.arizona.edu) is a molecular virologist, a philosopher and a novelist who is Professor Emeritus in the Department of Molecular and Cellular Biology at University of Arizona in Tucson. During his career, Hewlett has been awarded several research grants and awards. In addition to 30 scientific papers and book chapters, he recently published a novel. In addition to science classes, he team taught a course in science and theology, worked in a program on philosophy of medicine, and has been involved in a variety of programs on natural science and religion.

Gregory R. Peterson is Assistant Professor of Philosophy and Religion at South Dakota State University and is author of the recent book, *Minding God: Theology and the Cognitive Sciences*. His mailing address is Box 504, Scobey 336, SDSU, Brookings SD 57007. He may be contacted by email at greg_peterson@sdstate.edu

Philip Rolnick is now Associate Professor of theology and moral theology at the University of St. Thomas in St. Paul Minnesota. His published works include *Analogical Possibilities: How Words Refer to God* and articles on religion and science. He is currently finishing a manuscript on the theology of personhood.

Andy F. Sanders (a.f.sanders@theol.rug.nl) is Associate Professor of Philosophy of Religion at the University of Groningen, the Netherlands. He is the author of *Michael Polanyi's Post-Critical Epistemology. A Reconstruction of Some Aspects of "Tacit Knowing,"* (1988), as well as the author or editor of several other books and many scholarly journal articles on topics in philosophy, theology and religion and science. He has been a frequent contributor to *TAD* and provided a plenary address at the 2001 Polanyi Conference at Loyola University.