Top-down causation and emergence:
some comments on mechanisms

George F R Ellis, University of Cape Town

Abstract: *Both bottom-up and top-down causation occur in the hierarchy of structure and causation. A key feature is multiple realisability of higher level functions, and consequent existence of equivalence classes of lower level variables that correspond to the same higher level state. Five essentially different classes of top-down influence can be identified, and their existence demonstrated by many real-world examples. They are: algorithmic top-down causation; top-down causation via non-adaptive information control; top-down causation via adaptive selection; top-down causation via adaptive information control; and intelligent top-down causation (the effect of the human mind on the physical world). Through the mind, abstract entities such as mathematical structures have causal power. The causal slack enabling top-down action to take place lies in the structuring of the system so as to attain higher level functions; in the way the nature of lower level elements is changed by context, and in micro-indeterminism combined with adaptive selection. Understanding top-down causation can have important effects on society. Two cases will be mentioned: medical/ health care issues, and education, in particular teaching reading and writing. In both cases an ongoing battle between bottom-up and top-down approaches has important consequences for society.*

1: Complexity and Emergence

Physics underlies all complexity, including our own existence. How is this possible? How can our own lives emerge from interactions of electrons, protons, and neutrons?

The basis of complexity is ***modular hierarchical structures***, leading to emergent levels of structure and function based in lower level networks. Each of these aspects ("modular", "hierarchical", and "structure") is crucial in the emergence of complexity out of interactions between simpler units (Booch 1994; Ellis 2008). The basic principle is that when you have a complex task to perform, you break it up into sub-tasks that are each simpler than the overall project, requiring less data and less computing power, and assign these tasks to specific modules. Each module is again split up into submodules until you reach a base level where the necessary tasks are simple operations that can be carried out by simple mechanisms. This is the level where the real work is done, each of these components feeding its results into the next higher level components until the desired result emerges at the appropriate higher level. The modules at each level will interact with each other in some way: maybe just statistically, if they all carry out the same task, or maybe in the form of a complex interaction network, when they will generally each carry out different tasks. The result is a highly structured hierarchy of interacting entities.

A simplified version of the basic hierarchy of complexity and causality for natural systems (left) and for human beings (right) is given in Figure 1. There is a similar hierarchy for artificial systems (Simon 1992) - language, mathematics, computers, aircraft, cities, organisations, societal roles for example - that is not given here. Each level of the hierarchy is made up of interacting modules that are relatively strongly bound internally, with higher frequency and higher energy internal dynamics, interacting with other modules through weaker bonds and lower frequency interaction dynamics. The internal dynamical variables are hidden from outside view; external entities interact with the module through interface variables linked to the internal variables.

Examples are nuclei in atoms, cells in a human body, individuals in society, and subroutines in a computer program. The way they interact with each other at a specific level can be characterised by an interaction network showing which modules interact with which other modules through various possible interaction modes (inter-module forces, or matter, energy, and information exchange); this is the structure of the system. Emergence of a higher level system from the lower level modules takes place when reliable higher level behaviour arises out of the lower level actions taking place in the context of this structure, with lower level units grouped together to form higher level modules that one can identify as meaningful entities that persist over time and have identifiable laws of behaviour. Identifying modules in really interaction networks may not be obvious; various algorithms have been devised for this purpose. As implied above, a crucial feature is the possibility of recursion: each module may itself be made up of an interacting network of lower level modules (see for example the description of the hierarchical structure of life in Campbell and Reece, 2005).

Each of the different levels of the hierarchy function according to laws of behaviour appropriate to that level, and are describable only in terms of language suited to that level (the concepts that are basic to molecular biology, such as genes and proteins, cannot be described in the language of a particle physicist, such as quarks and gluons). Ideas applicable to lower level causation do not by themselves succeed in explaining the higher level behaviours, for the concepts employed are simply not appropriate to the higher level kinds of causation. Higher level entities, such as plans and intentions, have causal power in their own right, which partially determine what happens at lower levels in the hierarchy (billions of atoms and molecules move in accord with our intentions when we raise our arm). Here, we characterise as level as "higher" when it can be shown to influence another level ("a lower" level) by setting a context which guides the way the lower level actions take place.

Multifold causation takes place in such systems. A network of causal influences and constraints interact to produce an outcome. In order to understand such systems, we often take for granted most of these influences and concentrate on one of them, which we then label as `the cause', meaning the dominant cause. But a web of influences and multiple causations is in action all the time. Nevertheless, in order to understand what is going on, it is useful to single out particular links in this causal pattern, taking all the rest for granted; indeed we have to do so, for it is not possible to explicitly take into account all causal factors at all levels of the hierarchy. Note that as is discussed later, ``causal power" is not restricted to ``efficient causality."

There are limits to what can be achieved by bottom-up emergence alone. Self-assembly and self-structuring based in bottom-up action alone can lead to emergence of structures such as crystals and simple biological molecules, or of dynamical systems with attractors leading to entities such as stars and galaxies, and to more complex phenomena such as Bénard cells, patterns associated with the reaction-diffusion equation and sand piles, the dynamics of the Game of Life, properties of slime mould, existence of ant colonies, and the behaviour of flocks of birds. These however do not extend to truly complex systems such as a single living cell. It seems that developing very complex systems such as those occurring in biology requires top-down causation, needed in order to build up the necessary biological information (Kuppers 1994, Roederer 1995). This information cannot be derived in a bottom-up way, because it implicitly embodies information about environmental niches. It would be different in a different environment. Hence higher level conditions influence what happens at the lower levels, even if the lower levels do the work. This is what I characterise as top-down causation.

This viewpoint of interacting bottom-up and top-down influences is very helpful in understanding complex systems, and has significant implications for our models of society and consequent social policy. This paper aims to outline the importance of top-down causation in the functioning of complex systems, and particularly in humans. This significance is illustrated in other papers in this special issue. I will in particular make the case that symbolic systems are causally effective variables at the higher levels of causation, when the hierarchy is suitably extended to include such abstract entities (as is necessary in order to satisfactorily represent causal influences emanating from the social level). Two examples will be given to illustrate possible social implications of the viewpoint presented here.

2: Bottom-up and top-down effects

Both bottom-up and top-down causation occur in the hierarchy of structure and causation. Bottom-up causation is the basic way physicists think: lower level action underlies higher level behaviour, for example physics underlies chemistry, biochemistry underlies cell biology, and so on. As the lower level dynamics proceeds, for example diffusion of molecules through a gas, the corresponding coarse-grained higher level variables will change as a consequence of the lower level change, for example a non-uniform temperature will change to a uniform temperature. However while lower levels generally fulfill necessary conditions for what occurs on higher levels, they only sometimes (very rarely in complex systems) provide sufficient conditions. It is the combination of bottom-up and top-down causation that enables same-level behaviour to emerge at higher levels, because the entities at the higher level set the context for the lower level actions in such a way that consistent same level behaviour emerges at the higher level.

A key concept here is ***coarse-graining*** of lower level variables to give higher level variables (Penrose 1989, 2005), with consequent loss of detailed information, thus enabling higher level behaviour to emerge from lower level properties. One averages lower level properties and thereby determines higher level properties, for example the pressure and density of a gas emerge from the underlying molecular spatial and velocity distribution. The opposite is fine-graining - we look at the situation on finer and finer scales. However there is not enough information in the coarse grained view to determine what fine grained state we will discover –this an essential consequence of the information hiding that occurs when we adopt a higher level view. There are many fine-grained states that can realise any specific coarse –grained state (this is `multiple realisability' of the higher level state, see the next paragraph). Some physicists and philosophers claim that from a fundamental viewpoint, higher levels are "nothing more than" an aggregation of lower level phenomena: i.e. they will all emerge by coarse-graining. However some higher level causally effective variables cannot be obtained in this way, as they are demonstrably not coarse grainings of lower level variables. To accommodate the effect of these holistic higher level variables, the hierarchy must be understood not simply as one of increasing complexity or scale, but as one of lower to higher levels of causation, whether associated with physical entities or not.

Top-down causation takes place due to the crucial role of context in determining the outcomes of lower level causation. Higher levels of organization constrain and channel lower level interactions, paradoxically thereby increasing higher level possibilities. A key feature here is ***multiple realisability*** of higher level functions, and consequent existence of ***equivalence classes*** of lower level variables as far as higher level actions are concerned. An equivalence class identifies all lower level states that correspond to the same higher level state. For example, billions of different micro-states will correspond to the same higher level state of a gas, as characterised by its temperature, pressure, and density; the same higher state of a computer, as characterised by gate states and currents flowing in its components, can be realised by many different specific electronic states; numerous different molecular configurations give the same functional state of a neuron in a human brain. The possibility of coherent higher level action emerging from the lower level dynamics is based in the ***principle of equivalence classes***: the same higher level state leads to the same higher level outcome, independent of which lower level states instantiates the higher level state. This is the crucial feature characterising effects as being due to top-down causation (Auletta, Ellis, and Jaeger 2007). If different outcomes result from different lower level realizations of the same higher level state, we do not have reliable same level action resulting from top-down influences of the higher levels (the lower level actions do not mesh together to cause reliable higher level behaviour, as is the case in chaotic systems). One can also consider the possibility of higher level causation by choice among, or ongoing constraint on the future of, lower-level phenomena which may not be multiply realizable. However in practice there will not be many cases where the higher level state is not multiply realizable: this is a consequence of the huge number of lower level components that make up complex systems, because of the atomic nature of matter and the cellular nature of living beings.

Note that `top-down' causation is an interlevel concept: it applies between any two adjacent levels. Claiming it happens is agnostic as to whether there is or is not a topmost or lowest level.

Characterising top-down causation

To characterise some specific causal effect as a top-down effect, we must demonstrate that a change of higher level conditions alters the sequence of processes at lower levels; we do this by changing higher level conditions and seeing what happens at the lower levels (for example we decrease the volume of a gas and see that it makes molecules move faster). We should if possible demonstrate existence of equivalence classes of lower level effects that give the same higher level outcome (for example we show how to determine gas density, pressure, and temperature from integrals over molecular variables); this is solid evidence that top-down causation is happening, and shows how to coarse-grain lower level variables to obtain higher level effective variables. Emergent higher level behaviour occurs when such higher level variables determine outcomes without any recourse to lower level variables (for example, the gas laws relating pressure, density, and temperature are stated purely in terms of higher level variables); the number of lower level states realising a single higher level state determines the entropy of the system (Penrose 1989, 2005). However in some cases, there exist causally effective higher level conditions that cannot be obtained by coarse graining of any lower level variables (for example, the rules of chess control the way chess pieces are allowed to move on a chess board, thus they are causally effective; but they are not determined by any lower level variables).

Five kinds of top-down causation: Five essentially different classes of top-down causation can be identified, and their existence demonstrated by many real-world examples (Ellis 2008). They will be discussed in turn.

TDC1: Algorithmic top-down causation

Algorithmic top-down causation occurs when high-level variables have causal power over lower level dynamics through system structuring, so that *the outcome depends uniquely on the higher level structural, boundary conditions, and initial conditions*. The lower level variables determine the outcome in an algorithmic way from the initial and boundary conditions (for example the software loaded in a computer) as a consequence of the structural relations (for example the wiring in a computer, or interconnections of neurons); changing these conditions leads to different lower level events and dynamical outcomes. Provided the lower level interactions mesh together in a coherent way, the constrained operation of lower level forces, operating in a law-like/algorithmic way, leads to reliable higher level behaviour whose outcome depends on the nature of the constraints and initial conditions. These are often in the form of networks of interactions (Barabási and Oltvai 2004), usually including recurring network motifs (Alon 2007). These are higher level features because they cannot be described in terms of lower level concepts (for example, the specific connections between transistors in a computer cannot be described in terms of properties of electrons) and the system ceases to function if the higher level relationships are disrupted, even though the individual lower level elements are unchanged.

Examples: An excellent example – indeed the present day canonical one - is ***digital computers***: the low level gates and transistors act in accord with the data and programme loaded (word processor, music programme, image processing programme, etc), which is a high level concept whose structure and function cannot be explained in lower level terms. The hardware and software are each hierarchically structured in a symbiotic way so as to allow this higher level functionality (Tanenbaum 1990). A second example is the way the outcomes of many physical systems is determined by partial differential equations, where the outcome depends on the boundary conditions and initial conditions (see the articles by Denis Noble and Robert Bishop).

The mathematics and theory underlying these algorithmic effects is varied: it includes dynamical systems theory (Busemeyer 2011), partial differential equations theory (Courant and Hilbert 1962), numerical methods such as finite elements (Brenner and Scott 2007), statistical physics (Huang 1990), the analysis of computer algorithms (Knuth 2010), electronic circuit design (Jaeger 1997), and the analysis of network motifs (Alon 2007).

TDC2: Top-down causation via non-adaptive information control

In non-adaptive information control, higher level entities influence lower level entities so as to attain specific fixed goals through the existence of feedback control loops, whereby information on the difference between the system actual state and desired state is used to lessen this discrepancy. Unlike the previous case, *the outcome is not determined by the boundary or initial conditions; rather it is determined by the goals,* indeed the whole purpose of such systems is to make initial conditions irrelevant. A different outcome will occur if the goal is changed. Thus the nature of causality is quite different than the previous case, when feedback control systems are guided by goals, which are higher level entities. This contrasts dramatically with how physics is usually considered to operate, but is fully in accord with engineering and biological precepts. The goals are established through the process of natural selection and genetically embodied, in the case of biological systems, or are embodied via the engineering design and subsequent user choice, in the case of manufactured systems.

Examples: An excellent example is a thermostat controlling the temperature in a room***;*** the goal is set by setting a desired temperature on an input panel. All of biology embodies numerous genetically determined homeostatic systems, based on the principle of feedback control; in particular this is true of the human body: homeostasis is the key to physiological research (Rhoades and Pflanzer 1989). Thus for example we have inbuilt bodily systems that interact to maintain body temperature at 98.4F to high accuracy.

The mathematics involved is linear and non-linear control systems theory, e.g. DiStefano, Stuberub, and Williams (1995), including its applications to the biological context, e.g. Iglesias and Ingalis (2010).

TDC3: Top-down causation via adaptive selection
Adaptive processes (Holland 1992) take place when many entities interact, for example the cells in a body or the individuals in a population, and *variation takes place in the properties of these entities, followed by selection of preferred entities that are better suited to their environment or context*. Higher level environments provide niches that are either favorable or unfavorable to particular kinds of lower level entities; those variations that are better suited to the niche are preserved and the others decay away. Criteria of suitability in terms of fitting the niche can be thought of as *fitness criteria* guiding adaptive selection. On this basis a *selection agent* or *selector* (the active element of the system) accepts some of the variations and rejects the rest; these selected entities then form the current system state that is the starting basis for the next round of selection, ultimately leading to the emergence and nature of biological form. A different lower level structure will result if the higher level context is changed.

Thus this is top-down causation from the context to the system. An equivalence class of lower level variables will be favored by a particular niche structure in association with specific fitness criteria; if the top level conditions change, the outcome will change. Unlike feedback control, this process does not attain pre-selected internal goals by a specific set of mechanisms or systems; rather it creates systems that favor the meta-goals embodied in the fitness criteria. This is an adaptive process rather than a control process. It is the way new information is generated that was not present before (Kuppers 1994), and enables emergence of complexity with an increase of embodied information, for the process searches the solution space in a way that is not pre-ordained and adapts to the context. The outcome is usually not predictable either from the initial conditions or the meta-goals, because of the random element involved, although both clearly influence the outcome. This underlies all life, including cells and plants and animals, and is the basis for building up biological information – the foundational difference between physics and biology (Roederer 2005).

Example: Darwinian evolution is a specific example: the standard story on the evolution of life is that increasingly complex structures have evolutionary advantages precisely because they constrain the lower level interactions that they are subject to (Campbell and Reece 2005). Cell walls are the most striking case, but the principle applies from bio-polymers to societies. It results in DNA structuring via adaptive selection over geological timescales, with the meta-goal - the higher level `purpose' that guides the dynamics - being survival of populations of the organism (both higher level concepts). The development of DNA codings (the particular sequence of base pairs in the DNA) occurs through an evolutionary process which results in adaptation of an organism to its ecological niche; the selector is death, the implicit fitness criterion is survival. A different niche structure results in a different set of genes. As a specific example: a polar bear *Ursus maritimus* has genes for white fur in order to adapt to the polar environment, whereas a black bear *Ursus americanus* has genes for black fur in order to be adapted to the North American forest. The detailed DNA coding differs in the two cases because of the different environments in which the respective animals live.

This is a classic case of top-down causation from the large scale context to detailed biological microstructure - through the process of evolutionary adaptation, the environment (along with other causal factors) fixes the specific DNA coding. There is no way you could predict or explain this coding on the basis of biochemistry or microphysics alone. The survival of the organism is the fitness criterion, leading to existence of all those detailed conditions which must be fulfilled for survival to be assured. This meta-goal is the same for every organism because it is what leads to the existence of populations that fit environmental niches better than competitors. Note that the claim is not that the environment is the *only* relevant factor; rather it is that it is *one* of the causally effective factors. There will always be multiple causal factors, some bottom-up and some top-down; the final result comes from the confluence of these effects. How do you demonstrate this is top-down causation? - Change the niche structure (e.g. by changing the global climate), and a different population will adapt to it. The fact that lower level equivalence classes for the same higher level purposes are selected by higher level conditions and resulting niches is demonstrated by many examples of convergent evolution (Conway Morris 2005) and studies of microbiology (Jaeger, this issue)

A key feature of the biological world is that similar processes of adaptive selection take place not only on evolutionary time scales, but also on developmental and functional timescales (Gilbert 2006); for example brain plasticity is based in processes of neuronal group selection that underlie learning on a minute by minute basis (Edelman 1989; Ellis and Toronchuk 2005). One should note that the concept is of much wider applicability, however: it can occur in a once-off selection events in physics cases such as state-vector preparation (Ellis 2011). Repetition increases its effectiveness but is not necessary to the concept.

The mathematics in general cases is the mathematics of adaptive selection (Holland, 1992), but in specific cases it results in the standard equations of population genetics (Gillespie 2004) and molecular evolution (Kuppers 1985). Computational models are given by theories of artificial neural nets (Bishop 1999) and genetic algorithms (Mitchell 1998).

TDC4: Top-down causation via adaptive information control

Adaptive information control takes place when there is *adaptive selection of goals in a feedback control system*, thus combining both feedback control and adaptive selection. The goals of the feedback control system are irreducible higher level variables determining the outcome, but are not fixed as in the case of non-adaptive feedback control; they can be adaptively changed in response to experience and information received. The overall process is guided by fitness criteria for selection of goals. This allows great flexibility of response to different environments, indeed in conjunction with memory it enables learning and anticipation and underlies effective purposeful action (Gray 2011), as it enables the organism to adapt its behaviour in response to the environment in the light of past experience, and hence to build up complex levels of behaviour.

Example: The classical example is ***Associative learning*** in animals, such as Pavlovian conditioning: an animal responds to a stimulus such as a sound, which is taken as a sign of something else and causes physical reactions implemented by motor neurons (Gray 2011). The training is causally effective by top-down action from the brain to cells in muscles. The fitness criterion is avoidance of negative stimuli; change of the associated goals (through a change in the environment) results in change of behaviour. More generally, the mind works by adaptive prediction of what is likely to happen, updated on an ongoing basis (Hawkins 2004). This underlies most of our mental ability. For example, the process of perception is a predictive adaptive process using Bayesian statistics to update the current perception on the basis of prediction errors (Purves 2010). This includes prediction of the intention of others, which is the basis of theories of other minds (Gray 2011).

It is claimed by some that the mathematics of evolutionary game theory (Alexander 2009) will act as an adequate basis for understanding these processes. Personally I have doubts about how far this can succeed, because of the simplistic nature of the reductive models of human behaviour implied.

TDC5: Intelligent top-down causation (i.e. the effect of the human mind on the physical world)

Intelligent top-down causation is the special case of feedback control with adaptive choice of goals where *the selection of goals involves the use of symbolic representation to investigate the outcome of goal choices*. Here a symbolic system is set of structured patterns realised in time or space, that is arbitrarily chosen by an individual or group to represent objects, states, and relationships (Deacon 1997). It will generally involve hierarchical structuring and recursion, as is required if it is to be useful in understanding complex situations, and has the potential to enable quantitative as well as qualitative investigation of outcomes.

***Example*: *Aircraft Design*:** Plans for a Jumbo Jet aircraft result in billions of atoms being deployed to create the aircraft in accordance with those plans**.** This is a non-trivial example: it costs a great deal of money to employ experts in aerodynamics, structures, materials, fuels, lubrication, controls, etc. to design and then to manufacture the aircraft in accordance with those plans. The plan itself is not equivalent to any single person's brain state: it is an abstract hierarchically structured equivalence class of representations (spoken, drawn, in computers, in brains, etc.) that together comprise the design. It is clearly causally effective (the aircraft would not exist without it).

Symbolic representation and choice of goals underlies the causal efficacy of abstract entities such as action plans and the value of money, represented symbolically. Thus the key feature of this higher level of causation, distinguishing it from the general case of adaptive control systems, is its use of *language (spoken or written) and abstract symbolism*, extending to *the quantitative and geometrical representations of mathematical models*. These are all irreducible higher level variables of an abstract nature: they form equivalence classes of representations, *inter alia* because they can be represented in different languages, and in spoken or written form or in computers. They enable information to be stored and retrieved, classified and selected as relevant or discarded, processed in the light of other information, and used to make qualitative and quantitative projections of outcomes and plan future actions in a rational way, altering goals according to an intelligent understanding of past experiences and future expectations. Intentional action then enables one to implement the resulting plans, and so change the physical world. The outcome is thus the result of human agency.

A key feature is the causal power of images and formal and informal causal models of the natural and social worlds, ranging from mental images of what might happen to elaborate quantitative models of physical entities and societies (Boulding 1969). These abstract entities (which are shared among many minds) play a large part in formulating our understandings and consequent actions, and hence are causally effective in the real world as they help us attain our goals. This is based in subjective personal experience. Language, science, mathematics and myriad other artifacts and customs are social constructions that strongly influence the behaviour of individuals and cultures; these are examples of higher-level causal variables that are not coarse grained lower level variables. Of course we do not fully understand how the mind is able to plan and make choices resulting in top-down action as discussed here; the fact that we do not know how it works does not affect the fact that we are certain it does happen.

Example: ***The value of money.*** Physically, fiat money is just coins or pieces of paper with patterned marks on them. This does not explain its causal significance. The effectiveness of money, which can cause physical change in the world such as the construction of buildings, roads, bridges, and so on by top-down action of the mind to material objects, is based in social agreements that lead to the value of money and exchange rates. These are abstract entities arising from social interaction over an extended period of time, and are neither the same as individual brain states, nor equivalent to an aggregate of current values of any lower level variables (although they are causally effective through such states and variables).

Example: Roles, expectations, and values. Roles are socially determined abstract entities that are causally effective in structuring society. They are a key aspect of the causal power of social structures (Vass, 2009). Roles are developed by an adaptive process which is a combination of bottom-up and top-down interaction between society and the individuals who make up the society. They are then inculcated into the individual by top-down social processes, whereafter they become a core feature of individual psychology in relation to society, together with expectations guiding the choice of goals and actions and hence being causally effective in a top-down way from the mind to the body (Gray 2011). Thus *our understandings of meaning and purpose* are abstract entities that form a high level in the hierarchy of causation in the mind. The imperative to search for meaning is a key aspect of human nature (Frankl 2006), without which the entire edifice of science would not exist.

Roles embody social values, which, together with individual values relating to life purpose, guide the individual and communal choice of goals and the methods used to attain these goals. Thus the highest level adaptive goals are *values related to ethics, aesthetics, and meaning*, which are all causally effective in a top-down way by determining the nature of desirable and/or acceptable lower level goals. They are a set of abstract principles that are causally effective in the real physical world, indeed they crucially determine what happens. For example wars will be waged or not depending on ethical stances; large-scale physical devastation of the earth will result if thermonuclear war takes place, so the nature of our values has crucial effects on the way human activity impacts on society and the world. Values are irreducible higher level entities: there is no way they can result from coarse-graining of lower level variables.

Are there useful mathematical models of human behaviour? Many have tried to make such models. In addition to the claims of games theorists and mathematical models of evolutionary processes (Nowak 2006), the main thrust has been that of mathematical economics and financial mathematics, involving foe example the Brownian Motion model of financial markets, rational pricing assumptions, and the Black–Scholes model (Black and Scholes, 1973; Merton, 1973). Any such models can however only be partially successful, giving correct predictions some of the time; their unthinking use to guide policy can lead to disaster, as happened in the global financial crisis of 2008-2009 (Patterson 2010). All such models should be treated with great caution.

Complex Adaptive Systems: The last three classes of top-down causation are all examples of *complex adaptive systems* (Gell-Mann 2002, Holland 1992). This is the only way that biological information can be generated and incorporated into living systems. Its importance is because it can let a system adapt to ongoing changes in the environment, indeed it is the only way of doing so (Harford 2011). It is also the key to the way life can apparently violate the second law of thermodynamics. Adaptive selection can accumulate structure and information by selecting a subset of entities from a set of many variants, selecting only those lower level states that correspond to a higher level selection principle, thus embodying a form of top-down action.

This is an analogue of Maxwell's demon: a micro-entity that chooses molecules with high energy from an ensemble of molecules in a container, and lets them pass a trapdoor into a compartment, thus heating up the gas in the compartment and so locally violating the second law of thermodynamics, as negligible energy is used in the selection (Von Baeyer 1998). The second law remains globally valid because of the entropy increase in the environment. Darwinian selection in effect envisages a macro demon who acts down to the molecular level to select a specific sequence encoding desirable genetic information from an ensemble of nucleic acids. This again enables a local violation of the second law. Note that this process can take place once-off: in biology it is repeated many thousands of times, but in physical systems it may occur only once, for example in the process of state vector preparation in quantum physics (Ellis 2011).

Bottom-up and Top-down causation: It is the combination of these different kinds of top-down action with bottom-up causation that enables true complexity to emerge, particularly because feedback loops are set up between lower and higher level variables. For example social constructions such as roles influence individual behaviour (Berger and Luckman 1967), but they are not fixed: they are also subject to evolution, through bottom-up action of individuals in society on longer timescales.

3. Aristotle's forms of causation.

Reductionist analysis `explains' the properties of the machine by analysing its behaviour in terms of the functioning of its component parts (the lower levels of structure). Systems thinking tries to understand the properties of the interconnected complex whole (Churchman 1968, Flood and Carson 1990), and `explains' the behaviour or properties of an entity by determining its role or function within the higher levels of structure. For example, the question: `Why is an aircraft flying?' can be answered,

- In *Bottom-up terms*: it flies because air molecules impinge against the wing with slower moving molecules below creating a higher pressure as against that due to faster moving molecules above, leading to a pressure difference described by Bernoulli's law, this counteracts gravity;

- In terms of *same-level explanation*: it flies because the pilot is flying it, after a major process of training and testing that developed the necessary skills, and she is doing so because the airline's timetable dictates that there will be a flight today at 16h35 from London to Berlin, as worked out by the airline executives on the basis of need and carrying capacity at this time of year;

- In terms of *top-down explanation*: it flies because it is designed to fly! This was done by a team of engineers working in a historical context of the development of metallurgy, combustion, lubrication, aeronautics, machine tools, computer aided design, etc., all needed to make this possible, and in an economic context of a society with a transportation need and complex industrial organisations able to mobilise all the necessary resources for design and manufacture. A brick does not fly because it was not designed to fly.

- And why was it designed to fly? Because it will potentially make a profit for the manufacturers and the airline company! Without the prospect of that profit, it would not exist. This is the topmost cause for its existence.

These are all simultaneously true non-trivial explanations; *the plane would not be flying if they were not all true at the same time.* The higher-level explanations involving goal choices rely on the existence of the lower level explanations involving physical mechanisms in order that they can succeed, but are clearly of a quite different nature than the lower level ones, and are certainly not reducible to them nor dependent on their specific nature. The higher level goals can be realised in multiple ways. The bottom-up kind of explanation would not apply to a specific context if the higher-level explanations, the result of human intentions, had not created a situation that made it relevant.

This situation was captured by Aristotle in his *Metaphysics* (Ross, 1953) through his proposal of four different kinds of causation. According to Falcon (2006), they are**:**

- ***The material cause***: "that out of which", e.g., the bronze of a statue
- ***The formal cause***: "the form", "the account of what-it-is-to-be", e.g., the shape of a statue.
- ***The efficient cause***: "the primary source of the change or rest", e.g., the artisan, the art of bronze-casting the statue, the man who gives advice.
- ***The final cause***: "the end, that for the sake of which a thing is done", e.g., health is the end of walking, losing weight, purging, drugs, and surgical tools

The last is a *teleological explanation* - an explanation that makes a reference to *telos* or purpose. Additionally, circular causation is possible: things can be causes of one another - a relation of reciprocal influence.

I suggest a modern adaptation of these four kinds of causes is to consider causation in the hierarchical context considered here, identifying as specially significant the immediate lower level (Physical) cause, the same level (Immediate) cause, the immediate higher (Contextual) cause, and the topmost level of purpose or *Telos*, which activates the rest. We cannot identify an ultimate lower level cause because no one knows what the bottom level is (we have no fully successful ultimate theory of particle physics)

***Example*: *Physics Experiments*.** Successful completion of a physics experiment, such as observing particle production in the LHC collider, involves all these forms of causation. The physical cause is the particle interactions that lead to the production of new particles. The immediate cause is that the experimenters turn the accelerator and measuring equipment on at a particular time. The contextual cause is that the collider was designed and manufactured so that the collisions would take place and outcomes could be observed. The purpose might simply be that the experimenters want to understand the collision in the context of a theory of AdS/CFT duality, or it might be because they aspire to attaining a Nobel Prize.

Thus one can have top-down system explanations as well as bottom-up and same level explanations*,* all being simultaneously applicable. Indeed there will be numerous causal factors in any specific case forming a network of causes, including those identified here but also the overall historical and physical and social environment, without which the identified events would not take place (for example, the laws of physics are as they are, the Earth exists, scientists are able to do experiments, measuring apparatus can be devised reliably, and so on). An explanation usually takes most of this for granted and focuses just on one or two items that are the subject of attention, perhaps because they can be manipulated to alter the result.

The key point about causality in real-world contexts, then, is that ***simultaneous multiple causality (inter-level, as well as within each level) is always in operation in complex systems***. Claiming that any specific single cause is the only causation in action is fundamentally misleading, as it ignores the complex nature of the real causal web (Ellis 2010).

4. The Causal Efficacy of Human Symbol Systems

Some aspects of complex systems are emergent from their own internal nature and logic, but others rather arise from the way the nature of the external environment shapes human symbolic systems. A crucial point then is that the resulting higher level causal entities are not of a physical nature and are not reducible to or emergent from any physical entities. They do however have causal power. They are derived from the behaviour of the world around us and mirror the way that world behaves, hence they arise from the reality of the external world, which exists independent of the mind. They are thus not coarse grained lower-level variables. They are discovered and comprehended by the mind, so the top-down influence of operations at the level of symbolic systems is via intelligent top-down causation (TDC5).

Example*: *Mathematics Comprehension and Utilisation is a case of top-down causation from a world of mathematical abstractions to the human mind, being realized in details of neuronal connections, and then into the real world where it is causally effective both in terms of creating patterns on paper, and by underlying engineering and planning. Major parts of mathematics are discovered rather than invented (rational numbers, zero, irrational numbers, and the Mandelbrot set being classic examples). They are not determined by physical experiment, but are rather arrived at by mathematical investigation. They have an abstract rather than embodied character; the same abstract quantity can be represented and embodied in many symbolic and physical ways, and these representations form an equivalence class. The underlying mathematical truths are independent of the existence and culture of human beings (Penrose 1997; Changeux and Connes 1998);[1] it is plausible that the same features will be discovered by intelligent beings in the Andromeda galaxy as here, once their mathematical understanding is advanced enough (which is why these features are advocated as the basis for inter-stellar communication).

These features are discovered by humans, and represented by our mathematical theories; that representation is a cultural construct, but the underlying mathematical features they represent are not - indeed like physical laws, they are often unwillingly discovered, for example the irrationality of $\sqrt{2}$ and the number π. These mathematical verities are causally efficacious through the actions of the human mind: one can for example print graphic versions of the Mandelbrot Set in a book, resulting in a physical embodiment in the ink printed on the page.

***Example*: *Physics Theories*:** Maxwell's *theory* of electromagnetism (an abstract entity, described by Maxwell's equations, see Fleisch 2008) led to the development of radio, and then to existence of cell phones, TV, and so on, based in manipulation of physical materials composed of atoms and electrons. Maxwell's theory is not a physical entity, nor is it the same as any single person's brain state. It can be represented in many ways (on blackboards, in print, on computer screens, in spoken words) and in many formalisms (via 3-dimensional vectors or 4-dimensional tensors, for example). These various representations together form an equivalence class, as they all lead to the same predicted outcomes. How do you demonstrate top-down causation by this theory? - design an artefact such as a cell-phone through use of Maxwell's theory, and then construct it and operate it. The abstract theory will have altered physical configurations in the real world, and hence is causally effective. The theory is an irreducible higher level entity (it cannot be derived by coarse-graining any lower level variables) representing the nature of physical reality, in that is a representation of physical laws of behaviour that are eternal and omnipresent (physics is the same everywhere in the universe).

It is the accuracy of this representation of the way the world works that gives the theory its causal powers: it is demonstrably a good representation of the underlying physical reality (namely, the consistent regularities in the behaviour of matter that underlies what happens in the physical universe). Hence in this way the causal regularities in the physical world can be represented as a set of abstract patterns, resulting in a mental theory which is causally efficacious. *It is the underlying regularities in the behaviour of physical matter, independent of any human comprehension or mind, that is the ultimate source of this causal efficacy,* and hence *in this way* (enabled by TDC5) has causal powers in the physical world, for example by underlying engineering practice (as well by governing the physical behaviour of matter).

[1] Penrose and Connes propose that this makes best sense if one considers mathematics as existing in a Platonic sense (Penrose 1997, Changeux and Connes 1998). I concur (Ellis 2004), but this is of course a philosophically contentious position.

5. Room at the Bottom

Where does the causal slack lie enabling top-down action to take place? If the underlying physics is deterministic and determines all physics outcomes at the lower level, how is there freedom for higher level causation to be efficacious? Three key features are relevant.

Firstly, in considering specific physical and biological systems, it lies partly in the *structuring of the system* so as to attain higher level functions, for example the specific connections in a computer (which could have been different) act as constraints on lower level dynamics, thus channelling how they function; and partly in the *boundary conditions* together with *openness of the system*: new information can enter across the boundary and affect local outcomes. Together these features set the environment in which the lower level components operate, and so determine their outcomes (see for example Denis Noble's article).

Secondly, *top-down causation can change the nature of the lower elements*. There is not just a situation of invariant lower level elements obeying fixed physical laws; rather we have the nature of lower level elements being changed by context. Often this ensures that the lower level elements function so as to fulfil higher level purposes: this is an aspect of adaptive selection. Thus *the nature of micro-causation is changed by top-down processes*, profoundly altering the mechanistic view of how things work.

Example: Cell differentiation. Through the processes of developmental biology, cells get differentiated to perform specific functions; this changes their nature relative to other cells in an adaptive way (Gilbert 2006). Cells differentiate into neurons that get adapted to their location in the brain, into muscle cells adapted to their role in the heart, and so on. They each develop so as to fit into their allotted role in the body, creating the body and its biological form as they do so, and are then fine-tuned for their function. A particular case is the development of sensory neurons (Grey 2011) out of pluripotent cells.

Example: Humans in society. Individual minds develop in the context of their interactions with other minds, and brain development cannot be understood outside this context (Donald 2001). Individuals are shaped by society so that they fit into that society, for example learning a specific language and a variety of societal roles and expectations (Berger and Luckmann 1967). This is top-down causation from the society to the individual, and indeed to their synaptic connections: their brain is adapted to fit into the society in which it lives (Ambady 2011).

Thirdly, the required freedom lies in *micro-indeterminism* (random outcomes of microphysical effects), *combined with adaptive selection*: random outcomes at the micro-level allow variation at the macro-level, which then leads to selection at the micro-level but based in macro-level properties and meaning. Statistical variation and quantum indeterminacy provides a repertoire of variant systems that are then subject to processes of adaptive selection, based on higher level qualities of the overall system.

For this to work, one needs amplifying mechanisms in order to attain macroscopic variation from microscopic fluctuations. Some physical systems (such as photomultipliers and the human eye) amplify quantum effects to a macroscopic scale; classically, chaotic systems can amplify micro-fluctuations in initial data; some of the effects captured in Thom's catastrophe theory allow large amplification of microscopic changes; and some molecular biology processes (for example involving replication of mutated molecules) act as such amplifiers, even allowing quantum effects to change evolutionary outcomes (Percival 199). Because of quantum uncertainty, at a profound level the universe is indeterministic, allowing the needed causal slack for higher levels to be free of the tyranny of absolute control by lower level dynamics; and this affects biological processes (Lloyd 2011, Ball 2011). By itself that does not lead to emergence of higher level order; but it does allow this to occur through the process of adaptive selection. That adaptive selection process will act on equivalence classes of lower level, guided by higher level selection principles (Auletta Ellis and Jaeger 2007).

Because of the existence of random processes at the bottom, there is sufficient causal slack to allow all these kinds of causation to occur without violation of physical causation, for example developmental biology amplifies molecular level variation to system level changes.. That these random processes do indeed occur at the lower levels is indicated by many kinds of evidence (Glimcher 2005; Eldar and Elowitz, 2011; Chouard 2011). This mechanism can only work because of the large gap between macro and micro physics, together with the huge number of micro-components involved (atoms in a cell, cells in a human body, etc.): hence emergence of genuine complexity requires the vast numbers of entities entailed in physical reality.

Criticism and Response
In response to the argument above about the causes leading to an aircraft flying, Tim O'Connor has commented as follows: Note that the higher-level explanations appeal to (intentional) states long *prior* to the plane's flying. That is, the explanation works by setting the event to be explained in a larger spatiotemporal context. The reductionist might retort: If those prior intentional states are themselves wholly fixed by more fundamental physical facts that compose them, we could have in principle a completely physical (bottom-up) explanation spanning each step of the larger context to which you point. This would be an explanation wholly independent of high-level intentional explanations -- an appeal to facts that are themselves collectively responsible for their being a co-existing intentional level of explanation. Taking the widest scope possible (the universe as a whole), the fundamental physical facts and the laws that directly govern them asymmetrically determine the existence of higher-level systems and the forms of explanation they make possible. Or, at any rate, it is not clear that anything to which this paper appeals conflicts with this assertion. And if that is correct, then why is there not a perfectly good sense in which all action takes place down below?

As I understand it, this proposes that the larger spatiotemporal context of cosmology sets initial data that completely determines the present day situation and so explains all current lower and higher levels. My answer is two-fold.

Firstly, because of quantum uncertainty, such a proposal to explain present day details in terms of cosmological initial data cannot work even in principle: For example, quantum fluctuations can change the genetic inheritance of animals (Percival 1991) and so influence the course of evolutionary history on Earth. Indeed that is what occurred when cosmic rays – whose emission processes are subject to quantum uncertainty - caused genetic damage in the distant past (Scalo et al 2001). Consequently *the specific evolutionary outcomes on life on Earth (the existence of dinosaurs, giraffes, humans) cannot even in principle be uniquely determined by causal evolution from conditions in the early universe, or from detailed data at the start of life on Earth*. Quantum uncertainty prevents this, because it significantly affected the occurrence of radiation-induced mutations in this evolutionary history. The specific outcome that actually occurred was determined as it happened, when quantum emission of the relevant photons took place: the prior uncertainty in their trajectories was resolved by the historical occurrence of the emission event, resulting in a specific photon emission time and trajectory that was not determined beforehand, with consequent damage to a specific gene in a particular cell at a particular time and place that cannot be predicted even in principle. If our own existence cannot uniquely follow from that initial data, neither can any specific thoughts or intentions.

Secondly, if we disregard this impossibility, we are in effect faced with the proposal that the future occurrence of the battle of Trafalgar, the painting of the Mona Lisa, and the discovery of General Relativity theory by Albert Einstein are specifically written into the fluctuations on the last scattering surface in the early universe that we now observe through the WMAP satellite. I believe this is patently absurd. The only way these outcomes could have happened is for genuine higher level causal powers to have come into being with their own inherent logic, these then leading to these extraordinary outcomes (*inter alia* causing electrons and protons to move in brain in ways essentially determined by higher level causal factors). There is no way that they could be implied by physics per se.

Finally, O'Connor suggests that although it has to be conceded that biological top-down causation in cell differentiation does show that some *non-fundamental* levels of causes and explanations are not independent of those above them, the reductionist will claim that whatever the fundamental physical facts and laws turn out to be *will* be independent of higher-level entities that they make possible. However according to the best current "theory of everything", namely M-theory, the particles and forces that exist are not uniquely determined by fundamental laws but rather result from the specific string vacuum state that occurs (Douglas 2003). In that case a purely bottom-up explanation cannot work: the very nature of the effective laws of physics is environmentally dependent. It has been claimed by Susskind (2005) that the existence of intelligent life in the context of a multiverse thereby gives an explanation of why the laws of physics we experience are as they are – it is a selection effect resulting from our existence as observers in a multiverse. This is indeed a radical revision of the old idea of immutable physical laws underlying all physical causation.

A paper by Purves, Wojtach, and Lotto (2011), and comments by Simon Saunders (personal communication), express the view that in the end, one must acknowledge that it is still a fact that the essential work is happening at the micro level, whatever contextual effects may occur. Yes indeed; but by itself that action does not determine the outcomes, even when put in the full historical context, as just explained. I believe the above response to O'Connor, together with the other arguments in this paper, suffices to respond to these comments: for they show that while the lower levels do the work, the higher levels decide what is to be done. And it's just as well we don't have to understand the lowest level in order to make higher level predictions, because we don't even know what the lowest level is..

6: Implications

Understanding top-down action can have important effects on society. Two cases will be mentioned here: medical/health care issues, and education, in particular as regards language issues. In both cases an ongoing battle between bottom-up and top-down approaches has important consequences for welfare in society.

Health: the welfare of the body. There are in essence competing bottom-up (reductionist) and top-down (holistic) views of how to look after illness and disease, resulting in different treatment modalities. Bottom-up methods emphasize the role of microsystems in determining health outcomes, whereas the top-down methods place the emphasis on the ways that the operations of the mind can affect bodily health.

The bottom-up emphasis leads to treatment regimes that place their faith in medicines, drugs, and surgery; they emphasize the view that the way to heal the whole is to treat the parts. Top-down methods acknowledge that state of mind plays a significant role in physical health, and so takes that into account too. There are many ways that such top-down influences take place, an important one being the fact that many immune molecules are also neuro-modulators, allowing the mind to influence physical health through interaction of the brain and the immune system (Sternberg 2000). This is part of the way the mind influences health (Moyers 1995), indeed overall mental state crucially affects physical health. Some examples: small infants depend upon a caregiving adult for survival; they may suffer developmental problems, an array of emotional problems, and perhaps failure to thrive if deprived of such care (Berry 2001). In extreme cases they may lose weight and eventually die if they are not given individual attention and caring, even if all their bodily needs are met. Treatment paradigms should take this into account.[2] Children's survival rates after heart surgery are affected by whether or not they have animal companions (Friedmann 1980). In a family context, survival rates depend on family communication patterns (Gerhke and Kirschenbaum 1967).

As regards specifically mental health, there are competing paradigms of how to treat mental problems, represented by neurology, psychiatry, clinical psychology, psychotherapy, and cognitive behaviour therapy, the first being essentially bottom-up and the latter three top-down, with psychiatry in the middle. There may be different cases where one or other of these approaches is more appropriate. The challenge is to determine what combination of bottom-up and top-down interventions will give the best results in specific cases (Holmes 2001).

Education: learning to read and write. Closely related issues arise in education, with literacy teaching and learning providing an apt example. A bottom-up (part to whole) approach tackles detailed technical aspects of language such as phonics and handwriting first, and worries about the functional, communicative roles later, with

[2] In the 1970s, Dr Trudi Thomas ran a scheme called Sunshine Wards as St Matthew's Hospital, Keiskammahoek, to provide emotional contact to malnourished orphans in the Ciskei; this resulted in improved survival rates. This is described in her book *Their Doctor Speaks* (privately published: 1973). A current version of therapy for newborn children based on the same principles is called *Kangaroo care*; see http://en.wikipedia.org/wiki/Kangaroo_care for articles giving evidence of its medical benefits.

the assumption that the parts will then come together in some kind of 'building block' way, to form a meaningful whole. In this case, if pursued to an extreme, children are first given decontextualised exercises which ask them to recognise and sound out numerous letter –sound combinations such as ma, me, mi ma, mu, followed by phonetically regular words and even nonsense words such as tok, zat, fot . Only once children have accomplished learning this and other restricted texts, are they given opportunities to move on to engaging with meaningful texts. The mechanics of language is thus taught separately from and prior to the essential purpose of language, which is to make and convey meaning. Crucially, testing of children is done that involves recognition of meaningless words and hence omits the core functioning of language. Dehaene makes this explicit in his book (Dehaane2010, page 200):

" ***the child's brain, at this stage, is attempting to match the general shape of the words directly onto meaning, without paying attention to individual letters and their pronunciation – a sham form of reading***"

He defines reading the wrong way round, thereby misinterpreting the learning process! He wants the parts to work rather than the whole, and characterizes as sham reading, what is in fact both the aim of fluent readers, and the way that young children learn language. What he deplores is precisely what we want children to learn: to read from the whole, not the parts. The testing involved in associated reading programmes is of the same nature as this quote indicates: it penalizes attempts at conventional reading practices.

By contrast, one can proceed top-down, starting with the aim of conveying the idea that symbols convey meaning, showing children how this works in practice by immersing them in a reading culture, while at the same time encouraging them into a process of experimentation and successive approximation, as they come to adjust initial formally conventional symbol usage to more conventional symbolic use over time. In this way, ability to read and write emerges in a similar way to how babies learn oral language, by a trial and error process with feedback, but always concentrating primarily on the way that language is essentially about conveying meaning (Holdaway 1979, Krashen and Terrell 1983, Hall 1989). A beautiful illustration of how this occurs for young children's writing is given in Bloch (1997). In this case holistic understanding and meaning are the primary educational themes, and detailed issues such as grammar and syntax (and the other component parts) are tackled within the context of the whole. This understanding of reading as a psycholinguistic process leads to the *whole language* approach to teaching reading and writing.

This is a crucial ongoing debate with major practical implications for education. The bottom-up approach centred on grasping phonemes first is strongly supported inter alia by Adams (1990), Dehaene (2010). The holistic view (Goodman et al, 1986) strongly supports the top-down approach: children are inherently motivated to make sense of the complexities of their world, learning in a seemingly messy way that draws on all their knowledge and strengths to integrate bottom-up and top-down understandings predictively in order to do justice to this complexity. This is what happens in apprenticeships. It is also what happens for many children who grow up in well resourced, literate homes. Because the conditions of learning described by Cambourne (1995) are appropriate, and they are immersed in story reading and play with written language, such children learn, as if by osmosis many of the essential 'concepts of print' that are neglected in the first years of primary school. Children from home backgrounds which do not provide such experiences, such as in many African settings (Bloch2005), find themselves flailing when their introduction to print concentrates on the technicalities alone. Many tend to lose interest in what they see as meaningless activities, and this results in a lesser ability to read in a successful way.

This ties in to a much larger picture of how the brain functions in a top-down way. The brain is exquisitely constructed to search for meaning (Donald 2011) and to predict what is likely to happen (Hawkins 2004). This happens particularly in vision: it is not true that vision can be understood simply as data coming in from our eyes and being interpreted by the brain. Rather the brain is continually predicting what ought to be there, and filling in what it expects to see on the basis of only some of the data that it actually analyses at any one time. This can conclusively be show to be the case by analysing visual illusions (Frith 2007, Purves 2010). Hence a top-down process of interpretation, based on our expectations and facilitated by specific neuronal connections, modulates and shapes what we actually see. Similar processes happen in listening to music: expectation is a key feature of how we experience music (Levitin 2007, Huron, D 2007).

This is a crucial aspect of the way we read a text, which does not take place by reading each phoneme, assembling them into words, assembling those into phrases, and so on. Rather the eye skips over words, reading whole phrases at a time and filling in the bits that are not actually read. This can be demonstrated by miscue analysis and eye movement research (Flurkey, Paulsen and Goodman: 2008; Ebe 2008; Strauss, Goodman and Paulson 2009). You can experience it for yourself by carrying out the following simple exercise: read the following statement once or twice, and as you do so count the number of times the letter "F" appears:

FINAL FOLIOS SEEM TO RESULT FROM YEARS OF DUTIFUL STUDY OF TEXTS ALONG WITH YEARS OF SCIENTIFIC EXPERIENCE

How many times does F occur in this statement – 5? 6? 7? The answer is given at the end of the conclusion.

Our ability to read ambiguous texts derives from the fact that context sets the meaning and even the pronunciation of words: language is driven by word associations rather than individual words (Hooey 2001). Consequently, context drives the process of reading: it is not bottom-up, it is a psycholinguistic guessing game (Goodman 1967). This top-down driven process is a fundamental aspect of how the brain works, and is at the core of what reading is about. Because brain imaging studies have not been conducted on this holistic reading process, they do not yet have the capacity to show whether or not the phonics approach to teaching reading is superior (Strauss, Goodman and Paulson 2009).[3]

Other Applications: Obviously these contested themes have not been dealt with in adequate depth here. I have included the two examples in this section in order to indicate there is evidence demonstrating that the themes of this article might have significant practical consequences. This suggest this possibility must be taken seriously also in regards to other aspects of education, sometimes determining educational success, and other aspects of welfare, which is not determined in a simple bottom-up way.

7: Conclusion

Interaction of top-down and bottom- up causation as an integrating theme: The basic them of this paper is that it is the interaction of bottom-up and top-down effects that enables the emergence of true complexity. This realisation gives an integrative view of the relation between different levels of the hierarchy of complexity, both within each subject (particle physics to nuclear physics, nuclear physics to atomic physics, and so on; also between each of the levels of biology) and between subjects (physics to chemistry, chemistry to biology, psychology to sociology, and so on).

It is clear that top-down causation as envisaged here happens at the levels of physiology (Noble 2008) and in the in the case of vision (Purves 2010, Bressler et al 2011, Hansen et al 2011). Indeed there is a great deal of data supporting this view as regards the way the mind works in general, including the rapidly growing literature on cultural neuroscience (Harmon-Jones and Winkielmen 2007; Ambady 2011; Gavel 2011; Lederbogen et al 2011). This therefore supports the case made by Elder-Vass that it happens in society (Elder-Vass 2010). It can be demonstrated to happen in other cases, for example in microbiology (Jaeger, this volume) and fluid flows (Bishop, this volume). It also plays a significant role in quantum physics, see Ellis (2011).

In brief: the conclusion is that *there are other forms of causation than those encompassed by bottom-up causation due to physics and physical chemistry processes considered on their own*. The higher up the ladder of complexity one goes, the clearer this becomes; in particular, it is undeniable as regards the way the human mind operates, and

[3] Some books (for example Dehaene 2010) claim the opposite on the basis of neuroscience evidence, including brain imaging data. However these data do not in fact prove what they claim. The experiments that are claimed to lead to this conclusion do not test what is going on in the brain when meaningful reading is taking place; they relate only to reading meaningless texts or to some components of the reading process. This is not adequate data from which one can make pronouncements about the genuine reading process as a whole, where meaningful texts are read and understood.

in physiology. It is particularly clear in those cases where the higher level variables cannot be obtained by coarse-graining of any lower level variables, as is the case both in the causal power of money, and the causal effectiveness of Maxwell's theory of electromagnetism in the world of engineering.

A full scientific view of the world must recognise this fact, or else it will ignore important aspects of causation in the real world, and so will give a causally incomplete view of things. These forms of causation are based in the interaction of bottom-up and top-down effects: if we neglect either, we will be unable to understand genuinely complex systems.

Answer: The correct answer to the question in the last section (how many times does the letter F occur in the statement there?) is 8. If you got this stunningly simple algorithmic task wrong (as most people do), it is because we do not read a sentence word by word, phoneme by phoneme: the mind skips words and interpolates when reading! You don't even see the words "of" in this text, because the mind takes them for granted and does not read them.[4] This is an experimental proof of the fact that we read in a top-down way.

***

I thank Tim O'Connor and three referees for helpful comments that have considerably improved this paper, and Carole Bloch for crucial input as to how the reading process actually takes place.

***

[4] I am indebted to Kevin Dutton's intriguing book *Flipnosis* (Dutton 2010) for this example.

REFERENCES

Adams, M (1990) *Beginning to Read: Thinking and Learning about Print* (Cambridge MA: MIT Press)

Alexander, J M (2009) "Evolutionary Game Theory", *The Stanford Encyclopedia of Philosophy (Fall 2009 Edition)*, Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/fall2009/entries/game-evolutionary/.

Alon, U (2007) *An introduction to systems biology: design principles of biological circuits* (London: Chapman and Hall /CRC).

Ambady, N (2011) "The Mind in the World: Culture and the Brain" *Association for Psychological Science* **24**: 5-6, 49.

Auletta, G, Ellis, G F R, and Jaeger, L (2007). "Top-Down Causation: From a Philosophical Problem to a Scientific Research Program". *Journal of the Royal Society Interface* (2008) [arXiv:0710.4235].

Ball, P (2011) "The dawn of quantum biology" *Nature* **474**: 272-274.

Barabási, A-L, and Oltvai, Z N (2004). "Network biology: Understanding the cell's Functional organization". *Nature Reviews: Genetics* **5**: 101-114**.**

Berger, P and Luckmann T (1967). *The Social Construction of Reality: A Treatise in the Sociology of Knowledge* (Anchor, New York).

Berry, B (2001) "Bonding and Attachment in Maltreated Children Consequences of Emotional Neglect in Childhood" The ChildTrauma Academy: http://centerforchildwelfare.fmhi.usf.edu/kb/ChronicNeglect/ConseqOfChNeglect.pdf

Black, F, and Scholes, M (1973). "The Pricing of Options and Corporate Liabilities". *Journal of Political Economy* **81** (3): 637–654.

Bishop, C M (1999). *Neural Networks for Pattern Recognition* (Oxford: Oxford University Press).

Bloch, C (1997). *Chloe's Story* (Cape Town: Juta). Available at http://www.mth.uct.ac.za/~ellis/Chloe_story.pdf.

Bloch, C (2005) *Enabling Effective Literacy Learning in Multilingual South African Early Childhood Classrooms*. Cape Town: PRAESA University of Cape Town.

Booch, G (1994). *Object Oriented Analysis and design with Applications* (New York: Addison Wesley).

Boulding, K E (1969) *The Image: Knowledge in Life and Society* (Ann Arbor: University of Michigan Press).

Brenner, S C, and Scott, L R (2007) *The mathematical theory of finite element methods* (Heidelberg: Springer).

Bressler, S L, Tang, W, Sylvester, C M, Shulman, G L, and Corbett, M (2008) "Top-down control of human visual cortex by frontal and parietal cortex in anticipatory visual spatial attention" *The Journal of Neuroscience*: **28**: 10056-10061.

Busemeyer, J R (2011) "Dynamic systems" To appear in *ENCYCLOPEDIA OF COGNITIVE SCIENCE*, Macmillan [http://www.cogs.indiana.edu/Publications/techreps2000/241/241.html].

Cambourne, B (1995) Toward an educationally relevant theory of literacy learning: Twenty years of inquiry.The Reading Teacher Vol.49, No. 3 November 1995

Campbell, N A and Reece, J B (2005). *Biology* (San Francisco: Benjamin Cummings).

Changeux, J-P and Connes, A (1998). *Conversations on Mind, Matter, and Mathematics* (Princeton University Press).

Chouard, T (2011) "Breaking the Protein rules: If dogma dictates that proteins need a structure to function, then why do so many of them live in a state of disorder?" *Nature* **471**: 151

Conway Morris, S (2005). *Life's solution: inevitable humans in a lonely universe*. Cambridge, UK: Cambridge University Press.

Courant, R. & Hilbert, D. (1962), *Methods of Mathematical Physics*, *II* (New York: Wiley-Interscience)

Deacon, T (1997). *The symbolic species: the co--evolution of language and the human brain* (London: Penguin).

Dehaene, S (2010) *Reading in the Brain*: *The New science of how we read* **(**London: Penguin)

Deuschle, M, and Meyer-Lindenberg, A (2011) "City living and urban upbringing affect neural social stress processing in humans" Nature 474: 498-501.

DiStefano, J J, Stuberub, A R, and Williams, I J (1995) *Feedback and control systems* (McGraw Hill: Schaum Outlines).

Donald, M (2001) *A Mind so Rare: The Evolution of Human Consciousness* (W W Norton).

Douglas, M (2003) "The statistics of string / M theory vacua", *JHEP***0305**, 46 [arXiv:hep-th/0303194]

Dutton, K (2010) *Flipnosis: The art of split-second persuasion* (London: Random House).

Ebe, A (2008) “What eye movement and miscue analysis reveals about the reading process of young bilinguals” In *Scientific Realism in Studies of Reading*, Ed. A D Flurkey, E J Paulson, and K S Goodman (Taylor and Francis: London), 131-152.
Edelman, G M. (1989). *Neural Darwinism: The theory of group neuronal selection*. (Oxford: Oxford University Press).
Eldar, A and Elowitz, M B (2010) “Functional role for noise in genetic circuits” Nature 467: 167- 173.
Elder-Vass, D (2010) *The Causal Power of Social Structures: Emergence, Structure and Agency* (Cambridge: Cambridge University Press).
Ellis, G F R (2004): “True Complexity and its Associated Ontology”. In *Science and Ultimate Reality: Quantum Theory, Cosmology and Complexity*. Ed J D. Barrow, P C.W. Davies, and C L Harper (Cambridge: Cambridge University Press), :607-636
Ellis, G F R (2005). “Physics, Complexity, and Causality” *Nature* **435:** 743.
Ellis, G F R (2008). “On the nature of causation in complex systems” *Transactions of the Royal Society of South Africa* **63**: 69-84 [http://www.mth.uct.ac.za/~ellis/Top-down%20Ellis.pdf]
Ellis, G F R (2010) ``Fundamentalism in Science, Theology, and the Academy''. In *Human Identity at the Intersection of Science, Technology, and Religion* (Editors: Nancey Murphy and Christopher C. Knight (Ashgate), 57-76.
Ellis, G F R (2011) ``On the limits of quantum theory: contextuality and the quantum-classical cut''. http://arxiv.org/PS_cache/arxiv/pdf/1108/1108.5261v2.pdf.
Ellis, G F R and Toronchuk, J A (2005) “Neural Development: Affective and Immune System Influences.” In *Consciousness and Emotion*: *Agency, Conscious Choice, and Selective Perception* Edited by R D Ellis and N Newton (John Benjamins, 2005), pp. 81–119.
Falcon, A (2006), “Aristotle on Causality”. Stanford Encyclopedia of Philosophy, http://plato.stanford.edu/entries/aristotle-causality/.
Fleisch, D (2008) *A Student's guide to Maxwell's equations* (Cambridge: Cambridge University Press).
Flurkey, A D, Paulsen, E J, and Goodman, K S (2008) *Scientific Realism in Studies of Reading* (New York: Laurence Erlbaum).
Frankl, V (2006) *Man's Search for Meaning* (Beacon Press),
Friedmann, E, Katcher, A H, Lynch, J L, and Thomas, S A (1980) *Public Health Reports* **95**:307-312.
Frith, C (2007) *Making up the Mind: How the brain creates our mental world* (Malden: Blackwell).
Gerhke, S, and Kirschenbaum, M (1967) “Syrvival patterns in family conjoint therapy” *Fam Proc* **6**:67-80.
Gellman, M (2002). *The Quark and the Jaguar* (Abacus).
Gilbert, S F (2006): *Developmental Biology* (Sunderland, MA: Sinauer Associates).
Gillespie, J H (2004) *Population genetics: a concise guide* (Baltimore: Johns Hopkins University Press).
Glimcher, P W (2005). “Indeterminacy in brain and behaviour”. *Annual Review of Psychology*, **56**: 25.
Goodman, K S (1967) “Reading: a psycholinguistic guessing game”. In *Language and Literacy: The Selected Writiings of Kenneth Goodman*, *Volume 1* (Ed F V Gollasch) (Routledge and Kegan Paul: London) , 33-44.
Goodman, K S, Smith E B, Meredith R, Goodman Y M (1986). *Language and Thinking in School: A Whole-Language Curriculum* (Richard C Owen Publishers).
Gray, P (2011) *Psychology* (New York: Worth).
Hall, N (1989) *Writing with reason* (London:Hodder and Stoughton).
Hanson, J L, Chandra, A, Wolfe, B L and Pollak, S D (2011) “Association between Income and the Hippocampus” *PLoS One*: http://www.plosoone.org/article/info:doi/10.1371/jpurnal.pone.0018712.
Harford, T (2011) *Adapt: Why success always starts with failure* (London: Little Brown).
Harmon-Jones, E, and Winkielmen, P (2007) *Social Neuroscience: Integrating Biological and psychological Explanations of Human Behaviour* (New York: Guilford Press).
Hawkins, J (2004). *On Intelligence* (New York: Holt Paperbacks)
Holdaway, D. 1979. *The Foundations of Literacy* (Sydney:Ashton Scholastic),
Holmes, J (2001) “All you need is cognitive therapy?” *British Medical Journal* **324**: 288.
Hooey, M (2005) *Lexical Priming: A New theory of words and language* (London: Routledge).
Holland, J H (1992). *Adaptation in Natural and Artificial Systems* (Cambridge, Mass: MIT Press).
Huang, K (1990). *Statistical Mechanics*. (Wiley, John & Sons).
Huron, D (2007) *Sweet anticipation: Music and the psychology of expectation* (Cambridge, Mass: MIT Press).
Iglesias, P A and Ingalis, B P (2010) *Control Theory and Systems Biology* (Cambridge, Mass: MIT Press).
Jaeger, R (1997). *Microelectronic Circuit Design* (McGraw-Hill).

Knuth, D E (2010) *Selected Papers on Design of Algorithms* (Stanford, California: Center for the Study of Language and Information)

Krashen, S D and Terrell, TD (1983): *The natural approach: Language acquisition in the classroom* (San Francisco: Alemany Press).

Kuppers, B-O (1994). *Information and the origin of life* (Cambridge, Mass: MIT Press).

Lederbogen, F, Kirsch, P. Hadddad, L, Strie, F, Tost, H, Schuh, P, Wust, S, Preussner, J c, Riestchel, M,

Lloyd, S (2011) "A bit of quantum hanky-panky" *Physics World* January 2011: 26-29.

Levitin, D J (2007) *This is your brain on music: The science of a human obsession* (London: Plume).

Merton, RC (1973). "Theory of Rational Option Pricing". *Bell Journal of Economics and Management Science* **4**: 141–183.

Mitchell, M (1998) *An Introduction to Genetic Algorithms* (*Complex Adaptive Systems*)**.** (Cambridge, Mass: MIT Press).

Moyers, W (1995). *Healing and the Mind* (Doubleday: Main Street Books).

Noble, D (2008) *The Music of Life: Biology Beyond Genes* (Oxford: Oxford University Press).

Nowak, M A (2006). *Evolutionary Dynamics: Exploring the Equations of Life (*Harvard University Press).

Patterson, S (2010) *The Quants: How a new breed of Math Whizzes conquered Wall Street and nearly destroyed it* (New York: Crown Business).

Penrose, R (1989). *The Emperor's New Mind: Concerning Computers, Minds and the Laws of Physics* (New York: Oxford University Press).

Penrose, R (1997). *The Large, the Small and the Human Mind* (Cambridge: Cambridge University Press).

Penrose, R (2005). *The Road to Reality: A Complete Guide to the Laws of the Universe* (Knopf).

Percival, I (1991) "Schrödinger's quantum cat". *Nature* **351**, 357.

Purves, D (2010) *Brains: How they seem to work* (Upper Saddle River: FT Press Science).

Purves, D, Wojtach, W T., and Lotto, R B (2011) "Understanding vision in wholly empirical terms". *PNAS Early Edition* www.pnas.org/cgi/doi/10.1073/pnas.1012178108

Rhoades, R and Pflanzer, R (1989). *Human Physiology* (Fort Worth: Saunders College Publishing)

Roederer, J G (2005). *Information and its role in nature* (Berlin: Springer).

Ross, W D (1953) *Aristotle's Metaphysics*. 2 vols. (Oxford: Clarendon Press)

Scalo, J, Craig Wheeler, J and Williams, P (2001), "Intermittent jolts of galactic UV radiation: Mutagenetic effects" In *Frontiers of Life; 12th Rencontres de Blois*, ed. L. M. Celnikier [astro-ph/0104209].

Simon, H A (1992). *The Sciences of the artificial* (Cambridge, Mass: MIT Press).

Sternberg, E M (2000) *The Balance Within* (New York: W H Freeman).

Strauss, S L, Goodman, K S and Paulson. EJ (2009) "Brain research and reading: How emerging concepts in neuroscience support a meaning constructionist view of the reading process". *Educational Research and Review* **4**: 021-033.

Susskind, L (2005), *The Cosmic Landscape: String Theory and the Illusion of Intelligent Design* (New York: Little Brown and Company).

Tanenbaum, A S (1990) *Structured Computer Organisation* (Englewood Cliffs: Prentice Hall).

Vass, D E (2009). *The causal power of social structures* (Cambridge: Cambridge University Press).

Von Baeyer, H C (1998) *Maxwell's Demon: Why Warmth Disperses and Time Passes* (Random House).

---

gfre: 2011-09-01

Level 8	***Cosmology***	***Sociology/Economics/Politics***
Level 7	***Astronomy***	***Psychology***
Level 6	***Space, solar system science***	***Physiology***
Level 5	***Geology, Earth Science***	***Cell biology***
Level 4	***Materials Science***	***Biochemistry***
Level 3	***Physical Chemistry***	***Organic Chemistry***
Level 2	***Atomic Physics***	***Atomic Physics***
Level 1	***Particle Physics***	***Particle physics***

Figure 1*: The hierarchy of structure and causation.* *This figure gives a simplified representation of this hierarchy of levels of reality (as characterised by corresponding academic subjects)for natural systems (left) and human beings (right). Each lower level underlies what happens at each higher level, in terms of causation. There is no correlation between the left and the right hand columns above the level of chemistry, as emergence and causation is quite different in the two cases; but the first four levels are identical (life emerges out of physics!) For a more detailed description,* see *http://www.mth.uct.ac.za/~ellis/cos0.html*.