A Complex Systems Perspective on Neuroimaging Studies of Behaviour and its Disorders

Authors

Affiliations
1 Department of Neuroimaging, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, London, UK;
2 Centre for Psychedelic Research, Department of Brain Sciences, Imperial College London, London, UK;
3 Data Science Institute, Imperial College London, London, UK;
4 Centre for Complexity Science, Imperial College London, London, UK;
5 Global Digital Health Unit, School of Public Health, Imperial College London, London, UK;
6 School of Medicine, Jagiellonian University, Krakow, Poland.

Corresponding author:
Prof. Federico E. Turkheimer
Institute of Psychiatry, Psychology and Neuroscience, King’s College London
Centre for Neuroimaging Sciences
London SE5 9RT
Tel: +44(0) 203 228 3051
Email: federico.turkheimer@kcl.ac.uk

Running title: The Brain as a Complex System
**ABSTRACT**

The study of complex systems deals with emergent behaviour that arises as a result of nonlinear spatio-temporal interactions between a large number of components both within the system, as well as between the system and its environment. There is a strong case to be made that neural systems as well as their emergent behaviour and disorders, can be studied within the framework of complexity science. In particular, the field of neuroimaging has begun to apply both theoretical and experimental procedures originating in complexity science – usually in parallel with traditional methodologies. Here, we demonstrate that the use of such traditional models may distort the outcomes of neuroimaging experiments – hence affecting their interpretability and raising questions about their reliability. Therefore, we argue in favor of adopting a complex systems-based methodology in the study of neuroimaging, alongside appropriate experimental paradigms, and with minimal influences from non-complex systems approaches. Our exposition includes a review of the fundamental mathematical concepts, combined with practical examples and a compilation of results from the literature.
INTRODUCTION

One of the central aims of contemporary neuroimaging is to unify the models of neural structure and function in order to shed light, not just on the evolutionary organization of the brain, but also on its output, emergent behaviour, as well as on its disorders. Intuitively, one would expect the aetiology of brain disorders to be mainly driven by few dominant anatomical regions. However, evidence accumulated in recent years has unravelled a different picture. In fact, most studies indicate that the functional and structural variance in the human brain—during normal functioning, development, ageing and disease—tends to be widely distributed and can hardly be attributed to a set of specific regions (1-3). These findings have impacted the paradigms used to interpret neuroimaging data, which are gradually evolving from phrenological perspectives—that aimed to find associations between the variance in the data and parsimonious sets of regions (see (4, 5) for two contrasting views on the subject)—to embrace distributed analyses (6) that map the data into regional regulatory networks. This latter approach has been borrowed from the rich literature of complexity science, whose main aim is to progress our understanding of the essential features of complex systems. These two perspectives currently co-exist in neuroimaging analyses, and their tools are used, sometimes interchangeably, on similar datasets or within the same study (7-9).

In this work, we combine a number of mathematical concepts with practical examples and observations from the literature to build a case in favor of the use of complexity science analytics to study the brain through neuroimaging techniques. We aim to demonstrate that if the brain is a complex system, and there is strong evidence that it is, then the use of traditional models and associated analytics may distort the outcomes of imaging experiments. This issue, we believe, is one of the key drivers behind the lack of robustness, replicability and, ultimately, poor interpretability of a significant portion of contemporary neuroimaging studies. As a solution, we suggest that when a model of complexity is adopted, the choice of concepts and tools should be embraced together with appropriate experimental designs. We illustrate our position via mathematical concepts combined with practical examples, and a compilation of results from the literature.

THE BRAIN AS A COMPLEX SYSTEM

While the exact definition of a complex system is still under debate (10), there is consensus on four properties that all systems characterised as “complex” necessarily share. These properties are as follows (11):

1. **Multiplicity.** A complex system is made up of a large number of components, each of which interacts with its neighbouring components in relatively simple ways.
2. **Non-linearity.** The interactions between components are nonlinear, e.g. changes in the output are not always proportional to the corresponding changes of the input.

3. **Self-Organisation.** The system self-organises spontaneously, in the absence of any form of centralised control.

4. **Emergence.** The system exhibits emergent behaviour, i.e. the macroscopic behaviour cannot be understood purely in terms of the microscopic interactions (12). Rather, the emergent behaviour may result from: a) multi-scale organization; b) information processing capability; c) dynamical spatiotemporal patterns; d) evolution.

Of these four properties, multiplicity and non-linearity can be seen as the fundamental enablers that – in the right circumstances – allow self-organisation and emergence to happen.

The rest of this article analyses the idea of the brain as a complex system, exploring each of these four properties under the light of recent evidence provided by contemporary neuroscience. It is to be noted that each of these properties has been studied in the neuroscience literature (14-18), but – to the best of our knowledge – there have been no unified accounts which look at the four of them with their interrelationships and consequences.

### 1. **From Phrenology to an Extended View of Neuroimaging Models**

Let us start considering the nature of the various parts that make the brain, which is related to the first property of complex systems. One of the longest-standing challenges in neuroscience is to decipher how variation in neural systems – both structural and functional – maps onto variation in behavioural phenotypes. The initial approach used in the analysis of imaging data was heavily influenced by a highly localized perspective relating neural and phenotypic variability (4). Its use stemmed from the success of neurology/neuropsychology at making differential diagnoses, relating dissociable cognitive/perceptual disorders to apparent localized damage. Even though the localized damage was just a partial view of the underlying pathology, it facilitated the easy intuition that there is a clear mapping between form and function. As a consequence, it focused on the use of mass univariate testing to map brain correlates of such phenotypic variation (19, 20). This paradigm triggered a long and enduring controversy, facing those with the view that it is paramount to protect results from false positives by using stringent significance thresholds (21), against those concerned about the loss of information due to high rates of false negatives (22). To a large extent, this discussion stems from conceptual differences in the *a-priori* expectation of the distribution of the brain’s signals associated with a certain phenotype. Whereas the former position

---

1 Also sometimes referred to as hierarchical organisation, see discussion in 13. Simon HA. The architecture of complexity. Proc Am Phil Soc. 1962;106:467-82.
assumes that variability is highly concentrated in very few brain loci, the latter expects wider associations across brain networks.

Interestingly, this debate is reminiscent of a similar discussion that took place in the early 1900’s amongst geneticists, who were divided between the Mendelians – those inspired by Mendel’s work on pea genetics and saw phenotypes as the result of discrete, monogenic variants – and the biometricians, who did not accept that the inheritance of a continuous distribution of traits (e.g. height, intelligence, etc.) could be related to individual genes. At the time, R.A. Fisher resolved the issue by presenting his seminal work on the infinitesimal model (23), in which the inheritance of continuous traits could be explained by large numbers of variants, each of infinitesimal individual effect. In this theoretical work, Fisher conjectured that a large number of loci should be associated with what, at the time, were termed ‘complex traits’. Only about 50 years have passed since the first genetic markers allowed for the detection of variants that have major effects, and only 20 years since single-nucleotide polymorphism technology provided dense markers throughout the genome. The combination of technological advancements, cost reductions, and increased sample sizes has allowed for the gradual expansion of models describing the genetic foundations of complex traits. It is now generally accepted that common single-nucleotide polymorphisms with small individual effect sizes – generally below the statistical significance levels of experimental designs – account for most of the variability of common genetic traits (24-26).

Equally, complex disorders that are defined over a clinical spectrum demonstrate similarly broad genomic patterns (27). For example, recent work on the genetic architectures of schizophrenia has revealed that more than 70% of 1-Mb genomic regions contain variants influencing the risk of the disorder (28). In fact, a more general ‘omnigenic’ view of complex traits and their heritability and evolution is now driving most of the genetic contemporary discourse (29, 30).

Moving from the genotype to the intermediate phenotype, there is increasing evidence that some of the remarkable computational capabilities of the brain, such as object recognition – (still unmatched by machines) (31) – depend on the ability of neural systems to encode simple inputs into redundant widespread representations (31) and that this distributed architecture supports executive function (32). The conceptual models used in neuroimaging studies have evolved in the same direction. The field has been gradually shifting its focus from the analysis of to the analysis of covariance, e.g. moving from localized effects to distributed networks (33-37) However, a conscious, argumentative debate between these approaches has been lacking, and the two methods still co-exist. In particular, there remains the unanswered question of whether complex behaviour is supported by activity concentrated in few loci or by widely distributed networks. The latter view is supported by
substantial evidence: Gonzalez-Castillo and colleagues used very large sample sizes or very extended single subject designs to reveal brain-wide functional support for simple tasks in fMRI studies (38, 39). Similarly, Haxby and colleagues (40) demonstrated that the representations of faces and objects in ventral temporal cortex are widely distributed and overlapping and second, using structural data, Chu et al (41) demonstrated that feature selection (e.g. regional variation) was as effective as total brain volume at predicting the progression from minimally cognitively impaired cohorts to Alzheimer’s disease.

A complex system perspective on neural function results in predictions for the distribution of measured neural activity across the brain, which differs from predictions due to the localized perspective. For discussing some of these predictions, here we consider fMRI data from the Individual Brain Charting (IBC) project (42), which consists of a fMRI-curated dataset, acquired in a fixed environment from a permanent cohort of the same 12 participants during the performance of a dozen tasks, encompassing a wide range of motor and psychological domains. The results are presented in Figure 1, following the format in (29) where the z-scores obtained from the testing of each task-vs-control contrast (only positive z-scores considered) were sorted in descending order and plotted versus the log of their rank. Figure 1A and 1B depict simpler (mostly motor) and higher mental processing tasks, respectively.

Figure 1: The panels illustrate the vectors of z-scores extracted from task-vs-baseline contrast (positive values only) maps obtained from fMRI studies of simple/motor tasks (1A) or more complex activities (1B). The x-axis represents the logarithm of the z-score ranks. The red lines represent the same z-scores for the normal null distribution. The blue lines represent the James-Stein boundaries, i.e. the minimal decay of a z-score distribution that would justify the use of statistical thresholds for the estimation of the z-scored vector for each contrast (see Eq. [1]). The boundary curve has been arbitrarily scaled to the maximum z-score value for the group of tasks for which the theory suggests the use of shrinkage estimators for all vectors with smoother decays.
These plots are to be contrasted with the expected null distribution (red) and the decay (blue):

\[ z_k = C k^{-1/2}, \]

where \( z_k \) represents the z-score of rank \( k \); \( C \) is an arbitrary constant (which in the two plots in Figure 1 was fitted to the max z-score in each panel); and the exponent \(-1/2\) is a minimum bound on the rate of decay of the z-score vector – the estimation of which warrants the use of a statistical threshold, whereas vectors with a slower rate of decay require approaches based on weighted averages of the whole distribution (43, 44).

It may be seen that the curves in Figure 1 decay in a smooth manner – even in the simpler tasks, in which one would expect the signals to be described by a small number of coefficients with steeper decay rates - and do not satisfy the minimum requirement for the use of thresholding. In other words, the boundary in [3] is a way of formally stating that the use of thresholds to separate signal from noise in a multivariate vector is unwarranted when no such clear separation exists. In fact, any small change in experimental setting, imaging technology, pre-processing, as well as inter-individual variability (both neural activation, morphological or non-neural) could easily swap the ranks of the z-scores, thus limiting the reproducibility of the results (43). Furthermore, these plots confirm that the activity supporting the execution of a very large variety of tasks is distributed across much if not the entire neural system and that the brain therefore satisfies the first requirement in being classified as a complex system.

2. Interaction and Nonlinearity

Let us consider now non-linearity, the second key aspect of complexity. Nonlinearities within a complex system arise as a result of interactions between its elements. The brain offers several examples of nonlinear relationships between its physiological parameters. For example, there is the way in which V1 neurons respond to inputs (45) – where object recognition is performed by nonlinear operations on distributed, largely redundant representations (45) – or the response of evoked deep tissue pain (46) – where the response is linearly related to pain intensity in certain regions, but is nonlinear in others.

However, the point we want to make in this section is that nonlinearities in complex systems are generated by the interactions amongst elements when changes are applied to the organic whole. To illustrate this perspective, let’s start with a simple theoretical illustration, which will then showcase how these inferences might apply to a real scenario. As a simple example,
let us consider a population of rabbits that reproduce regularly with a certain birth-rate $R$. The relationship between the number of rabbits at time $t$, denoted by $N_t$, and the corresponding quantity at time $t+1$ can be expressed as

$$N_{t+1} = RN_t. \quad [2]$$

In particular, if one considers the case $R = 2$ and runs three iterations of the model, starting from one rabbit at $t=0$ we obtain eight rabbits at $t=3$, while if we start with 2 rabbits the final total is sixteen. This confirms that this system is linear and that the totality of the system is equal to the sum of the parts, as two rabbits produce twice the offspring of a single one.

The linear model above can be readily turned into a nonlinear complex system with interacting elements if we introduce a constraint wherein offspring need to compete for resource. Then, the model in (2) turns into the simple, but effective Verhulst model (47):

$$N_{t+1} = R \left( N_t - \frac{N_t^2}{P_{\text{max}}} \right), \quad [3]$$

where $P_{\text{max}}$ represents the maximum capacity of the environment to keep the offspring alive. Setting $P_{\text{max}} = 10$, we run 3 iterations of this algorithm with an initial state of one rabbit to obtain four offspring. Running the same model with three iterations and starting with two rabbits we obtain five offspring. In this case, the offspring of two rabbits is not five-times the offspring of one rabbit (see example illustrated in Figure 2).

**Figure 2:** The panel above illustrates a simple linear growth model whereas the offspring of two rabbits is always twice the number of offspring of a single rabbit, whatever the number of iterations.

In the logistic model below, a resource constraint makes the rabbits compete for resources and, this time, the offspring of two rabbits is not twice the offspring of one.
The above example demonstrates three important phenomena. Firstly, a simple resource constraint on the system introduces a highly non-trivial interaction between elements, so that the totality is not anymore equal to the sum of the parts. Secondly, this interaction translates into a non-linear behaviour observed within the system. Thirdly, the reproduction rate in (3) depends on the initial number of rabbits $N_t$.

The insights developed over the simple example can be used to understand more complicated issues related to brain development, brain ageing, and their disorders. For example, it has proven difficult for neuroimaging studies in autism to replicate both structural and functional results (48, 49), whereas the most consistent finding is the heterogeneity of imaging measures (50). One particularly puzzling finding is related to brain development of subjects within the autistic spectrum: these populations exhibit increased brain volumes in young children, whereas the rate of growth decreases subsequently during adolescence to meet the normal developmental curve between 10 and 15 years of age, with the volume of many structures finally declining atypically into adulthood (51-53). These findings are paradoxical from the perspective of a linear system; however, it is likely that a suitable nonlinear model of growth could explain why a more numerous initial state would experience a strong bias downwards in its developmental rate, especially when the whole system is placed under exogeneous pressure.

More generally, it is often noticeable in neuroimaging that individual developmental trajectories seem to provide much better resolution to categorize psychiatric disorders than cross-sectional data (54). This phenomenon is known in the literature of complexity science and dynamical systems as non-ergodicity, which characterises dynamical systems in which the statistics computed over single trajectories are not necessarily equal to the statistics across the population (55). This is equivalent to say that trajectories that start from different initial conditions may exhibit radically different dynamics. Non-ergodicity has dramatic consequences in the context of ageing and age-related disorders, where imaging studies have shown that cross-sectional analyses produce findings that are often in discrepancy with longitudinal analyses on the same subjects (56). In particular, longitudinal analyses unveil complex nonlinear trajectories of structural and functional data of anatomically parcellated regions, which can have radically different dynamics challenging most available explanatory models (57-60). Notably, during normal ageing, some brain areas may exhibit little to know age effect on regional cerebral volume, while others show a complex, non-linear relationship with age. This is mirrored by the significant intra-subject variability of cognitive decline –
see, for example, the findings of the Nun Study (61), in which cognitive domains within the same individual deteriorate at very different rates (62). When it comes to connecting the intermediate phenotype (i.e., the brain) with environmental modifiers on the one hand and genetic factors on the other, recent analysis of very large datasets have demonstrated that individual variability of MRI measures across ageing is spread across a large number of regional networks, or modes – each associated with overlapping environmental factors and genes (63). Interestingly, the biological and genetic factors involved in ageing accrue over three fundamental processes: metabolism, stress (particularly its effects on cellular senescence), and immunity. However, each of these processes is in itself a complex outcome of interactions amongst multiple genes and environmental factors, such as nutrition, environmental/psychosocial stress, and infections (64, 65).

Motivated by the above evidence, we propose here that the complex developmental and neurodegenerative patterns observed in structural and functional data are the result of non-linear interactions generated by systemic pressures, which stem from metabolic and immune challenges punctuated across times (66-68). For example, metabolism is a fundamental determinant of neural function, as the brain accounts for ~2% of total body mass, yet burns ~20% of its energy, which is obtained through oxidative glucose metabolism, thus rendering the brain in constant demand of oxygen supply from cerebral blood flow (69). In fact, the human brain is, from a metabolic perspective, a scalable version of a primate brain, as the energy budget of the whole brain per neuron is approximately fixed across species and brain sizes (70) and hence its development has been strictly controlled by metabolic constraints (71). Brain metabolism is mostly devoted to sustaining synaptic activity (72) and metabolic rates influence neural circuitry and activity patterns by exerting selective pressure toward metabolically efficient wiring patterns (73), grey/white matter segregation (74), neuronal morphology (75), and neural codes (76, 77). The metabolic demands of neural circuitry render the system highly susceptible to mitochondrial impairments, oxidative stress and deficits in glucose metabolism and perfusion. Therefore, the resulting global metabolic pressure stemming from noxious metabolic events translates into complex, non-linear interactions amongst the system’s constituents. Furthermore, imaging, as well as cognitive and behavioural outcomes are strongly dependent on the timing of events across the age-span (78).

Following the same principles, we propose that immunity-based second processes, that are recurrently associated to either developmental or neurodegenerative disorders, may generate system-wide pressures that then translate into complex outcomes. In fact, either maternal infections during gestation, or activation of the innate or adaptive immune system caused by
stressful events, trauma or infections across a lifespan, have demonstrated to be the cause of psychiatric or degenerative conditions (79-84). The exact mechanism by which the immune system mechanistically affects brain function is still undetermined. One model proposes that peripheral cytokine concentrations may cause leakages in the blood-brain barrier, with activation of the immune system and consequential interference in transmission and plasticity (81, 85, 86). Alternatively, we have recently proposed an indirect model of peripheral versus central immunity activation, where the mild peripheral inflammation observed in subjects with psychiatric conditions causes stiffening of the endothelial layers. This in turn causes a reduced permeability of the blood brain barrier to the transmission of small molecules, which (in chronic cases) leads to disturbed homeostasis and reduced functional activity (87).

3. SELF-ORGANIZATION AND CRITICALITY

The third key property of complex systems is self-organization. Self-organization takes place whereas some form of global order arises as a consequence of the local non-linear interactions between parts of an initially disordered system (88). Importantly, the organization should emerge in a spontaneous manner, provided energy is available, and not as a result of the actions of a centralized controller or an external guidance. Therefore, the self-structuring events are often triggered by seemingly random fluctuations, which are dynamically amplified by some form of feedback enabled by the underlying non-linearities. The resulting organization is wholly decentralized, being distributed over all the components of the system, making it robust and able to endure (i.e. self-repair) or recover from substantial perturbations (89). The process of self-organization of a dynamical system can be visualized as a trajectory in $n$-dimensional phase space (where $n$ is the number of its elements) that evolves towards equilibria that can be described in terms of limit cycles. The evolution of the system is constrained to remain on the limit cycles and this constraint implies some form of coordination between its constituent elements (90, 91). A substantial amount of evidence, based on both human and animal experiments, indicates the important role of neural plasticity in the maturation of perceptual and cognitive processes. In other words, neural structures underlying these functions require sensory input for their maturation and genetic instructions are not sufficient to specify neuronal connections with significant precision (92, 93).

Importantly, the brain (as many natural complex systems) displays macroscopic patterns of activity with spatial and/or temporal scale invariance. The latter is a property which results in dynamical systems, e.g. those exhibiting turbulent flow (94), evolving in such a way that (except at the spatial or temporal extremes) it is not possible to distinguish the scale at which
the system is being observed. This self-similar (or “fractal”) character suggests that the self-organizing process that is ongoing during the development of the central nervous system (CNS) drives the system towards the inflection point of a second-order phase transition (11, 88). Pier Bak used the sand pile model as an exploration of self-organized criticality (SOC) (95). A sand pile can be organized by a simple rule that sees the grains falling randomly on a grid and accumulating until they topple sideways once they have reached a certain height (96). Once the grid starts filling up, avalanches of all sizes start cascading with an overall size-probability distribution which demonstrates the characteristic power law behaviour observed in self-similar dynamical systems (Figure 3).

Figure 3: A simple sand-pile experiment simulated on a 20x20 grid (left) whereas sand is dropped randomly on the square, and it topples sideways, as well off the grid, when a set limit for the height of a sand column is reached (4 grains in this case). The right panel shows the total sand in the grid through time, as well as the size of the avalanches falling off it.

Many natural systems tend towards such power law distributions, as these distributions allow for maximal adaptivity and other functional benefits (97). As in the case of the sand pile model, any perturbation to the system can be passed across scales and dissipated quickly – this endows the system with stability (98). Another example of this phenomenon is the energy cascade in turbulent hydrodynamic systems, in which kinetic energy is passed between eddies of decreasing size until it is distributed, at the smallest scales, to the viscosity of the liquid (99).

Other important requirement for self-organization is the abundance of available energy, in the same way that the sand pile needs a constant supply of grain to sustain itself in a critical state. Equally, the brain devotes up to 80% of its energy budget to maintaining its intrinsic activity (i.e. brain’s activity in the absence of an overt motor, perceptual or cognitive task), whereas evoked activity in the brain does not use more than 1% of its metabolic reservoir (100, 101). This feature has been dubbed the Brain’s Dark Energy (102).
Self-organization rests upon four basic requirements (103): interactions between multiple constituent elements, strong non-linearity, possibly involving positive and negative feedback, a dynamic behavior striking a balance between exploitation and exploration, and availability of energy, which allows to overcome the natural tendency to increase entropy (e.g. disorder).

The above requirements can be readily observed in the organization of cortical structures. As summarized in (104), the brain (primarily the cortex) is made up of a multitude of elementary units, each of which consists of two neurons – a pyramidal excitatory unit coupled with a gamma aminobutyric acid (GABA) interneuron that, via a delayed negative feedback, transforms the firing of the unit into an oscillation in the gamma frequency (105), the so-called pyramidal interneuronal network gamma (PING) network. These elementary units are weakly connected via further GABA interneuron projections into and across the individual neural columns (106). Therefore, local oscillations in the gamma frequency are the binding force that generates rapid metastable dynamics, moving between synchronization and de-synchronization, which spread across the cortex. These elementary computational motifs have been observed across various brain regions, as well as across scales (104), reaching the long-range synchrony of neural activity. This supports the view of the brain as SOC system (107-110), with a distinctive fractal signature both in its structure and function (104, 111, 112). The fractal distribution of neural output, also known as the 1/f noise, can be observed in the behavioural repertoire, from the motor regions to the higher cognitive areas, capturing the range between basic physiological responses and creative abilities in the human brain (113, 114). Conversely, evidence suggests that this scale-invariance in brain output may be lost during anaesthesia-induced loss of consciousness (115).

4. **EMERGENCE**

The fourth property of complex systems is that they are characterized by emergent behaviours. Complex systems, i.e. organisms or environments made up of very large numbers of elementary agents (e.g., ants in colonies, birds in a flock, or arrays of molecules or weather elements), can produce an impressive array of sophisticated behavior that cannot readily be explained solely by the physical properties of their constituent components. As a matter of fact, emergence – although still lacking a precise definition – has come to be recognized as a key ingredient of any system studied under the umbrella of complexity science (12).
The concept of emergence is credited to the physiologist and philosopher George H. Lewes in his book Problems of Life and Mind (116). At the time, Lewes was concerned with reductionism, i.e. why all science doesn't ultimately reduce to physics. Victorian England in times of Lewes saw important advances in mechanics; however, chemical reactions (e.g. the production of water from oxygen and hydrogen) seemed to require explanations that were qualitative different from those that contemporary physics could provide. Building on this idea, Lewes drew a difference between mechanical effects and chemical effects—which he called “emergents”. In his words:

“... The emergent is unlike its components in so far as these are incommensurable, and it cannot be reduced to their sum or their difference...” ((116), pp.413).

The notion of emergence was first brought into the neurosciences by Roger W Perry, the 1981 Nobel laureate for physiology and medicine, in the context of the mind-body problem. According to Perry, emergence “... is the idea, in brief, that conscious phenomena as emergent functional properties of brain processing exert an active control role as causal detents in shaping the flow patterns of cerebral excitation. Once generated from neural events, the higher order mental patterns and programs have their own subjective qualities and progress, operate and interact by their own causal laws and principles which are different from and cannot be reduced to those of neurophysiology.” (117). In other words, the behaviour of the whole cannot be explained purely by focusing only on local spatiotemporal interactions between its elements in isolation— new theories are needed to explain these phenomena. Similarly, higher order cognitive function operates via laws that cannot be reduced to those of neurophysiology – instead they operate in the domain of psychology.

Despite their popularity, strong emergent approaches have been the source of numerous controversies. From a philosophical perspective, strong emergence has been accused of being logically inconsistent (118) and sustained on illegitimate metaphysics (119). Additionally, diverse concerns with respect to strongly emergent approaches in neuroscience have been raised, including arguments related to their plausibility in regard to empirical evidence (120), or to potential limits of their falsifiability (121). Finally, the fact that consciousness is usually the only proposed example of strong emergence in the natural world (see e.g. (122)) presumes an exceptionalism of consciousness that could be hard to justify. These difficulties are avoided by “weak emergent” approaches, which rely on computational models that start from biologically plausible elementary units and build higher order outputs, using tools and concepts borrowed by complexity science (12). These models are defined as “weakly” emergent, are not limited by ontological faults and have been proven to
be scalable and able to realistically simulate the hierarchies of brain output and have now reached a level of maturity that enables predictions in the clinical realm (104, 120, 123-127). More technically, weak emergence can be broadly categorised into computational irreducible and causal emergence approaches. The former, introduced by Bedau in (119) based on ideas from Wolfram (128), posits that emergence relates to limitations that hold in practice but not in principle; specifically, emergent properties would be generated by elements at microscopic levels in such complicated ways that they cannot be derived via explanatory shortcuts, but only by exhaustive simulation. While theoretically attractive, these approaches currently lack of ways of operationalizing their definitions, and hence their practical value is limited. In contrast, causal emergent approaches identify formal methods to assess the causal power exerted by macroscopic properties of a system. Causal emergence has been operationalized using Granger Causality (160), Pearl’s do() calculus and effective information (158,159), and most recently by Integrated Information Decomposition (162).

5. The Tools
Having examined neural function in terms of the four main properties associated with complex systems, we now address the question of how the obtained insights can be operationalised. Realising the large number of degrees of freedom that characterise the brain, one suggested approach is to shift the focus from the analysis of variance to the analysis of covariance. In other words, inference should be driven by the way the elements of the systems interact, and less so on how much they vary individually. The statistical literature provides numerous multivariate methods to study data covariance, which have been already been adapted to use in neuroimaging. A first approach involves using methods such as principal component analysis (PCA) (34), independent component analysis (ICA) (37), and partial least squares (33). Although not aimed at complexity analysis per se, these tools can be used to explore patterns in the data, as well as to derive measures of complexity. For example, Morgera Information Complexity is a measure of the extent to which covariance is distributed across principal components (129). A second approach is to represent covariance as networks, which are a popular tool to capture and explore covariance structures (6). Networks can be built on multiple metrics, ranging from standard correlation coefficients to mutual information or model-based indices that are generally better suited to capture non-linear interactions (130). Non-linear metrics are often preferred, as they avoid spurious patterns that can be caused by unjustified linear assumptions (131). Once the web of interactions has been modelled as vertices and edges, the topology of the structure can be further explored, both at the macro-scale level – using graph theoretical measures that summarize the network properties (35) – and at the meso-scale level where the topological stability of the
structure can be further quantified (132, 133). A third approach to analyse covariant structures is to explore multivariate correlations, which capture high-order statistical properties that go beyond pairwise phenomena (134-137). High-order interactions can be used to build hypergraphs, which are networks where links can connect more than two nodes (138). Hypothesis analysis is an active area of ongoing research, which is opening novel avenues for neuroimaging research (see e.g. (139)).

Importantly, the brain is a non-linear system and, as such, its behaviour is dynamic. The discussions above have clarified the importance of experimental designs that move beyond stationary cross-sectional data paradigms and instead attempt, whenever practical, the acquisition of individual trajectories. This also enables the analysis of initial conditions, which are often relevant for complex systems; we have already shown how the artificial splitting of life-trajectories in development-adulthood-ageing may result in faulted inferences. The consideration of trajectories is also relevant at smaller timescales, including the duration of a BOLD MRI study where higher resting-state functional activity before a stimulus can result in less activation or more deactivation (140). It is often useful to characterise dynamics in terms of “states” or attractors, that is, regional patterns of activity between which the system oscillates (89). The number of these states and their stability fully define such dynamics and can be quantified using standard analytics (141-145) (see Figure 4) and have been used successfully in recent work on a broad spectrum of psychiatric disorders (146-152).
Figure 4: Brain states are more or less likely to be visited by the dynamics (red trace) over time by their relative energy/stability. Brain states are defined as intervals of temporal coherence (adapted from (153)).

Furthermore, if one accepts SOC as a model for brain, then the use of fractal geometry can be justified in studying brain structural data (111, 112) and its deviations from the SOC model. Fractal principles enable parsimonious methods to explain rich structures, holding great potential to explain brain structure and function. SOC approaches, however, face a number of important challenges (154), including the analysis of applicability of principles developed to study homogeneous systems (made e.g. by equivalent spins) over heterogeneous systems such as the brain. Also, there is a need to operationalize the beneficial properties of criticality, such as efficiency and robustness, so that they can render verifiable predictions. It must be clarified that these challenges do not put in question the core contribution of SOC into neuroscience, namely, that by neglecting the right details neuroscience might be capable of finding general underlying principles (155, 156).

Finally, the study of complex systems requires the quantification of two key properties; diversity and emergence. The first one can be described by a measure of entropy, for example Shannon’s entropy (e.g. $\Sigma_i p_i \log p_i$ where $p_i$ is the probability of state $i$), or variations of it including the Excess entropy (157, 158) and the entropy rate – this being closely related to the Lempel-Ziv complexity (159) and the statistical complexity of epsilon machines (160, 161). The second one, emergence, is a unique attribution of complex systems; the availability of one such measure would be a key tool in understanding the information flows between the various hierarchies/layers of models. So far, attempts to operationalize the various models of emergence have had limited applicability (162-165). This work is, however, fundamental to the progress of complexity science in general and it application to the neurosciences in particular, and novel modelling approaches and effective measures are appearing in the recent literature (166). These quantitative approaches allow researchers to frame hypotheses about emergence in a formal, rigorous manner, and test these hypotheses on fine-grained data-driven alternatives to traditional all-or-none classifications, opening a broad range of possibilities for applications.

6. Conclusions

We reviewed neural systems from the perspective of complexity science and matched the required properties of complex systems with evidence from contemporary neuroscience. The use of complexity science is not new to studies of the brain. However, its use has often been intertwined with standard analytical approaches that stem from older conceptual
frameworks. What is now required is the acceptance of these new frameworks both in terms of experimentation and analysis. Complexity science is itself experiencing ongoing progress, and some of the required tools are perhaps yet to be discovered. Nonetheless, the potential insights to be gained by this paradigm are sufficiently promising to render this a very worthwhile endeavour.
REFERENCES


87. Turkheimer F, Althubaity N, Nettis MA, Cousins O, Dima D, et al. Increased serum peripheral C-reactive protein is associated with reduced small-molecule brain perfusion in healthy volunteers and subjects with major depressive disorder. 2020;2020.06.29.20138438.
103. NBonabeau E, Dorigo M, Theraulaz G. Swarm intelligence: from natural to artificial systems: OUP USA; 1999.


163. Hoel EP. When the map is better than the territory. Entropy. 2017;19.
