**Mental Machines**[[1]](#footnote-2)

David L Barack

Columbia University

Abstract

Cognitive neuroscientists are turning to an increasingly rich array of neurodynamical systems to explain mental phenomena. In these explanations, cognitive capacities are decomposed into a set of functions, each of which is described mathematically, and then these descriptions are mapped on to corresponding mathematical descriptions of the dynamics of neural systems. In this paper, I outline a novel explanatory schema based on these explanations. I then argue that these explanations present a novel type of dynamicism for the philosophy of mind and neuroscience, componential dynamicism, that focuses on the parts of cognitive systems that fill certain functional roles in producing cognitive phenomena.

1. Introduction

Two prominent approaches to the analysis of cognition are dynamicism and computationalism. Dynamicism analyzes cognition in terms of interacting subsystems of the brain, body, and environment and advocates for the explicit use of dynamical systems theory to explain cognition. But this approach ignores the subsubsystems, the parts of the subsystems of the cognitive system such as the parts of the brain. In contrast, computational approaches analyze and explain cognitive phenomena in terms of such parts and their functions. In this paper, I resolve some of this tension by presenting an explanatory schema that draws on both dynamicism and computationalism.

Not only does this schema partly reconcile these two positions to understanding the mind, it also provides novel insight into cognitive phenomena. The account of explanations in cognitive neurobiology, the study of cognition through the investigation of activity in single neurons and neural populations, suggests a sketch of an argument for a particular way to understand cognition. Cognitive systems are componential dynamical systems, a view I call componential dynamicism, which I elaborate below. Hence, this paper has two mutually supportive goals. First, I aim to outline a novel explanatory schema present in cognitive neurobiology. Second, this explanatory schema illustrates how to partly reconcile dynamicist and computational approaches to cognition and how to analyze cognitive phenomena in a way that is most consilient with neurobiological practice. In this essay, I first make an induction, looking at the description and explanation of cognitive phenomena in cognitive neurobiology and educing an explanatory schema from this discussion. I then draw on this explanatory schema to present three arguments for componential dynamicism. I conclude that cognitive systems are componential dynamical systems.

Cognitive neurobiologists explain mental phenomena using an increasingly rich array of dynamical systems. In these explanations, cognitive capacities are decomposed into a set of functions, each of which is described mathematically, and then these descriptions are mapped on to corresponding mathematical descriptions of the dynamics of cognitive subsubsystems, such as the parts that make up the brain. The recent turn to dynamical systems underlying cognitive phenomena, however, departs from previous characterizations of dynamicism by analyzing the brain into dynamical systems whose coordinated activity results in cognitive phenomena. In this paper, I will argue that this componential dynamicism is a novel and needed approach to understanding cognition.

Past dynamicist approaches to cognition (see, e.g., Port and van Gelder 1995; van Gelder 1995; Chemero 2011; Zednik 2011) emphasized interactions between agents, whether the brains and bodies of organisms (Stepp, Chemero et al. 2011) or simulated systems (Beer 2000), and their environment (Port and van Gelder 1995). These approaches focus on the coupled, reciprocal causal influence of agents on the environment and the environment on agents (Clark 2013). Finally, they describe cognitive systems with the use of global parameters and collective variables (Walmsley 2008; Zednik 2009; Shapiro 2013; Lamb and Chemero 2014). These hallmarks amongst others have served in the past to individuate dynamicism from computational approaches to understanding cognition.[[2]](#footnote-3)

Computational approaches to cognition emphasize computational operations inside agents to explain cognitive phenomena (Putnam 1967; Fodor 1975; Dennett 1981; Lycan 1981). In contrast to dynamicist approaches, computational ones describe the transformation of input representations into output by processes inside the system (Newell 1980), focus on internal variables and local parameters that are transformed in light of input to and the goals of the system (Anderson 1990), and often eschew environmentally coupled causal interactions (Grush 1997). This focus on local transformations of representations is consistent with the mechanistic approach to explanation in biology and neuroscience (Craver 2007; Kaplan 2011; Kaplan and Craver 2011; Craver and Darden 2013). The different emphases of dynamicism and computationalism have often been contrasted in the philosophical literature (van Gelder 1995; Bechtel 1998; Beer 2000; Wheeler 2005).

Recent research uncovering the neural circuits of many different cognitive phenomena both diverges from and has ground in common with both dynamicism and computationalism. Cognitive neurobiology uses dynamical systems theoretic models to explain cognitive phenomena, like dynamicism, but with a focus on the parts of the system, like computationalism. To illustrate, consider the plight of a monkey living in the Taï forest of the Ivory Coast. In those African forests, leopards prey upon many different species of monkeys (Zuberbühler and Jenny 2002). These leopards are diurnal and hide in order to ambush monkey groups (Zuberbühler, Jenny et al. 1999; Jenny and Zuberbühler 2005). Forest leopards are large spotted cats, and detecting their presence requires distinguishing the camouflaged feline from the forested habitat in which the monkeys and leopards live (Bailey 1993; Jenny 1996). The dappled coats of the cats present a cryptic problem for primates: how to distinguish this dangerous predator from the foliage concealing its movements?

A clue to the systems underlying such life-or-death decisions can be found in contemporary research into the cognitive and neurophysiological processes of motion perception and decision-making. Central to the description and explanation of these cognitive capacities are dynamical systems that perform the computations necessary for cognition and are instantiated by the subsystems of the brain. To detect the direction of motion, a number of functions must be executed, including setting the prior probabilities on the various hypotheses (viz., the different possible directions of motion), processing evidence from the noisy sensory stream, establishing a threshold for making a decision, summing the evidence for different hypotheses, detecting when the summed evidence crosses the threshold, selecting a response, initiating a response, and so forth. In the case of motion detection, a dynamical system plays a key role in integrating motion evidence for making perceptual decisions (Roitman and Shadlen 2002; Gold and Shadlen 2007). This system is instantiated in a brain region, the lateral intraparietal area, which integrates the evidence for the direction of movement until a particular threshold is reached. Upon crossing the threshold, the decision is made and action is initiated while the system resets to a baseline. This simple system partly executes the functions underlying the ability to make perceptual decisions, specifically executing the function for the integration of evidence, in such degraded sensory conditions.

There are many different dynamical systems instantiated in the brain to execute the functions that compose cognitive capacities. I contend that this dynamical turn presents a novel philosophical theory of cognition, one that retains dynamicism’s focus on dynamical systems theoretic models but that endorses the functions of components found in computationalism. In this essay, I first present the classic dynamicist view, and distinguish between systemic and componential dynamical systems. The classic view appeals to systemic dynamical systems to explain cognitive phenomena. This essay will advocate on behalf of componential approaches. I then discuss a series of case studies from cognitive neurobiology including perceptual decisions, strategic decisions, attention, and reward encoding. These case studies will illustrate a novel dynamicist explanatory schema in cognitive neurobiology and establish its scope. I then argue that the dynamical systems appealed to in such explanations imply that cognitive systems are also componential dynamical systems.

1. Componential Dynamicism

Recent research in cognitive neurobiology is revealing componential dynamical systems in the brain. More precisely, explanations of cognitive phenomena refer to such dynamical systems. In order to understand these explanations and what they imply about the nature of cognitive systems, componential and systemic dynamical systems must be distinguished. I begin by introducing the concept of a dynamical system. I next outline the extant form of dynamicism in the literature, systemic dynamicism. I illustrate the view with an example, the systemic dynamicist explanation of motor coordination. I finally distinguish between componential and systemic dynamical systems and tell a novel story about the same example. In the next section, I turn to a lengthy discussion of the case study of perceptual decision making.

* 1. Dynamical Systems

The following discussion will rely heavily on the concepts and terms used to describe dynamical systems. A system is a set of individuals, the elements that make up the system. Dynamical systems are sets of individuals that change over time or with respect to one another.[[3]](#footnote-4) These individuals can be objects, properties, or relations. This description of dynamical systems is general enough to capture almost all systems, as almost any collection of objects, properties, or relations involves such change.

Dynamical systems are described using dynamical systems theory, a branch of mathematics that describes the ways that systems change and that contains concepts and tools for understanding this change.[[4]](#footnote-5) A basic concept in dynamical systems theory is the state space, the set of all possible states that a system can occupy. A state for such a system will be defined as the set of determinate properties for the determinable types of properties of the system. These states are mathematically described by a set of values for all the variables and parameters of the equations that describe the system, the system’s state equations. These state equations are either difference equations, describing how the variables change in discrete time or with respect to another variable, or differential equations, describing the change in continuous time or again with respect to another variable.[[5]](#footnote-6)

Systems start in a particular state and evolve over time through any number of other states.[[6]](#footnote-7) This movement through the system’s state space is called a trajectory. The system’s state space can possess structure, such as when trajectories tend to converge on or near a single state, called an attractor. Attractors can be points in the system’s state space, called a point attractor, or sequences of states through which the system repeatedly passes, called a ring attractor. There are other types of attractors as well. Attractors can be stable, points toward which, within a certain range, the system evolves along all dimensions of variation in the system, or unstable. There are numerous other properties of dynamical systems in addition to attractors (Strogatz 2001). A dynamical system is described in terms of the system’s trajectories through the system’s state space, which I call the system’s evolution. The system’s evolution is mathematically described by the system’s state equations.

* 1. Systemic Dynamicism and the Description of Dynamical Systems

Traditionally, dynamicism in cognitive science views cognitive systems as dynamical systems composed of brains, bodies, and environments that stand in coupled interaction. Call this systemic dynamicism, as cognitive phenomena arise from the interactions of these subsystems. Port and van Gelder provide the best succinct description of systemic dynamicism, whose

“…core is the application of the mathematical tools of dynamics to the study of cognition. Dynamics provides for the dynamical approach… a vast resource of powerful concepts and modeling tools. But the dynamical approach is more than just powerful tools; … it is a worldview. The cognitive system is… a dynamical system. It is… the whole system comprised of nervous system, body, and environment. The cognitive system is… a structure of mutually and simultaneously influencing change. Its processes… unfold in the real time of ongoing change in the environment, the body, and the nervous system. The cognitive system does not interact with other aspects of the world by passing messages or commands; rather, it continuously coevolves with them.” (Port and van Gelder 1995, p. 3)

Using dynamical systems theory, cognitive phenomena correspond to trajectories through the possible states that the coupled subsystems (the brain, body, and environment) can take. Hence, cognitive phenomena result from the interactions of brain, body, and world as these subsystems traverse their joint state-space (Beer 1995; Port and van Gelder 1995; van Gelder 1995; Beer 2000; Wheeler 2005; Chemero and Silberstein 2008; Zednik 2008; Chemero 2011; Stepp, Chemero et al. 2011; Zednik 2011; Zednik 2015).

Three main properties characterize systemic dynamicism for this discussion. First, cognitive phenomena result from the interaction of three subsystems, the brain, body, and world, as “…the true cognitive system is a single unified system embracing all three” (van Gelder 1995), p. 373). Systemic dynamicism rejects any focus on the parts within these interacting subsystems, the subsubsystems such as the internal states and processes of the brain, body, or environment. Instead, systemic dynamicism emphasizes how these subsystems interact and evolve over time.

Second, these interacting subsystems are described holistically, using “differential equations that are defined over a small number of collective variables and global control parameters” (Zednik 2009, p. 2298). The features of these models such as variables, parameters, and the functional relationships *inter alia* describe such systems. Collective variables stand for changing properties of the whole, interacting set of subsystems, typically denoting some relation between two or more elements of the set. Global parameters stand for relatively stable properties of the set of subsystems that control its behavior. Hence, these collective variables and global parameters avoid reference to parts of the whole system.[[7]](#footnote-8)

Third, these subsystems are coupled, with the state of any one of the subsystems influencing and being influenced by the others (what Clark refers to as “continuous reciprocal causation” (Clark 2000; cf. Grush 1997 on the ‘strong coupling thesis’). Stepp et al. advocate for viewing cognition as “the ongoing, active maintenance of a robust animal–environment system, achieved by closely coordinated perception and action” (Stepp et al. 2011, p. 432). This close coordination is achieved via coupling between the brain, body, and world. As Clark puts it, the dynamicist view “is of two coupled complex systems (the agent and the environment) whose joint activity solves the problem” facing the agent (Clark 1997, p. 98). For this discussion, when two dynamical systems are coupled, the systems mutually determine each other’s properties. The properties of one system (whether its parts or the entire system) help determine the properties of another system (in part or whole), and the latter system also simultaneously helps determine the properties of the former. Coupling is reflected in the state equations for the system or for its parts; sets of equations are coupled when a variably *y* is a function of another variable *x* (e.g. *y* = *αx* + …) and, in turn, *x* is a function of *y*. Systemic dynamicism maintains that coupling occurs between the brain, body, and world, which implies that the state equations for each contain terms referring to the states of the others. In sum, on systemic dynamicism, cognitive phenomena result from the coupled coevolution of brain, body, and environment as described by the mathematical tools of dynamical systems theory.[[8]](#footnote-9)

As illustration of the systemic dynamicist approach to cognition, consider the Haken-Kelso-Bunz (HKB) model of sensorimotor coordination (Haken, Kelso et al. 1985; Kelso 1995).[[9]](#footnote-10) The HKB model describes coordination between sensory and motor systems such as occurs when two fingers on opposing hands oscillate back-and-forth. Subjects are instructed to waggle their fingers in this fashion. The movement of the fingers tends to be either in phase, moving back and forth at the same time, or out of phase, moving in opposite directions. The in-phase or out-of-phase motion are the two stable point attractors of the system. The relation between the fingers is captured by the relative phase of motion of the two fingers, *ϕ*, a collective variable ranging over both fingers. The behavior of the system can be predicted and described by the following simple equation:

where *ϕ̇* is the first derivative of this relative phase, and *α* and *β* are fit parameters whose ratio *β*/*α* is a global control parameter controlling the effect of the relative phase. Roughly, this equation predicts that for values of the global control parameter *β*/*α* > 0.25, *ϕ* values of 0 and *± π* are both stable attractors, whereas for lower values of *β*/*α*, only 0 is stable (see Figure 1). These relative phase values *ϕ* correspond to in-phase (0) or out-of-phase (± *π*) stable states observed in bimanual coordination.

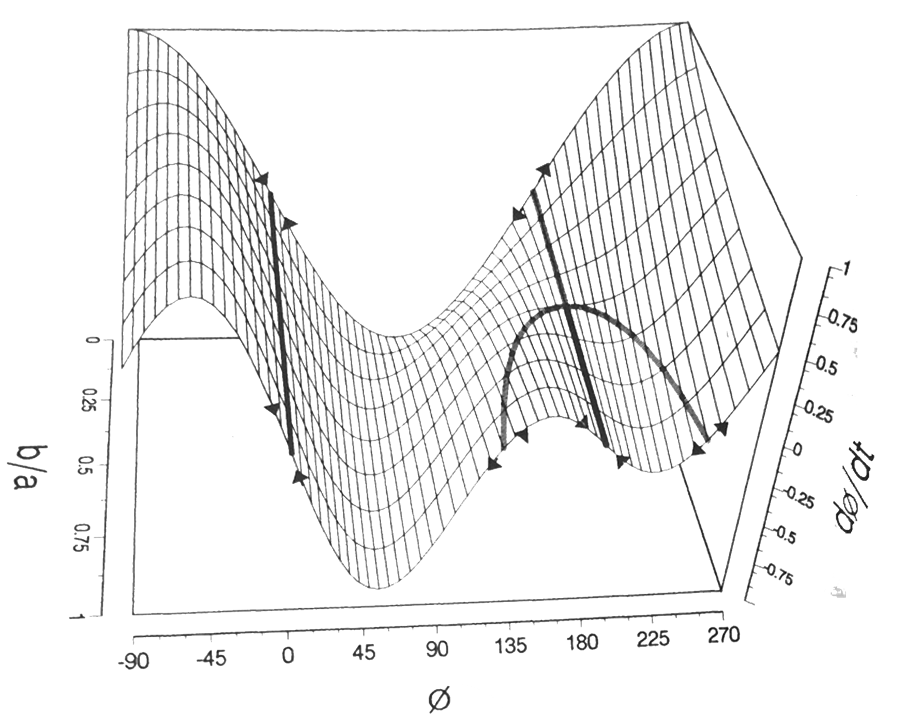


Figure 1: Phase diagram of the influence of the global parameter (b/a) values over the values of the relative phase (ϕ), the collective variable for the system. The curved manifold represents how the global parameter and relative phase influence the change in the phase (dϕ/dt). The arrows on the depicted manifold correspond to the direction in the state space that the system tends to move. The thick black lines are the sets of points at which dϕ/dt = 0, called fixed points. If the arrows point toward the black lines, then the points are attractors; if the arrows point away, then the points are repellors. Adapted from Kelso 1995, p. 57.

The HKB explanation of bimanual coordination focuses on the fingers and not the parts, accounts for their behavior in terms of collective variables and global parameters, and invokes coupling between the fingers.[[10]](#footnote-11) On the HKB model, the description of the interaction of the two fingers (equivalent to the focus on brain, body, and environment for cognitive systems writ large) explains bimanual coordination. The parts of the subsystems—that is, the parts of the fingers—are absent from the explanation. In addition, the system is mathematically described utilizing collective variables, specifically the relative phase of the two fingers. This relative phase tends towards certain values, the attractor points in the phase space. The shape of this phase space is governed by global control parameters, specifically the ratio of the phase weights, *β*/*α*. Finally, the explanation relies on the fact that the two fingers are coupled, which permits the use of a collective variable for the relative phase *ϕ* to denote the system state.

* 1. Componential vs. Systemic Dynamicism

Dynamical systems can be more or less systemic, and more systemic dynamical systems are contrasted with less systemic, or componential, dynamical systems. The behavior of a componential dynamical system is explained in terms of its subsystems and parts[[11]](#footnote-12), the subsubsystems of the system, whose functions together account for the capacities of the larger containing system (cf. Cummins 1975).[[12]](#footnote-13) Systemic dynamical systems, in contrast, are not explained in terms of their parts. In addition, componential dynamical systems are described using individual instead of collective variables and local instead of global control parameters. Finally, coupling is an optional characteristic of componential dynamical systems, unlike systemic ones.

To illustrate, consider a car. A car can be described as a systemic dynamical system, including such properties as the car’s speed, position, fuel levels and others. The state of the car’s parts, such as the axles, fuel valve, or carburetor, and the functions internal to the car, such as the injection of fuel, the ignition of the fuel, or the transfer of energy from the pistons to the driveshaft, are elided in such a systemic description. This systemic dynamical system also includes environmental properties, such as the weather, distance to the nearest gas station, how many other cars are on the road, and others. Furthermore, the influence of these environmental properties on the car is described by their effect on the car as a whole. The specific details of how increases in traffic result in increased braking and decreased speed are foregone in favor of a general description that relates the car’s velocity to the local traffic density. The more general systemic description will capture not just how the environment influences the car and its properties, but also how the car influences the environment. The presence of the car on the road contributes to the amount of traffic; its exhaust contributes to the atmosphere; and so on.

These properties can be described using collective variables. Consider a system composed of two cars. The inter-car distance can be used to describe these cars, where the inter-car distance is a collective variable that stands for the distance between the cars. The inter-car distance will vary as a function of velocity and location as well as of the local density of cars on the road. A property like the amount of traffic acts as a global control parameter that controls the collective, inter-car distance variable. Other collective variables and global control parameters can be imagined.

Some of the interactions between properties will be coupled: the state of the car partly determines the state of the environment and vice versa. This coupling can be captured by states of the car and environment that mutually determine one another and that are described using variables that appear in their state equations. Similarly, two cars may be coupled, where their distances and velocity are mutually determining as well, such as in the inter-car distance case above.

A car can also be described as a componential dynamical system. The parts inside the car are represented in the componential dynamical description and the change in these parts is reflected in the equations describing the evolution of those parts. Such a description contains variables that stand for the various parts, such as the engine, the steering mechanism, the wheels and other parts. Specific environmental influences, typically restricted to interfaces between the car and the environment, may also be included. Such a description does not consist primarily in describing the subsystems of the car-environment system. The car’s behavior is not described in terms of the interaction of the car with other cars and the environment, but rather in terms of the subsubsystems in the car. Hence, this componential description does not contain collective variables or global control parameters, which would require specifying other cars or properties of the environment. Finally, coupling between parts may occur, but the descriptions of the parts and their evolution usually does not necessitate coupling. Coupling between the car and the environment is also left out altogether.

The HKB model of sensorimotor coordination described above is a good example of a systemic dynamical system. However, the case of sensorimotor coordination can also be described as a componential dynamical system. A componential dynamical system description would include the peripheral or central nervous system, the musculature, and other tissues. The states of the fingers, such as their location and velocity, are determined by the states of these parts.[[13]](#footnote-14)

In addition to sanctioning parts of dynamical systems, componential dynamical systems are not described using collective variables and global parameters. Unlike the main form of the HKB model above, in the componential dynamical system, the variables and parameters governing the parts of the fingers determine each finger’s position, velocity, and acceleration. These individual variables refer to specific properties of each finger. Individual control parameters govern how each variable changes. These individual variables and local control parameters govern the amplitude and frequency of waggling of each finger.

Finally, coupling is optional in componential dynamical systems, though it is present for the case of sensorimotor coordination. Thus, the two fingers are coupled in virtue of the parts of the first finger, which determine the position of that finger, being determined by the position of the second finger and vice versa. On this redescription, the explanation of the individual fingers’ movements is a function of a global control parameter only derivatively, in virtue of how the parts of the fingers determine the fingers’ motions. This reflects the effect of coupling: when systems are coupled, their behavior mutually influences each other, such that a change in even a local control parameter will affect the state of the other system.

In sum, systemic dynamical systems are those that are described at the level of the whole system; focus on the interaction of the subsystems including the environment; forego partly or entirely the parts of the subsystems; and that require coupling between the subsystems. In contrast, componential dynamical systems are those that are described in terms of subsystems and parts; describe the system’s overall capacities in terms of the functions of its parts; focus on the interaction of the parts within the subsystems; often forego mention of the whole surrounding environment and potentially system-wide variables and effects; and that often lack coupling either with the environment or between large subsets of parts in describing the system.[[14]](#footnote-15) Componential dynamical systems’ focus on parts and local interactions fits well with computationalism.

1. Cognitive Neurobiology and Neurodynamics

I will now argue that cognitive neurobiology prominently features neurodynamical systems in explanations of cognitive phenomena. In this section, I first define neurodynamical systems and next illustrate the claim that cognitive systems contain neurodynamical systems with an in-depth discussion of the example from perceptual neuroscience. I then discuss the role of functions in these explanations. The discussion of functions is needed in order to understand how cognitive neurobiology constructs explanations for cognitive phenomena. I argue that neurodynamical systems execute the functions that result from an analysis of cognitive capacities. I distinguish between two senses of function, compositional functions and functional role functions, and describe how the ascription of a functional role function to a dynamical system requires mapping a functional role function on to a compositional function of a dynamical system. In the next section, I present the explanatory schema and discuss its scope. In the penultimate section, I turn to an examination of the implications of this explanatory framework, and argue that the presence of neurodynamical systems in cognitive systems implies that cognitive systems are componential dynamical systems.

* 1. Neurodynamical Systems

Neurophysiological systems have many different types of properties and relations, including spatial, temporal, and other kinds, as well as many different types of objects, including neurons, ion channels, and so forth. These systems also feature dynamical properties, properties of how these systems change. These dynamics are present across spatial scales, ranging from quantum effects related to synaptic vesicle fusion and conformational changes in the receptor proteins that regulate the flux of ions across the neuronal membrane to whole organismic effects from interacting with other members of the species or the environment. These dynamics are also present across temporal scales, from very fast sub-millisecond changes in ion channels to multi-year changes in the anatomical pathways in the brain.

Organized sets of dynamical properties of neurophysiological systems are neurodynamical systems. Since these dynamical properties are defined in terms of changes in objects, properties, or relations of neurophysiological systems, neurodynamical systems are token identical to subsets of dynamical properties of neurophysiological systems.[[15]](#footnote-16) A neurodynamical property can be a change in some property of a neurophysiological system, such as a change in neuronal membrane voltage (the changes that underlie neuronal firing rates), or it can be a change in such changes, such as changes in firing rates. Hence, in the brain, dynamical systems are instantiated[[16]](#footnote-17) by neurophysiological systems, in that the properties of neurodynamical systems are changes in neurophysiological properties or changes in the changes in such properties. But not just any collection of dynamical properties will do; neuroscientists pick out and describe organized sets of dynamical properties. These organized sets are sets of ordered n-tuples of dynamical properties, those properties that stand in particular relations to each other. In addition, as elaborated below, these dynamical systems perform functions for cognitive systems. Neurodynamical systems are organized sets of dynamical properties of neurophysiological systems that perform functions for cognitive systems. The discussion that follows will utilize this notion of neurodynamical system.

* 1. A Case Study from Cognitive Neurobiology: Perceptual Decisions in Noisy Conditions

To illustrate and motivate singling out neurodynamical systems, I will review in greater depth the case study from cognitive neurobiology presented in the introduction. The case study is of perceptual decisions under noisy sensory conditions, such as the monkey in the jungle being stalked by the leopard, who must decide if that dappled pattern in the foliage is a predator.[[17]](#footnote-18) Neuroscientists study these decisions using the random dot motion task (RDMT), in which subjects are presented with a visual display of moving dots, only some fraction of which move in the same direction (motion coherence). Different fractions of dots move coherently on different trials. To make a decision, the subject looks at a target as determined by the perceived motion of the dots, receiving a juice reward for a correct decision. How do cognitive neurobiologists explain this decision making capacity?

In primates, motion properties of visual stimuli are encoded in area V5/MT, an area in the occipital cortex of the brain (Zeki 1974; Zeki 1991; Britten, Shadlen et al. 1992). The lateral intraparietal area (LIP), an eye movement control region in the parietal cortex (Platt and Glimcher 1999), receives motion information from area MT (Britten, Shadlen et al. 1992; Britten, Shadlen et al. 1993). LIP then integrates this information over time to contribute to the decision about the direction of motion of the stimulus (Figure 2). The study of LIP neuronal responses during the RDMT has revealed the presence of a neurodynamical system, the integrate-to-bound system, for integrating this motion evidence (Roitman and Shadlen 2002; Gold and Shadlen 2007).

The integrate-to-bound dynamical system starts at some initial state, smoothly transitions through a series of adjacent states, and then terminates at a particular point that is identical across different initial states and trajectories. The system continuously changes state as a smooth, non-saltatory function[[18]](#footnote-19) of changes in environmental or internal properties until a particular boundary is reached. For example, in the RDMT, different clouds of dots may have different proportions of coherent motion (dots moving in the same direction). For motion in a neuron’s preferred direction, the neuron will integrate this motion evidence in proportion to its coherence. This is reflected in different trajectories through the state space of the integrate-to-bound system. Further, different clouds of dots (that is, different patterns of sensory input) may have the same coherence; the integrate-to-bound system may follow approximately the same trajectory in those cases. Regardless, these different trajectories all lead to the same attractor point in the system’s state space. From this terminal state, the system can reset and begin the integrative activity anew. This integrate-to-bound pattern of dynamics can be described in part using the concepts and tools of dynamical systems theory. This system is a one-dimensional attractor with a single stable fixed point: the system under input is drawn toward one particular point in its state space (Strogatz 2001). The system’s change in state is an integration, a trajectory through the system’s state-space toward an attractor point that correlates with some variable. In sum, an integrate-to-bound dynamical system is a dynamical system that starts in an initial state, smoothly and continuously transitions through a series of adjacent states as a function of some other property, arrives at a threshold state, and then resets.

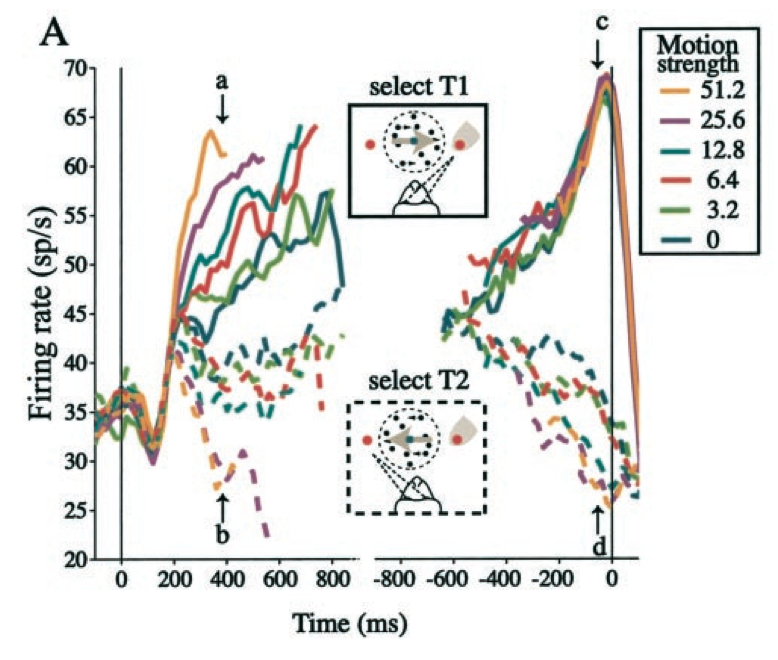


Figure 2: Average recorded firing rates from cells in the parietal cortex of monkeys while they make a perceptual decision in different evidential conditions. In the left plot, average firing rate of the recorded neuronal population is on the y-axis, and time is on the x-axis; time ‘0’ is time of onset of stimulus. A larger proportion of dots moving in the same direction (“Motion strength”) results in a steeper increase in activity of these neurons (left plot). In the right plot, average firing rate is also plotted against time. The increase in firing crescendos at a common boundary (labeled ‘c’) for different motion strengths, just prior to making a movement (time ‘0’, right plot). Adapted from Roitman and Shadlen 2002, p. 9482.

* 1. Functional Neurodynamics

Neurodynamical systems perform functions for cognitive systems. Understanding the sense in which dynamical systems perform certain functions for cognition requires a distinction between different notions of function (for related discussions of function, see Clark 1980; Young, Hilgetag et al. 2000; Bergeron 2007; Anderson 2010). A full discussion of the different senses of function is out of place for the present, but a sketch of the intended notions is needed to fully specify my account of componential dynamicism.[[19]](#footnote-20)

First, there is a compositional sense of function. Elements of systems have many properties and stand in many relations that can help to make up systems. Different types of these compositional properties or relations include causal (the set of causes and effects of the elements of the system) and constitutive (the contribution of some element to the ordered set of objects, properties, and relations that are token identical to a system at a time or over some interval).[[20]](#footnote-21) Elements and their properties and relations can simultaneously contribute to many different systems. A compositional function of an element is a description of its contribution to some system. The system is the set of relations, transformations, and causes and effects of the elements that constitute the system at a time or over a temporal interval. Once a particular system is chosen for analysis (such as the mind, the brain, or the body), the elements of the system have compositional functions with respect to that system, or (equivalently for my purposes) the element performs a compositional function for the system (cf. Clark 1980; Neander 2017). Call these c-functions.

A c-function of an element includes the set of relations in which it stands to other elements. A brick in a house stands in numerous relations to other elements of the house, such as laying under some other brick, being next to a third brick, and so forth. Unlike the brick, other elements of the house may undergo transformations. A door, for example, undergoes characteristic transformations such as opening and closing. Finally, elements can have causal effects in the system; for example, the door’s opening or closing may result in changes in air pressure or in access to the house. These changes are also part of the whole that is the house, and an element’s c-functions include the set of transformations that are included in the description of the whole. In sum, a c-function of an element is the set of relations in which the element stands, transformations which the element undergoes, and the element’s causes and effects within the system. Hence, the dynamics of elements are included in the set of c-functions.[[21]](#footnote-22)

To illustrate c-functions, consider the heart. The heart has many compositional properties including a range of causal properties. The heart helps to compose the body in that it contributes many compositional properties, including causes, to the body. For example, the contraction of the heart causes blood from the right ventricle to enter the pulmonary trunk and from there to travel to the lungs. So, one of the heart’s c-functions is to pump for the body.

Second, there is a functional role sense of function. Systems have cognitive capacities that are often analyzed into a collection of functions, with the functions mapped on to the various elements of the system (cf. Bechtel and Richardson 2010). The execution of each such function in the collection just is the execution of the capacity. Call these f-functions.

F-functions are closely related to Cummins’ causal role functions. Cummins defines the causal role function ϕ of an entity x “…relative to an analytical account A of s's capacity to ψ just in case x is capable of ϕ-ing in s and A appropriately and adequately accounts for s's capacity to ψ by, in part, appealing to the capacity of x to ϕ in s” (Cummins 1975, p. 762). However, unlike Cummins’ analysis, f-functions do not make a claim regarding some components (Cummins’ entity x). Instead, in Cummins’ terms, f-functions are the subcapacities that result from an analysis of some capacity of a system.

Consider the case of physiological respiration. The body has the capacity to respire. This capacity is the explanandum capacity. The capacity to respire is analyzed into a set of subcapacities, including the subcapacity to circulate metabolic resources such as oxygenated blood. Those subcapacities are the f-functions. The body has the capacity to respire partly as the result of the f-function of the circulation of oxygenated blood.

C-functions and f-functions are connected. The element that fills an f-function is token identical to some element that fills some c-functions. Furthermore, in virtue of filling some c-function and relative to some capacity analysis, that element fills an f-function. First, the constitutive, dynamic, or causal properties that underlie the element’s c-function are the same as underlie the element’s f-function. In the heart example, the causes that underlie the c-function, the pumping, are identical to the causes that underlie the f-function, the circulation of blood, for that particular case. There are other causes, such as those that result from an artificial heart, that could underlie circulation of the blood in counterfactual cases. Second, the heart fills some f-function for physiological respiration because it fills some c-function. In circulating the blood, the contraction of the heart forces blood to the lungs for the exchange of oxygen and carbon dioxide, essential for physiological respiration. Similarly, the heart’s contractions force blood throughout the body for systemic circulation of oxygenated and de-oxygenated blood. The heart’s pumping, the series of contractions and relaxations, are the relevant causes for the body to explain this circulation. The heart fills the role of circulating the blood (an f-function) for the body by the causal contribution of pumping (a c-function) for the body. More generally, the capacity of the system is explained and described by the f-functions of the system’s elements, and those f-functions are filled because of the elements’ c-functions.

In order to map f-functions on to c-functions in cognitive neurobiology, both c- and f-functions are mathematically described. The formalization of the c-function mathematically describes the constitutive, transformational, and causal properties of the elements corresponding to the c-function. Call this mathematical description of the c-function, the mc-function. Similarly, the formalization of the f-function mathematically describes the contribution of the subcapacity to the overall capacity of the system. Call this mathematical description of the f-function, the mf-function. The mf-function is mapped on to the mc-function.

The perceptual decision-making case study illustrates this somewhat complex mathematized mapping. This mapping is not a perfect isomorphism and need only be approximate. Consider again the above problem of determining the direction of motion of a field of moving dots. Detecting the direction of motion of a noisy visual stimulus involves a number of f-functions including evidence integration, a summary of the evidence for or against particular hypotheses about motion direction. A mathematical model[[22]](#footnote-23) describes an evidence sampling and summarizing process for helping to determine the direction of motion during the RDMT. The first step in the model is to set the prior odds of the dots moving left or right. Then, by sampling the motion signal over some small period of time, evidence is gathered from the field of moving dots (Shadlen and Newsome 2001). This evidence is used to update the odds that the dots are moving left or right as summarized by a log-odds ratio. The process iterates until some evidential threshold is reached.[[23]](#footnote-24) Once a threshold is reached, a decision is made.

A connection between the mathematical descriptions supports the conclusion that the integrate-to-bound dynamical system in LIP performs the evidence integration function for the cognitive system. The mathematical operation underlying the integration has been framed as an integration (*sensu* calculus), an exponentiation, or a mixture of different functions (Usher and McClelland 2001; Wang 2002; Mazurek, Roitman et al. 2003; Carandini and Heeger 2012).[[24]](#footnote-25) Regardless of the operation used in specifying the state equation, the state equations respect the dynamical systems theoretic description of the integrate-to-bound system as a one-dimensional attractor with continuous, integrating dynamics. Specifically, these state equations feature a baseline starting point, an increase in value, and some threshold. The mf-function is mapped on to this mathematical description. In the case of detecting motion direction for a noisy stimulus, the mf-function is the mathematical model described above. First, these cells have a ‘reset’ point or baseline firing rate just after motion onset which is the starting point of the integration, just as the model starts with a prior log-odds ratio that is the same across different strengths of evidence (Figure 2, left panel, the dip in firing around 100 ms). Second, the activity of cells in LIP varies with the motion strength of the stimulus with steeper increases in firing rate for stronger motion, described as differences in the trajectory through the system’s state space, just as the model dictates a sequential evidence integration process (Figure 2, left panel, the different integration trajectories coded by color under the label ‘a’). Third, the activity of these cells converges on a common firing rate across different motion conditions just prior to a decision, described as a single fixed point, just as the model’s evidence sampling process halts when a threshold is reached, the same across different strengths of evidence (Figure 2, right panel, point labeled ‘c’). Each of these qualitative properties correspond to quantitative properties of the different mathematical models. In sum, the mathematical description of the evidence integration function maps on to the mathematical description of the integrate-to-bound dynamics for making noisy perceptual decisions.

1. Explanations of Neurocognitive Phenomena

The above considerations motivate the construction of a schema for neurodynamical explanations of cognitive phenomena. To begin, I outline this schema using the case study of perceptual decisions as illustration. Then, I discuss the scope of this explanatory schema by describing four other case studies from recent cognitive neurobiology. This will set the stage for the following arguments that cognitive systems are componential dynamical ones.

* 1. A Novel Explanatory Schema

Explanations of neurocognitive phenomena contain a series of steps.[[25]](#footnote-26) The explanandum is the cognitive capacity of the system. The explanation for these capacities involves describing a set of subcapacities and parts of the system that possess those subcapacities. The cognitive system is analyzed into subsystems (like the brain) and subsubsystems (like the brain’s regions), the parts that instantiate dynamical systems. These dynamical systems perform c-functions for the cognitive system, and these c-functions are then described mathematically as mc-functions. The cognitive capacity is also analyzed into a set of subcapacities, some of the f-functions of the cognitive system. These are described mathematically as mf-functions. For each such f-function, some part performs the f-function for the cognitive system. This part is identified by mapping the mf-function on to an mc-function. This mapping grounds the inference that the dynamical system has a particular f-function for the cognitive system.[[26]](#footnote-27) This is repeated for each such f-function into which the explanandum cognitive capacity was analyzed. The resulting set of functional neurodynamical systems is then taken to compose the subsystem of the cognitive system with the explanandum capacity. (See Figure 3 for a depiction of the explanatory structure.)

Capacity C

Subsubsystems

(c-functions)

Dynamical System 1

Dynamical System 2

.

.

.

Dynamical System N

Subcapacities

(f-functions)

mc-functions

mc-function 1

mc-function 2

.

.

.

mc-function N

mf-functions

Subsystem S

mf-function

A

mf-function B

.

.

.

mf-function Z

Cognitive System

f-function

A

f-function

B

.

.

.

f-function

Z

Figure 3: Schematic of the structure of explanations in cognitive neurobiology. Double arrows are descriptions, single arrows are on to mappings. The cognitive system is divided into a set of subsystems such as the brain (only one depicted here, subsystem S). These subsystems are divided into subsubsystems, dynamical systems that are ascribed c-functions. These c-functions are mathematically described by mc-functions. The cognitive system has a number of capacities (only one depicted here, capacity C), each of which is analyzed into a set of subcapacities, the f-functions. These f-functions are mathematically described by mf-functions. The mf-functions are then mapped on to the mc-functions.

This schema is well-illustrated by the explanation of the capacity for perceptual decision making under noisy conditions. First, the cognitive capacity to discriminate motion is decomposed into a set of f-functions, including the integration of visual evidence over time. Second, the relevant f-function for motion discrimination for the explanation is specified; in this case, the f-function is evidence integration. Third, this f-function is described mathematically, yielding an mf-function that uses motion evidence to update the odds of competing hypotheses about the direction of motion. Fourth, the integrate-to-bound dynamics, the c-function, of the physical subsubsystem are specified, an integrating trajectory through the subsubsystem’s state space towards an attractor point.[[27]](#footnote-28) These dynamics are a c-function because they partly constitute the total dynamics of the brain. Fifth, those dynamics are described mathematically, which in the case of the integrate-to-bound system includes specifying a state equation, the mc-function. Sixth, the mf-function is mapped on to the mc-function. The mapping of the mf-function on to the mc-function grounds the inference that the dynamical system executes some f-function for the system. This structure for explanations of cognitive phenomena generalizes beyond perceptual decision making.

This explanatory schema supports an inference about what subsystems of the brain (or other parts of the cognitive system) have subcapacities in the cognitive system. But the schema is not deductive. The different steps provide defeasible support for the conclusion that some part performs a function for the cognitive capacity being explained. Given a subcapacity of some cognitive capacity and its mathematical description, some dynamics of a subsubsystem and its mathematical description, and a mapping from the subcapacity on to the dynamics, we defeasibly infer that the part possess the subcapacity. The support is defeasible because the steps could all be true and the conclusion, that the part fills some f-function, false.

* 1. Scope of the Schema

A set of objections regards the scope of the explanatory schema. There are three different versions of the scope objection. The first version objects to the application of the explanatory schema to simpler cognitive processes such as sensation or perception.[[28]](#footnote-29) The second version objects to the application of the explanatory schema to other cognitive processes besides perceptual decision making. The third version objects to the application of the explanatory schema to other dynamical systems besides the integrate-to-bound system. As I illustrate below, all three scope objections underestimate the schema.

On the first scope objection, simpler cognitive processes such as the encoding of sensory features like motion properties of visual stimuli do not seem to utilize dynamical systems in the way that perceptual decisions based on those stimuli might. In reply, the explanatory schema outlined above holds even for understanding the functional role of individual neurons such as the responses in area MT that encode motion properties of visual stimuli. Those neurons selectively respond by increasing their firing rates to stimuli in their visual receptive field, a circumscribed area in the visual field that drives activity in those neurons. Sensory encoding, such as encoding of motion, refers to functions that involve transformations of magnitudes such as speed or direction. The set of f-functions include the indication of motion speed and direction of stimuli. To demonstrate that area MT encodes motion, both of these specific f-functions would need to hold of the neurodynamics of that area. Speed and direction encoding are typically mathematically described using Gaussian tuning curves (the mf-function) (Britten and Newsome 1998).[[29]](#footnote-30) These properties are taken to be encoded by the dynamics of neurons, such as changes in firing rates (the c-functions). The firing rates of individual neurons exhibit a variety of linear and non-linear effects and are often modeled using a linear-non-linear leaky integrate-and-fire model (Paninski, Pillow et al. 2004) or a generalized linear model using a log-linear link function (Aljadeff, Lansdell et al. 2016) (the mc-function). MT cells signal the speed and direction of motion stimuli: they respond more to movement in a particular direction (Zeki 1974) and at a particular speed (Maunsell and Van Essen 1983). The inference from the dynamics of MT neuronal responses to their functional role in signaling motion speed and direction is partly justified by the mapping of the Gaussian tuning curve on to the firing rate model (that includes the variables to which the neuron is tuned).[[30]](#footnote-31) The general idea here is that while neurons may be sensitive to a range of variables for a number of reasons, in order to justify the claim that their function is to encode a particular property, as though they were a sensor constructed to detect that property, their dynamic response to changes in that property must exhibit a certain organization. This organization is succinctly captured by tuning curves, and the conclusion that the neuron signals the property is the result of successfully mapping a tuning curve model for that property on to the model of the neuron’s dynamics.

While I contend that the model can be extended to sensory encoding, the first case study of perceptual decisions may appear too parochial, limiting the scope of the schema to perception and sensation. On the second version of the scope objection, the objector disputes an extension of the schema to other cognitive phenomena. In reply, the schema can be extended to other types of decision making and even other cognitive phenomena.

An extension of the same integrate-to-bound dynamical system to strategic decision making helps explain a second cognitive phenomenon. When foraging in an environment where rewards are clustered in patches and deplete as they are harvested, animals face the patch leaving problem: determining when to leave the current depleting patch to travel to a new one (Stephens and Krebs 1986). Neuroscientists who study these decisions have used the patch leaving task (Hayden, Pearson et al. 2011; Barack, Chang et al. 2017), where subjects decide whether to continue foraging in a simulated patch or to leave the patch and incur a time-out penalty to mimic the travel time to a new patch. Patch leaving decisions require a number of functions, such as keeping track of the value of the current option, the value of switching, and trading off these values. These decisions are partly mathematically described as tracking average reward rates to set decision thresholds (Charnov 1976; Stephens and Krebs 1986; Kacelnik, Vasconcelos et al. 2011; Barack and Platt 2017). Neuronal recordings from the anterior cingulate cortex (ACC), a medial prefrontal cortical structure, revealed all three elements of an integrate-to-bound system and hence a different neurophysiological instantiation of it. The longer monkeys foraged in a patch, the greater the peak response in ACC neurons around the time of a decision, akin to an integration. For similar travel times to a new patch, the firing rates in those neurons also rose to a common threshold for different patch leave times. Finally, the initial firing rates at the beginning of the patch were the same. This dynamical system can be mathematically described,[[31]](#footnote-32) and some aspect of mathematical models of foraging decision thresholds can then map on to this description of the dynamics (see, e.g., Barack and Platt 2017).

This is an active area of research, and the science is far from settled. On some accounts, changes in the relative value of foraging elsewhere (Kolling, Behrens et al. 2012; Kolling, Wittmann et al. 2016), one f-function, map on to the integrate-to-bound dynamics, while on others, changes in decision difficulty, a different f-function which enters into a cost-benefit computation to determine whether to stay or leave the patch (Shenhav, Straccia et al. 2014), map on to the dynamics. Note that filling in the step in the schema that identifies precisely which aspect of the function for the cognitive capacity maps on to which aspect of the dynamics explains why certain experiments are designed and undertaken in cognitive neurobiology, as illustrated by this debate over activity in the ACC during foraging. I take this as a mark in favor of my approach, in virtue of explaining why the dynamics of ACC activity continues to be of relevance to cognitive neurobiologists.

This case illustrates that the dynamical system in the initial case study also appears in the explanation and description of foraging decision making. However, a third version of the scope objection maintains that the explanatory schema is limited to only a particular dynamical system, the integrate-to-bound system. In reply, the scope of the explanatory schema is not limited to only integrate-to-bound systems. To bolster the case that a range of dynamical systems play a role in explanations of cognitive phenomena, I turn now to a second dynamical system, the divisive normalization system, and its role in attention and reward encoding in the brain.[[32]](#footnote-33)

Neuronal activity in sensory areas, such as visual areas V1 or MT, often exhibits nonlinear characteristics when coding environmental variables (Andersen and Mountcastle 1983; Heeger 1992; McAdams and Maunsell 1999; Treue and Martinez Trujillo 1999; Heuer and Britten 2002). Divisive normalization (DN) has been proposed to account for these effects (Heeger 1992; Carandini, Heeger et al. 1997; Cavanaugh, Bair et al. 2002). DN describes the activity of a particular neuron as driven by the stimulus in the neuron’s receptive field and depressed by the pooled responses across a population of nearby neurons.

To illustrate, consider the DN model of attention (Reynolds and Heeger 2009).[[33]](#footnote-34) Attention comes in many guises, but in one form, attention is selective processing, famously characterized by William James as “taking possession by the mind, in clear and vivid form” the object of attention (James 1890/1950, p. 403). In cognitive neurobiology, the functions related to the selection of the object of attention include a relative increase or decrease in sensory or perceptual signal strength (the relevant f-function). This is often mathematically described as some type of gain modulation (additive or, more often, multiplicative) of the signal (the mf-function) (Reynolds and Heeger 2009).

The DN model of attention is a mathematical description of observed neural dynamics while attending to a stimulus. Under the influence of attention, neural responses to the attended stimulus increase or decrease relative to the unattended stimulus (the c-function). The DN model of attention describes these dynamics by invoking the interaction of three different influences on neural activity: sensory stimulation, neuronal suppression, and attention (Figure 3). DN is usually utilized to mathematically model the responses from individual neurons and describes how a neuron’s firing rate will change in response to changes in its stimulus drive, the activity elicited in a neuron by a stimulus, or to changes in the suppressive effect from other neurons. The sensory stimulation characterizes the preferred spatial position and orientation of a stimulus for driving a neuron’s responses. The suppression characterizes the spatial positions and other features contributing to a suppression of a neuron’s response as a result of the activity of other neurons responding to stimuli. Attention acts as a gain for the effect of the stimulus on each neuron in the population, and increases the stimulus drive before the suppressive action by nearby neurons. This normalization is described in terms of a division of the stimulus drive by the summed activity of neurons that suppresses that drive (the mc-function).[[34]](#footnote-35) Just as in the first case study, the inference that neurons perform attentional functions for cognitive systems is grounded in mapping the mathematical description of the function (an increase or decrease in sensory signals) on to the mathematical description of the dynamical system (DN).

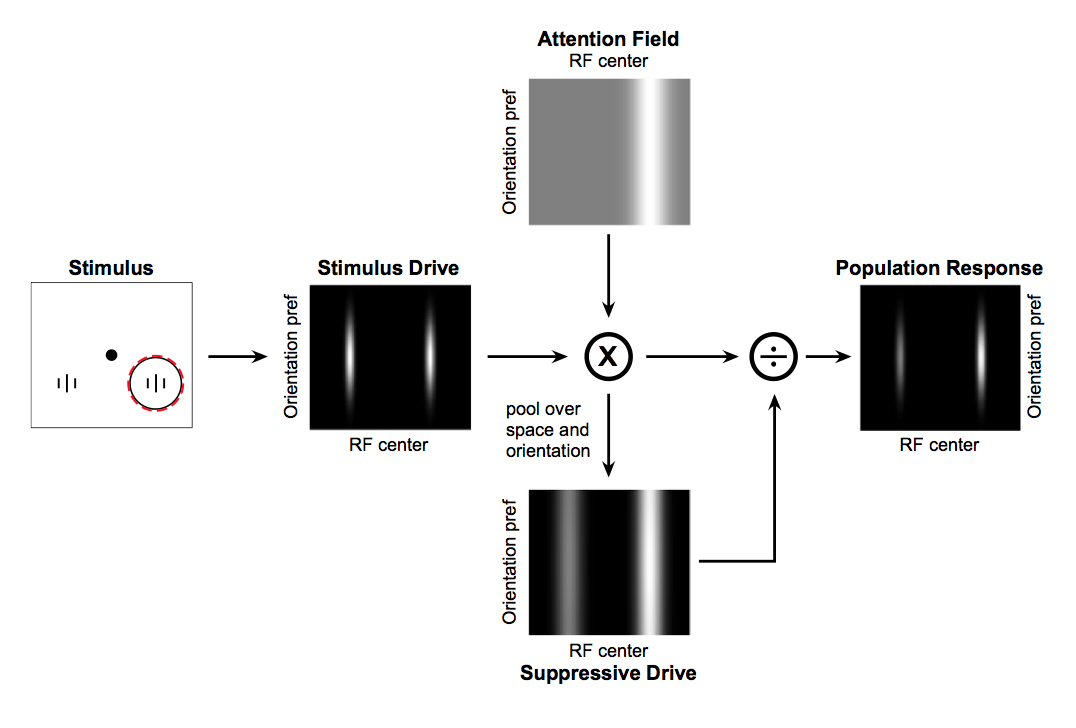


Figure 4: The divisive normalization model of attention (adapted from Reynolds and Heeger 2009, p. 169). The left-most panel displays the stimulus presented to an animal on a computer screen in a typical attention task, in this case, two vertical oriented gratings (the two sets of the three dashed lines). The dashed circle in that panel is the neuron’s receptive field (RF). The other four panels reflect the strength of the factors influencing neural activity or the neural activity itself as a heatmap. This schematic displays the orientation preferences for the RF centers in only one dimension for simplicity. The ‘Stimulus Drive’ panel shows the strength of the influence of the stimulus on neural activity. ‘Orientation pref’ is the preferred stimulus orientation for neurons (e.g., horizontal or vertical) for the possible stimulus gratings. ‘RF center’ refers to the different possible centers of the receptive fields of neurons in the population. Brighter colors correspond to larger modulation. The ‘Attention Field’ panel depicts the influence of attention for different stimulus orientations at different RF center locations. The attention field multiplicatively scales the stimulus drive of a neuron for a particular orientation. The ‘Suppressive Drive’ panel indicates how the attention field and the stimulus drive influence each neuron in the population. The product of the attention field and the stimulus drive is divided by this suppression to yield the response from the neural population, depicted in the ‘Population Response’ panel. For the depicted display with two stimuli, even though the stimuli have the same orientation, the effect of attention is to enhance the responses for neurons at a particular location on the screen while depressing the responses of neurons elsewhere.

While the system’s use for cognitive processes is well-illustrated by its application to attention, divisive normalization might be tailored specifically for attentional phenomena. In reply to this objection, divisive normalization has also been recently used to help explain reward-based decision making. Such decisions involve a host of f-functions including context-based reward encoding (the relevant f-function), the encoding of reward sizes relative to a particular environment or cognitive task. This reward encoding is typically mathematically described using either a tuning curve for reward size or as a set of ordered responses to rewards (the mf-function). Louie and colleagues investigated reward encoding in the lateral intraparietal area (LIP) by varying the number of targets and their associated reward sizes in a cued saccade paradigm (Louie, Grattan et al. 2011). Presented with an array of one to three targets in one of seven different spatial configurations, monkeys made a saccade to an instructed target. In order to assess firing rates under different reward conditions, the reward amounts for the targets inside and outside the receptive field (RF) of individual neurons was varied systematically across blocks. The total reward modulates LIP responses, with LIP neurons’ firing inversely correlated with total reward, consistent with DN dynamics (the c-function). A divisive normalization model, that computes the firing rate of those neurons as a function of the ratio of the rewards inside a neuron’s RF to those outside, best described the neural data (the mc-function) (Louie, Grattan et al. 2011). Hence, LIP neurons contextually encode rewards related to targets in the environment. Once again, a mapping between a mathematical description of an f-function and a mathematical description of a dynamical system justifies the inference that the dynamical system fills a function for the cognitive system.

**Table 1.** Summary of the case studies discussed herein. See text for more details.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Example | c-function | mc-function | f-function | mf-function |
| Perceptual decision making under noise | Integrate-to-bound dynamics | Mathematical integration or exponentiation | Evidence integration | Sequential sampling model |
| Encoding of motion stimuli | Single-neuron response dynamics | Generalized linear models or linear-non-linear models | Signaling motion speed or motion direction | Tuning curves |
| Foraging decision making | Integrate-to-bound dynamics | Mathematical integration | Relative value coding or decision difficulty | Subtraction between the values of options or probabilistic computation over differences in values |
| Attention | Divisive normalization | Mathematical division | Increased signaling of stimuli | Gain modulation |
| Value encoding | Divisive normalization | Mathematical division | Value signaling | Tuning curves or sets of ordered responses |

These various examples are summarized in table 1. I contend that, in addition to the integrate-to-bound and divisive normalization systems, there are a range of dynamical systems being used for different cognitive capacities. These include synaptic reverberation (Wang 2002; Wang 2008; Strait, Blanchard et al. 2014), center-surround inhibition (Mink 1996; Hikosaka, Takikawa et al. 2000; Mink 2003; Kiyonaga and Egner 2016), and others. The discovery of such a suite of dynamical systems filling functions for cognitive systems is an important development in the neuroscientific explanation of cognitive phenomena.

1. Against Systemic Dynamicism

Above, I outlined an explanatory schema that centrally featured dynamical systems. I illustrated that schema with examples drawn from research on the neurodynamics underlying a range of cognitive phenomena. In this section, I present three arguments that cognitive systems are componential dynamical systems. First, neurodynamical systems are componential dynamical systems and when different such systems are combined to explain cognitive capacities, their componential features don’t change. This suggests that cognitive systems are componential dynamical systems. Second, the methodology of cognitive neurobiology suggests attempts to manipulate local, individual dynamics, not systemic ones. Finally third, the nature of scientific debates in cognitive neurobiology supports a componential interpretation of cognition. Unlike past dynamicist approaches (such as Port and van Gelder 1995; van Gelder 1995; Chemero 2011) and others), then, the present account implies cognitive systems are componential dynamical systems.

First, the models used in cognitive neurobiology lack the properties of systemic dynamical systems models. The models used to mathematically describe neurodynamical systems lack collective variables and global control parameters. The integrate-to-bound system in LIP can be described without a role for systemic properties in governing the evolution of the system, and so can be described independently of the state of the body or environment. Collective variables are usually not present in the models describing the neurodynamical systems posited to explain cognitive phenomena.[[35]](#footnote-36) Those models typically feature individual variables instead. The state of the integrate-to-bound system instantiated in LIP neurons is described by the firing rate model for those neurons and lacks a variable referring to the system, environment, or other parts of the brain besides the input. In addition, in systemic dynamicist models, global parameters control the behavior of the collective variables. In contrast, dynamical systems in cognitive neurobiology use local parameters. The integration of evidence over time is a particular function filled by neurons in area LIP and is partly controlled by local parameters, such as the rate of rise proportional to the strength of evidence specific to each neuron. Furthermore, coupling is not a necessary aspect of the description of the dynamics. Indeed, the integrate-to-bound system is treated as though it is isolated from other areas of the brain.

Different neurodynamical systems explain different subcapacities of the cognitive system. The neurodynamical systems (integrate-to-bound, divisive normalization, and even simple neuronal responses such as are found in area MT) are executing functions that are only part of the cognitive process being explained. In perceptual decision making, the integrate-to-bound system explains evidence integration but not the setting of priors on the probability of motion direction, setting thresholds for making a decision, or the execution of a decision upon threshold crossing. In strategic decision making, the integrate-to-bound system explains only one of the numerous functions necessary to make such decisions, possibly tracking the changing value of switching strategies, assessing opportunity costs, or evaluating the comparison between instantaneous and average reward rates. In attention, DN explains gain modulation but does not explain (amongst other aspects of attentional phenomena) selection of the object of attention. In reward-based decision making, DN explains reward encoding but does not explain selection of the choice set or the execution of the decision. In each case, the cognitive phenomenon is fully described only once taken together with other functions and other parts. That is why the function is an f-function and not a system capacity.

These explanations of different subcapacities are then combined in order to explain the capacities of the system. Research into many of the subcapacities is still in the early days. However, at present, the combination of the different explanations need not yield the introduction of collective variables, global control parameters, or coupling, at either the level of the brain’s parts or at the level of the brain, body, and environment. Take for example a combination of the activity of MT neurons sensitive to motion speed and direction with integrate-to-bound dynamics for LIP neurons that integrate motion evidence. LIP dynamics are described as partly driven by input from MT neurons, which in turn are described as sensitive to those stimulus properties. This combination lacks reference to the whole brain, and furthermore does not contain any collective variables or global control parameters. Indeed, it is unclear in what sense the system could even be described using a collective variable. Coupling between these two systems could be included if the LIP dynamics feedback on to the MT neurons. However, even in that case, coupling with the body or environment does not occur. Granted that the models lack systemic features and that their combination need not introduce them, cognitive systems seem to be componential dynamical systems composed of subsystems and subsubsystems, the neurodynamical systems themselves.

The manipulations used to probe cognitive phenomena also support the claim that cognitive systems are componential dynamical systems. Cognitive neurobiology engages in robust productive research programs assuming that cognitive systems are made of componential dynamical systems. For example, certain experiments suggest themselves, such as excitatory studies (cf. Craver 2007; Bechtel and Richardson 2010) that attempt to increase the activity of neurons in the relevant regions through electrostimulation (or some other means) to examine their influence on the behavioral responses of the system. Though such studies have not been run in area LIP, they have been run in other areas during the random dot motion task (Salzman, Britten et al. 1990; Gold and Shadlen 2000). The research programs also suggest inhibitory studies that reduce the activity of neurons in relevant regions. Recent results from such studies in area LIP have in fact called into question the causal role of the region in generating perceptual decisions (Katz, Yates et al. 2016).

These studies support the conclusion that cognitive systems are componential. Componential cognitive systems have subsystems and parts (the subsubsystems) that possess the subcapacities into which cognitive functions are analyzed. Stimulation or inhibition internal to the parts, such as occur when cognitive neurobiologists stimulate or inhibit subsystems of the brain, will directly disrupt the performance of these subcapacities without any intervening steps in the explanation. On a systemic view, the explanation of the effect of such manipulations is more complex. Stimulation or inhibition might manipulate cognitive processing on the systemic view, but only in virtue of such effects subsequently changing the relationships between the brain, body, and environment. This indirect effect is a more complex explanation than the direct effects implied by componential dynamicism. Considerations of simplicity favor componential dynamicism.

Further justification for seeing the systems as componential comes from an analysis of the nature of debates about the observed dynamics in area LIP. An alternative to the integrate-to-bound system, a discrete step system, has been proposed to explain the observed activity in LIP, and recent modeling suggests that discrete stepping dynamics may better describe this activity (Latimer, Yates et al. 2015; for response see Shadlen, Kiani et al. 2016). But these are debates about what componential model better describes the underlying dynamical system. Similarly, different models for the integrative activity (specifically, the mathematical description of the integration) also reflect debate about what model best describes the underlying dynamics. These debates are hard to explain on a systemic understanding of cognitive systems.

In sum, the dynamical systems studied by cognitive neurobiologists imply that cognition is componential, not systemic. Unlike systemic dynamical systems which use collective variables and global parameters and lack parts, dynamical systems in cognition are parts of cognitive systems modeled using individual variables and local parameters. The best explanation for these modeling practices of cognitive neurobiologists is that the targets of these models are componential dynamical systems. Furthermore, the types of experiments pursued to uncover how cognitive systems operate suggest componential systems. Finally, the debates in the field regarding neurodynamical systems suggest they are componential dynamical systems.

All three arguments for componential dynamicism support the hypothesis that cognitive systems have parts, the neurodynamical systems that execute f-functions for cognition. The combination of different neurodynamical systems need not yield coupling, collective variables or global control parameters, or the introduction of environmental or bodily influences. Hence, cognitive systems are componential dynamical systems themselves. Componential dynamicism is similar to systemic dynamicism in the use of dynamical systems theory to explain cognitive phenomena, but shares with computationalism a focus on parts and their behavior.

1. Conclusion

In this essay, I provided some prima facie plausibility for a new kind of dynamicism, componential dynamicism. In virtue of being dynamicist, cognitive systems are described using the concepts and mathematics of dynamical systems theory. In virtue of being componential, cognitive systems are characterized as having subsystems and parts that execute functions for them. I described a novel explanatory schema in cognitive neurobiology that contains these two elements. The parts are the dynamical systems being revealed by recent work in cognitive neurobiology. The mapping of the mathematical description of the functions for cognitive capacities on to the mathematical description of the neural dynamics grounds the claim that these dynamical systems perform particular functions for cognitive systems and thereby help explain cognitive phenomena.

While I have argued for a place in conceptual space for a componential dynamicist approach to cognition, one that sanctions components of cognitive systems while retaining a dynamical bent, there may be cases where a systemic dynamicist approach is appropriate. And, in some cases, a mixed componential and systemic dynamicist approach may be appropriate. The framework outlined herein is meant to be driven by empirical findings on the ground, and the type of dynamical system used to explain a particular cognitive phenomenon is tailored to suit each case.

Dynamicism historically focuses on coupled interactions between brains, bodies, and environments described using dynamical systems theory. Computationalism focuses on the functions and behaviors of parts of cognitive systems. These two views of the mind have often been contrasted. In this essay, I have attempted to weave together threads from both approaches. Componential dynamicism shares the use of dynamical systems theory with past dynamicism, but shares a focus on the functions and behaviors of parts with computationalism.

Componential dynamicism respects the state of the scientific art. However, it may not be the only approach to do so. At best, the account herein represents an empirical argument for componential dynamicism from current cognitive neurobiological research. However, I think other arguments can be marshalled for the view. The view captures certain regularities in cognitive systems that escape other views of cognition. Additionally, componential dynamicism shares tantalizing similarities to some recent interpretations of systems biology (Levy and Bechtel 2013; Green, Levy et al. 2015), possibly heralding a unification of cognitive neurobiological research with other branches of biology. Discussing these other arguments must await another time.

So how does the monkey detect its dappled predator, the leopard? The monkey’s perceptual subsubsystem, and in particular the motion detection system, is organized as a set of cooperating neurodynamical systems that have the capacity to execute on-the-fly transformations of sensory data. One part of the monkey’s brain, the lateral intraparietal area (or its analogue) integrates motion evidence from visual area MT by instantiating an integrate-to-bound dynamical system. These dynamics combine with dynamics from other processing systems in the monkey such that together, the perceptual and other cognitive subsubsystems execute the drift diffusion model for perceptual decisions, and determine that in fact that dappled mix of shifting shadows is—without doubt—a leopard.

Footnote refs: (Strawson 1959; Pour-El 1974; Cummins 1975; Fodor 1975; Horgan and Tienson 1992; Horgan and Tienson 1994; Kelso 1995; Horgan, Horgan et al. 1996; Giunti 1997; Bechtel 1998; Craver 2001; Chemero and Silberstein 2008; Zednik 2008; Anderson 2010; Kaplan 2011; Kaplan and Craver 2011; Stewart 2011; Carandini and Heeger 2012; Kolling, Behrens et al. 2012; Eliasmith 2013; Anderson 2014; Shenhav, Straccia et al. 2014; Latimer, Yates et al. 2015; Piccinini 2015; Rice 2015; Katz, Yates et al. 2016; Kolling, Wittmann et al. 2016; Chirimuuta 2017; Huneman 2017; Neander 2017)

**Bibliography**

Aljadeff, J., B. J. Lansdell, A. L. Fairhall and D. Kleinfeld (2016). "Analysis of neuronal spike trains, deconstructed." Neuron **91**(2): 221-259.

Andersen, R. and V. Mountcastle (1983). "The influence of the angle of gaze upon the excitability of the light- sensitive neurons of the posterior parietal cortex." The Journal of Neuroscience **3**(3): 532-548.

Anderson, J. R. (1990). The adaptive character of thought, Psychology Press.

Anderson, M. L. (2010). "Neural reuse: A fundamental organizational principle of the brain." Behavioral and Brain Sciences **33**(04): 245-266.

Anderson, M. L. (2014). After Phrenology, Oxford University Press.

Bailey, T. N. (1993). The African leopard: ecology and behavior of a solitary felid, Columbia University Press.

Barack, D. L., S. W. C. Chang and M. L. Platt (2017). "Posterior cingulate neurons dynamically signal decisions to disengage during foraging." Neuron **96**(2): 339-347.e335.

Barack, D. L. and M. L. Platt (2017). Engaging and Exploring: Cortical Circuits for Adaptive Foraging Decisions. Impulsivity, Springer**:** 163-199.

Bechtel, W. (1998). "Representations and cognitive explanations: Assessing the dynamicist's challenge in cognitive science." Cognitive Science **22**(3): 295-318.

Bechtel, W. and R. Richardson (2010). Discovering Complexity, MIT Press.

Beer, R. D. (1995). "A dynamical systems perspective on agent-environment interaction." Artificial intelligence **72**(1): 173-215.

Beer, R. D. (2000). "Dynamical approaches to cognitive science." Trends in Cognitive Sciences **4**(3): 91-99.

Bergeron, V. (2007). "Anatomical and functional modularity in cognitive science: Shifting the focus." Philosophical Psychology **20**(2): 175-195.

Britten, K. H. and W. T. Newsome (1998). "Tuning bandwidths for near-threshold stimuli in area MT." Journal of Neurophysiology **80**(2): 762-770.

Britten, K. H., M. N. Shadlen, W. T. Newsome and J. A. Movshon (1992). "The analysis of visual motion: a comparison of neuronal and psychophysical performance." J Neurosci **12**(12): 4745-4765.

Britten, K. H., M. N. Shadlen, W. T. Newsome and J. A. Movshon (1993). "Responses of neurons in macaque MT to stochastic motion signals." Vis Neurosci **10**(6): 1157-1169.

Carandini, M. and D. J. Heeger (2012). "Normalization as a canonical neural computation." Nature Reviews Neuroscience **13**(1): 51-62.

Carandini, M., D. J. Heeger and J. A. Movshon (1997). "Linearity and normalization in simple cells of the macaque primary visual cortex." J Neurosci **17**(21): 8621-8644.

Cavanaugh, J. R., W. Bair and J. A. Movshon (2002). "Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons." Journal of neurophysiology **88**(5): 2530-2546.

Charnov, E. L. (1976). "Optimal foraging, the marginal value theorem." Theor Popul Biol **9**(2): 129-136.

Chemero, A. (2011). Radical embodied cognitive science, MIT press.

Chemero, A. and M. Silberstein (2008). "After the Philosophy of Mind: Replacing Scholasticism with Science\*." Philosophy of Science **75**(1): 1-27.

Chirimuuta, M. (2017). "Explanation in computational neuroscience: Causal and non-causal." The British Journal for the Philosophy of Science: axw034.

Clark, A. (1980). Psychological Models & Neural Mechanisms, Oxford University Press.

Clark, A. (2000). Mindware: An introduction to the philosophy of cognitive science, Oxford University Press.

Clark, A. (2013). Mindware: An introduction to the philosophy of cognitive science, Oxford University Press, Inc.

Craver, C. F. (2001). "Role Functions, Mechanisms, and Hierarchy." Philosophy of science **68**(1): 53-74.

Craver, C. F. (2007). Explaining the brain, Oxford University Press Oxford.

Craver, C. F. and L. Darden (2013). In search of mechanisms: Discoveries across the life sciences, University of Chicago Press.

Cummins, R. (1975). "Functional Analysis." Journal of Philosophy **72**(20): 741-765.

Dennett, D. C. (1981). Brainstorms: Philosophical essays on mind and psychology, MIT press.

Eliasmith, C. (2013). How to build a brain: A neural architecture for biological cognition, Oxford University Press.

Fodor, J. A. (1975). The language of thought, Harvard University Press.

Giunti, M. (1997). Computation, dynamics, and cognition, Oxford University Press.

Gold, J. I. and M. N. Shadlen (2000). "Representation of a perceptual decision in developing oculomotor commands." Nature **404**(6776): 390-394.

Gold, J. I. and M. N. Shadlen (2007). "The neural basis of decision making." Annu Rev Neurosci **30**: 535-574.

Green, S., A. Levy and W. Bechtel (2015). "Design sans adaptation." European Journal for Philosophy of Science **5**(1): 15-29.

Grush, R. (1997). "Yet another design for a brain? Review of Port and van Gelder." Mind as motion. Philosophical Psychology **10**: 233-242.

Haken, H., J. S. Kelso and H. Bunz (1985). "A theoretical model of phase transitions in human hand movements." Biological cybernetics **51**(5): 347-356.

Hayden, B. Y., J. M. Pearson and M. L. Platt (2011). "Neuronal basis of sequential foraging decisions in a patchy environment." Nat Neurosci **14**(7): 933-939.

Heeger, D. J. (1992). "Normalization of cell responses in cat striate cortex." Vis Neurosci **9**(2): 181-197.

Heuer, H. W. and K. H. Britten (2002). Contrast Dependence of Response Normalization in Area MT of the Rhesus Macaque.

Hikosaka, O., Y. Takikawa and R. Kawagoe (2000). "Role of the basal ganglia in the control of purposive saccadic eye movements." Physiological reviews **80**(3): 953-978.

Horgan, T., T. Horgan and J. Tienson (1996). Connectionism and the Philosophy of Psychology, Mit Press.

Horgan, T. and J. Tienson (1992). "Cognitive systems as dynamical systems." Topoi **11**(1): 27-43.

Horgan, T. and J. Tienson (1994). "A nonclassical framework for cognitive science." Synthese **101**(3): 305-345.

Huneman, P. (2017). "Outlines of a theory of structural explanations." Philosophical Studies: 1-38.

James, W. (1890/1950). The Principles of Psychology. New York, Dover Publications.

Jenny, D. (1996). "Spatial organization of leopards Panthera pardus in Taï National Park, Ivory Coast: is rainforest habitat a ‘tropical haven’?" Journal of Zoology **240**(3): 427-440.

Jenny, D. and K. Zuberbühler (2005). "Hunting behaviour in West African forest leopards." African Journal of Ecology **43**(3): 197-200.

Kacelnik, A., M. Vasconcelos, T. Monteiro and J. Aw (2011). "Darwin’s “tug-of-war” vs. starlings’“horse-racing”: how adaptations for sequential encounters drive simultaneous choice." Behavioral Ecology and Sociobiology **65**(3): 547-558.

Kaplan, D. M. (2011). "Explanation and description in computational neuroscience." Synthese **183**(3): 339-373.

Kaplan, D. M. and C. F. Craver (2011). "The Explanatory Force of Dynamical and Mathematical Models in Neuroscience: A Mechanistic Perspective\*." Philosophy of science **78**(4): 601-627.

Katz, L. N., J. L. Yates, J. W. Pillow and A. C. Huk (2016). "Dissociated functional significance of decision-related activity in the primate dorsal stream." Nature **535**(7611): 285-288.

Kelso, J. A. S. (1995). Dynamic Patterns: The Self-Organization of Brain and Behavior. Cambridge, MA, MIT Press.

Kiyonaga, A. and T. Egner (2016). "Center-surround inhibition in working memory." Current Biology **26**(1): 64-68.

Kolling, N., T. E. J. Behrens, R. B. Mars and M. F. S. Rushworth (2012). "Neural mechanisms of foraging." Science **336**(6077): 95-98.

Kolling, N., M. K. Wittmann, T. E. Behrens, E. D. Boorman, R. B. Mars and M. F. Rushworth (2016). "Value, search, persistence and model updating in anterior cingulate cortex." Nature Neuroscience **19**(10): 1280-1285.

Lamb, M. and A. Chemero (2014). Structure and application of dynamical models in cognitive science. Proceedings of the Annual Meeting of the Cognitive Science Society.

Latimer, K. W., J. L. Yates, M. L. Meister, A. C. Huk and J. W. Pillow (2015). "Single-trial spike trains in parietal cortex reveal discrete steps during decision-making." Science **349**(6244): 184-187.

Levy, A. and W. Bechtel (2013). "Abstraction and the organization of mechanisms." Philosophy of science **80**(2): 241-261.

Louie, K., L. E. Grattan and P. W. Glimcher (2011). "Reward Value-Based Gain Control: Divisive Normalization in Parietal Cortex." The Journal of Neuroscience **31**(29): 10627-10639.

Lycan, W. G. (1981). "Form, function, and feel." The Journal of Philosophy **78**(1): 24-50.

Maunsell, J. H. and D. C. Van Essen (1983). "Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation." Journal of neurophysiology **49**(5): 1127-1147.

Mazurek, M. E., J. D. Roitman, J. Ditterich and M. N. Shadlen (2003). "A role for neural integrators in perceptual decision making." Cereb Cortex **13**(11): 1257-1269.

McAdams, C. J. and J. H. Maunsell (1999). "Effects of attention on the reliability of individual neurons in monkey visual cortex." Neuron **23**(4): 765-773.

Mink, J. W. (1996). "The Basal Ganglia: Focused Selection and Inhibition of Competing Motor Programs." Prog Neurobiol **50**(4): 381-425.

Mink, J. W. (2003). "The basal ganglia and involuntary movements: impaired inhibition of competing motor patterns." Archives of neurology **60**(10): 1365-1368.

Neander, K. (2017). A Mark of the Mental: In Defense of Informational Teleosemantics, MIT Press.

Newell, A. (1980). "Physical Symbol Systems\*." Cognitive Science **4**(2): 135-183.

Paninski, L., J. W. Pillow and E. P. Simoncelli (2004). "Maximum likelihood estimation of a stochastic integrate-and-fire neural encoding model." Neural computation **16**(12): 2533-2561.

Piccinini, G. (2015). Physical computation: A mechanistic account, OUP Oxford.

Platt, M. L. and P. W. Glimcher (1999). "Neural correlates of decision variables in parietal cortex." Nature **400**(6741): 233-238.

Port, R. F. and T. van Gelder (1995). Mind as motion: Explorations in the dynamics of cognition, MIT press.

Pour-El, M. B. (1974). "Abstract computability and its relation to the general purpose analog computer (some connections between logic, differential equations and analog computers)." Transactions of the American Mathematical Society **199**: 1-28.

Putnam, H. (1967). The mental life of some machines. Intentionality, Minds and Perception. H.-N. Castaneda, Wayne State University Press.

Reynolds, J. H. and D. J. Heeger (2009). "The normalization model of attention." Neuron **61**(2): 168-185.

Rice, C. (2015). "Moving beyond causes: Optimality models and scientific explanation." Noûs **49**(3): 589-615.

Roitman, J. D. and M. N. Shadlen (2002). "Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task." J Neurosci **22**(21): 9475-9489.

Salzman, C. D., K. H. Britten and W. T. Newsome (1990). "Cortical microstimulation influences perceptual judgements of motion direction." Nature **346**(6280): 174-177.

Shadlen, M. N., R. Kiani, W. T. Newsome, J. I. Gold, D. M. Wolpert, A. Zylberberg, J. Ditterich, V. de Lafuente, T. Yang and J. Roitman (2016). "Comment on “Single-trial spike trains in parietal cortex reveal discrete steps during decision-making”." Science **351**(6280): 1406-1406.

Shadlen, M. N. and W. T. Newsome (2001). "Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey." J Neurophysiol **86**(4): 1916-1936.

Shapiro, L. A. (2013). "Dynamics and cognition." Minds and Machines **23**(3): 353-375.

Shenhav, A., M. A. Straccia, J. D. Cohen and M. M. Botvinick (2014). "Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value." Nature neuroscience **17**(9): 1249-1254.

Stephens, D. W. and J. R. Krebs (1986). Foraging Theory. Princeton, NJ, Princeton University Press.

Stepp, N., A. Chemero and M. T. Turvey (2011). "Philosophy for the rest of cognitive science." Topics in Cognitive Science **3**(2): 425-437.

Stewart, J. (2011). Single variable calculus: early transcendentals, Brooks/Cole Publishing Company.

Strait, C. E., T. C. Blanchard and B. Y. Hayden (2014). "Reward value comparison via mutual inhibition in ventromedial prefrontal cortex." Neuron **82**(6): 1357-1366.

Strawson, P. F. (1959). Individuals. London, Methuen.

Strogatz, S. (2001). Nonlinear Dynamics And Chaos: With Applications To Physics, Biology, Chemistry, And Engineering (Studies in Nonlinearity), Westview Press.

Treue, S. and J. C. Martinez Trujillo (1999). "Feature-based attention influences motion processing gain in macaque visual cortex." Nature **399**(6736): 575-579.

Usher, M. and J. L. McClelland (2001). "The time course of perceptual choice: the leaky, competing accumulator model." Psychological review **108**(3): 550.

van Gelder, T. (1995). "What Might Cognition Be, If Not Computation?" The Journal of Philosophy **92**(7): 345-381.

Walmsley, J. (2008). "Explanation in dynamical cognitive science." Minds and Machines **18**(3): 331-348.

Wang, X. J. (2002). "Probabilistic decision making by slow reverberation in cortical circuits." Neuron **36**(5): 955-968.

Wang, X. J. (2008). "Decision making in recurrent neuronal circuits." Neuron **60**(2): 215-234.

Wheeler, M. (2005). Reconstructing the cognitive world: The next step, MIT press.

Young, M. P., C. C. Hilgetag and J. W. Scannell (2000). "On imputing function to structure from the behavioural effects of brain lesions." Philosophical Transactions of the Royal Society of London B: Biological Sciences **355**(1393): 147-161.

Zednik, C. (2008). Dynamical models and mechanistic explanations. Proceedings of the 30th annual conference of the cognitive science society.

Zednik, C. (2009). "The Varieties of Dynamicism." Proceedings of the 31st Annual Conference of the

Cognitive Science Society.: 2298-2303.

Zednik, C. (2011). "The Nature of Dynamical Explanation\*." Philosophy of Science **78**(2): 238-263.

Zednik, C. (2015). Heuristics, descriptions, and the scope of mechanistic explanation. Explanation in Biology, Springer**:** 295-318.

Zeki, S. (1991). "Cerebral akinetopsia (visual motion blindness)." Brain **114**(2): 811-824.

Zeki, S. M. (1974). "Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey." The Journal of Physiology **236**(3): 549.

Zuberbühler, K. and D. Jenny (2002). "Leopard predation and primate evolution." Journal of Human Evolution **43**(6): 873-886.

Zuberbühler, K., D. Jenny and R. Bshary (1999). "The predator deterrence function of primate alarm calls." Ethology **105**(6): 477-490.

1. Acknowledgements: This work has benefitted greatly from a number of comments over many years. The work draws on my dissertation at Duke University, and I would like to thank Walter Sinnott-Armstrong, Alex Rosenberg, Karen Neander, Felipe De Brigard, and Michael Platt for their thoughtful input. I would also like to thank Gaultiero Piccinini, Chris Peacocke, William Lycan, Ann-Sophie Barwich, Jorge Morales, Nemira Gasiunas, Michael Weisberg, and the University of Pennsylvania philosophy of science reading group for critical comments. [↑](#footnote-ref-2)
2. In particular, the discreteness of the system (see, e.g., van Gelder 1995) and the seriality of the processing (see, e.g., Wheeler 2005) have also distinguished the approaches. Discreteness and seriality, however, are deeper issues that require their own discussion, and so I skip them herein. [↑](#footnote-ref-3)
3. For this essay, I adopt an object, property, and relation ontology. On such a view, objects are analyzed as bundles of properties with some substrate. This approach is adopted for convenience only, and the claims can be suitably recast in other terms. In order to cast a wide net, in this section I have framed the view in terms of individuals in the logical sense (cf. Strawson 1959). I thank a reviewer for pushing me on this point. [↑](#footnote-ref-4)
4. This distinction between dynamical systems and dynamical systems theory is often overlooked. I explore and defend this distinction elsewhere. [↑](#footnote-ref-5)
5. As such, dynamical systems theory covers both classical (see, e.g., Piccinini 2015) and non-classical (see, e.g., Pour-El 1974 or Giunti 1997) accounts of computation. [↑](#footnote-ref-6)
6. Systems can also evolve with respect to other variables besides time as noted, but to simplify I will often focus on time. [↑](#footnote-ref-7)
7. Some dynamicists might object to such a characterization of dynamical systems as ignoring parts of these systems. Specifically, Bechtel classifies connectionist models that refer to internal operations as dynamical systems that do refer to parts (Bechtel 1998). Bechtel says that “the question of how [dynamical] models comport with mechanistic explanatory objectives does not arise, since connectionist models are models of mechanisms. The [dynamicist] approach is employed by these theorists to analyze how these mechanisms behave” (Bechtel 1998, p. 312). This analysis involves decomposing the mechanism into parts and localizing the operations of the mechanism to the activity of these parts (Bechtel 1998; Bechtel and Richardson 2010). However, not all philosophers agree with this broad characterization of connectionism’s relation to dynamicism. For example, Zednik disagrees, arguing that “[i]nsofar as connectionist dynamical models do not… capture the dynamics of coupled brain-body-environment systems, but merely describe “internal” neural dynamics, they differ substantially” from other connectionist models that focus on “the behavior of the whole… brain-body-environment system” (Zednik 2008, p. 1457-1458). These latter systems “can be understood as the interactive operations between two simpler subsystems: the agent’s motion M as the operation of subsystem A (the agent) on the one hand, and the sensory stimulus function S as the operation of subsystem E (the environment) on the other” (Zednik 2008, p. 1458). Notably, the behavior of these models is explained by an “understanding of the way in which individual parts of the system—the agent on the one hand and the environment on the other—interact” (Zednik 2009, p. 2301). There is thus some debate as to the appropriate interpretation of certain connectionist systems, and some such systems do refer to parts or parts of parts while others do not. My componential dynamicism is consistent with connectionist approaches that focus on internal dynamics. [↑](#footnote-ref-8)
8. A notable exception to this systemic dynamicist approach is Horgan and Tienson’s Dynamical Cognition (Horgan and Tienson 1992; Horgan and Tienson 1994; Horgan and Tienson 1996). Their view is substantially different from other dynamicist approaches and I will not be addressing it herein. [↑](#footnote-ref-9)
9. For extensive discussion, see Kelso 1995. The model has received much philosophical attention (for recent discussion, see Chemero and Silberstein 2008; Zednik 2008; Zednik 2009; Zednik 2011; Kaplan 2011; Kaplan and Craver 2011). [↑](#footnote-ref-10)
10. Kelso himself would probably endorse parts, seeing as his research program includes both understanding how the neural mechanisms give rise to the dynamics captured by the HKB model and applying the lessons gleaned from dynamical analysis of behavioral phenomena to the brain. I do not have space to fully address Kelso’s own dynamicism (Kelso 1995), which is a remarkably deep and distinct application of dynamical ideas to analyzing cognitive phenomena and which should also be seen as an exception to the systemic approach. Kelso’s view is that there are dynamical laws that underlie many complex systems, including cognitive ones, and that cognitive phenomena result from these laws. In addition to this nomological component, his view has a mechanistic component, seeing the dynamical systems as mechanistic parts. For the time being, suffice to say that Kelso is not a target of my criticism. [↑](#footnote-ref-11)
11. From now on, I will exclusively use the term ‘parts’ to refer to subsubsystems of the system. [↑](#footnote-ref-12)
12. See below, §3.3, for a lengthier discussion of functions and componential dynamical systems. [↑](#footnote-ref-13)
13. Note that the parts of the fingers like their nerves or muscles do not necessarily determine the relative phase. Usually, those properties will determine the fingers’ movements. But there could be ways of setting the value for the relative phase other than the musculature, nerves, and other physiological parts that typically cause finger movements. Imagine, for example, that a mechanical device is used to oscillate the fingers, mimicking the behavioral output of the usual physical parts. On systemic dynamicism, this mechanical system is described in the same way as when the musculature and nerves control the movements. Systemic dynamicism cannot distinguish between the two systems, one organically controlled, the other artificially, because the role of the parts is excluded by systemic dynamicism. The inclusion of a function describing the motive forces on the fingers could distinguish between different sources of movement. Such a description will distinguish different sources of movement unless the proximal causes of the finger movements—viz., muscles and nerves vs. artificial means—have precisely the same force functions in all contexts. But even then, the equations corresponding to such functions are interpreted; so the variables refer to those proximal causes, which will be organic in the first case and artificial in the second. [↑](#footnote-ref-14)
14. This division of types of dynamical systems is not meant to be a stark contrast. There can be systems with mixed dynamical descriptions. A car can be described systemically, componentially, or in a mixed fashion, with some terms for its environment and systemic properties and other terms for the properties of the parts of the car and environment. Thus, the two types of dynamical description stand at extremes on a scale of dynamical descriptions, with most such systems lying somewhere in the midst of the spectrum. [↑](#footnote-ref-15)
15. From this point on, I will focus on changes in properties, but this should not be read as excluding changes in objects or relations as underlying dynamical properties. [↑](#footnote-ref-16)
16. In fact, neurophysiological systems implement dynamical systems in a technical sense that I explore in other work. For the time being, instantiation here should be understood in the sense of token identity, so that if a neuron instantiates a dynamical system, then a subset of the neuron’s dynamical properties are an instance of the dynamical system; if a neural population instantiates a dynamical system, then a subset of the population’s dynamical properties are an instance of the dynamical system; etc. In addition, the neuron, neural population, etc. are themselves dynamical systems. [↑](#footnote-ref-17)
17. See Gold and Shadlen (2007) for extensive discussion of this research. Note that many aspects of this case are still actively researched. In particular, whether neurons in area LIP truly exhibit the integrate-to-bound dynamics discussed below is hotly debated (Latimer et al. 2015). For my present purposes, the fact that the details are still not settled does not matter, as I am merely illustrating how such explanations are constructed. [↑](#footnote-ref-18)
18. When describing this state change mathematically, both the smooth and non-saltatory nature of the change is captured by a continuous mathematical function. The mathematical function does not contain jump or point discontinuities, corresponding to a lack of saltatory changes in the system. The function is also continuously differentiable, corresponding to a lack of abrupt changes in the system. While there is also typically some monotonicity in the trajectories, the trajectory’s evolution is correlated with some input such that if the input changes sign, then the trajectory can switch from an increase to a decrease or vice versa. A description of a system’s evolution wherein the system resides in the same state despite changes in some variable does not count as an example of integration. [↑](#footnote-ref-19)
19. In other work, I intend to present a deeper analysis of functions. [↑](#footnote-ref-20)
20. Causal functions may be clearest example of compositional functions. Constitution is taken to be distinct from causation for a number of reasons, including that the former is synchronic and the latter diachronic. Whether a component of a physical system can help constitute that system without causal powers is an interesting open question, though this has recently been argued by a range of philosophers (see, e.g., Huneman 2017 or Chirimuuta 2017). [↑](#footnote-ref-21)
21. Here, I deliberately include both causal and constitutive roles for the part in c-functions. [↑](#footnote-ref-22)
22. This model is often dubbed the drift diffusion model for historical reasons. [↑](#footnote-ref-23)
23. The system could also run out of evidence (e.g., if the dot display disappears) or some other stopping criterion could be met that results in a decision. [↑](#footnote-ref-24)
24. Carandini and Wang both discuss the neurons as computing an exponentiation function (Carandini and Heeger 2012; Wang 2002). In a case of pure exponential growth, for a given strength of evidence, the relative change in the firing rate, or the change in firing rate divided by the current rate, is constant (Stewart 2011). However, upon removal of the stimulus, LIP neurons will not maintain their firing rates indefinitely, and adding state-dependent leak and recurrent excitation exponential decay terms captures this decay (Usher and McClelland 2001; Wang 2002). Furthermore, the integration is noisy, requiring the addition of a noise term. Thus, neurons such as those in area LIP are ‘noisy leaky integrators’; their firing rate is a function not only of the change in firing but also contains a noise term and a loss (or ‘leak’) term (e.g. see Usher and McClelland 2001 or Wang 2002; for a philosophical discussion, see Eliasmith 2013, p. 43ff). The leak is dependent upon the state of the system, and so the internal system dynamics contribute to the state of the system. Wang’s model is even more complex but in short contains an additional recurrent excitation term that is driven by the state of the system. These noise and leak terms effectively curtail the integrative growth function as well as enforcing a biologically plausible decay in the firing rate at some time constant absent any driving force, while the differences in the strength of evidence correspond to differences in the relative growth rate constant driving the neuron’s activity. [↑](#footnote-ref-25)
25. Exactly how many steps depends on the number of subcapacities of the capacity. The set of subcapacities often implies some ordering to their execution. Some subsets of subcapacities may be executed serially or in parallel, whereas other subsets must be executed seriatim. [↑](#footnote-ref-26)
26. Does the part need to be causally active in the cognitive system? I will remain neutral on this issue herein; some explanations of cognitive capacities might reflect purely structural features of systems such that the parts’ causal powers are not relevant (see, e.g., Chirimuuta 2017 or Huneman 2017). In the examples herein, causality is important, but I do not want to beg any questions about other cognitive phenomena. [↑](#footnote-ref-27)
27. For this discussion, the part of the brain, area LIP, is a subsubsystem because the subsystem is the brain itself. [↑](#footnote-ref-28)
28. Perhaps these processes should not be considered cognitive. I construe the term ‘cognitive’ quite widely, so I set aside this concern. [↑](#footnote-ref-29)
29. The computations executed by these neural dynamics are often magnitude to magnitude transformations or functions on magnitudes. I note that these computations are not obviously like the representational transformations posited by classical computational theories like the language of thought (Fodor 1975). [↑](#footnote-ref-30)
30. Hence, showing that e.g. a generalized linear model containing some independent variable can significantly capture some of the variance in neuronal firing rates is insufficient to show that the neuron’s function is to signal that variable. In order to argue that the neuron’s function is in part to signal the variable, there must be at least the additional step of modeling the neuronal dynamics, such as by fitting a Gaussian tuning curve. [↑](#footnote-ref-31)
31. Despite instantiating the same dynamical system, the physiological systems are different in LIP and ACC. In the case of LIP and the RDMT, integrative activity occurring on the timescale of hundreds of milliseconds during a single trial is instantiated by an increase in firing by individual neurons. In the case of ACC and the patch foraging task, integrative activity occurring on the timescale of tens of seconds over many trials is instantiated by an increase in the peak activity of individual neurons during a trial. These distinct timecourses indicate distinct physiological mechanisms instantiating the integrate-to-bound system. Furthermore, the integrate-to-bound in LIP is instantiated via a noisy but continuous-like increase in firing rates, whereas the integrate-to-bound in ACC is instantiated in a series of discontinuous transient increases in firing rates. These distinct firing rate patterns also indicate distinct physiological mechanisms. These two differences suggest that distinguishing between the dynamical system and the physical mechanism that possesses the dynamics is necessary to understand how cognitive neurobiologists go about explaining cognitive phenomena. In addition, the differences in the integrative activity will be reflected in models that include greater neuronal details. More abstract models, however, that focus just on the integrating peaks of activity may still use the same mathematical operations as used to describe the integrate-to-bound activity in area LIP. [↑](#footnote-ref-32)
32. This will also illustrate how the schema applies well beyond decision making to a range of cognitive capacities. [↑](#footnote-ref-33)
33. I cannot do justice to the two decades long research project, and the dozens of studies, that have gone into the development of the DN model of attention. An excellent review is in (Reynolds and Heeger 2009). [↑](#footnote-ref-34)
34. The basic mathematical description of divisive normalization is

    for response of a cell *R*, input to the jth cell *Djn*, and normalization pool summed over the normalization input *Dkn* (Carandini and Heeger 2012, p. 54). The parameters *γ*, *n*, and *σ* are fit to the data, with *σ* controlling how quickly the firing of a cell reaches its maximum. When *σ* is very large, the normalizing input has little effect on the firing rate, and when *σ* ~ 0, the input is largely determined by the normalizing pool of activity. If *σ* is roughly equivalent to the normalization input *Dkn*, the cell is only moderately determined by the activity of the normalization pool. [↑](#footnote-ref-35)
35. Some models do contain such parameterizations. What’s key here is that such models need not; and, as the illustration above showed, certain central cases of explanation of cognitive phenomena by neurodynamics do not. [↑](#footnote-ref-36)