
Self-Organization and Self-Governance

Philosophy of the Social Sciences
XX(X) 1–25
© The Author(s) 2010
Reprints and permission: <http://www.sagepub.com/journalsPermissions.nav>
DOI: 10.1177/0048393110363435
<http://pos.sagepub.com>



J. T. Ismael^{1,2}

Abstract

The intuitive difference between a system that choreographs the motion of its parts in the service of goals of its own formulation and a system composed of a collection of parts doing their own thing without coordination has been shaken by now familiar examples of self-organization. There is a broad and growing presumption in parts of philosophy and across the sciences that the appearance of centralized information-processing and control in the service of system-wide goals is mere appearance, i.e., an explanatory heuristic we have evolved to predict behavior, but one that will eventually get swept away in the advancing tide of self-organization. I argue that there is a distinction of central importance here, and that no adequate science of complex systems can dispense with it.

Keywords

self-organization, self-governance, selves, dynamics, intentional systems theory, Dennett

The Advancing Tide of Self-organization

When we see behavior that looks coordinated and directed at an end, we tend to assume that there is a guide, that is, a central intelligence collecting information, formulating goals, and choreographing the movements of its parts. So, for example, armies are led by central command, cars are guided by drivers, countries are run by governments. But the intuitive difference between a system that choreographs the motion of its parts in the service of goals of its own

Received 4 March 2009

¹University of Sydney, Centre for Time, Sydney, Australia

²University of Arizona, Tucson, USA

Corresponding Author:

J. T. Ismael, Dept. of Philosophy, SS213, University of Arizona, Tucson, AZ 85750

Email: jtismael@u.arizona.edu

formulation and a system composed of a collection of parts doing their own thing without coordination has been shaken by now-familiar examples of self-organization. We have learned that systems can exhibit remarkably complex forms of apparently coordinated, goal-directed behavior without any real consolidation of information or collective decision making. Colonies of insects and “smart crowds” provide famous examples. And many have assumed that self-organization is the key to incorporating the coordinated patterns of apparently goal-directed activity we see on the human and social level into a seamless dynamical description of nature. Dennett has explicitly argued, what many have tacitly supposed, that explanations that treat complex systems as decision makers are simply heuristic. Self-organizing explanations would reveal the real dynamical underpinnings of behavior even at the level of the individual.

In the human and social sciences, however, many continue to feel the need for a distinction between systems for whom the attribution of deliberative rationality—that is, systems that form goals and reason about how to achieve them—and those in which it is only an *as-if* description that captures some gross regularities in behavior. Is this distinction a part of our folk theory of the world that should be swept away by the advancing tide of self-organization? Is the appearance of centralized information-processing control in the service of system-wide goals always merely appearance, a sort of *as-if* explanation that results from viewing at a low level of resolution, an explanatory heuristic we have evolved to predict behavior, but not an accurate representation of real processes?

The question is of relevance both to the nature of human intelligence and to the social sciences where we speak of groups of certain kinds as though they exhibit a form of deliberative rationality in a nonmetaphorical way. We talk of deliberative democracy and debate about the details of decision procedures for democratic societies. Corporations debate how collective decisions should be made. The applicability of game-theoretic models of human rationality, international relations, and corporate activity treats the agents involved as rational deliberators. Decision theorists take themselves to be describing real processes governing the behavior of agents at different levels of organization. I am going to argue that there is a distinction of importance here, though it is perhaps more subtle than the pretheoretic distinction. I shall introduce the distinction with a pair of contrasting schemas, try to make it precise, and then suggest some philosophical contexts on which it can shed some light.

Centralization and Self-Modeling

I shall use the labels “self-governance” and “self-organization” as terms of art to mark the two sides of this distinction. Let us get an example of self-governance in front of us to begin. Consider a ship that locates itself on a map

and uses the information contained in the map to navigate. By “the ship” I mean not simply the physical vessel, but the whole complex system including the crew and the instruments and computer networks that support it. None of the crew members has any global vision or plan for the vessel; each performs his own task following a simple script that requires him to take care of a certain piece of equipment, perform a certain calculation, or pass a certain piece of information on, oblivious to activity in other parts of the ship.¹ There is a captain and collection of commanding officers, none of whom individually, we shall suppose, possesses a plan or vision of the ship as a whole, but who collectively use the information passed to by the crew, in conjunction with an accoutrement of tools and procedures they only partly understand, to keep the ship on course. Navigation, we shall suppose, goes in cycles; sightings are made, instrument readings are taken, the ship’s location is plotted on a map; this is then compared against a destination, a direction is chosen, and commands are issued to crew that set sail and rudder positions, until the next cycle of self-location.² If there is a functional center of the navigational activity on the ship, a place where all of the information is brought together and transformed into a plan of action, it is the map. The map provides the representational space for a computation that guides the movements of the ship, serving as a causal nexus where information collected by the information-gathering components of the ship is brought together and from which commands that control the machinery of the ship are issued.³

I shall call the computational cycle that occurs in the space of the map the “self-representational loop” because it involves self-location and deliberation. We can make it as simple or complex as we like. When new instrument

¹The scripts are local and procedural. For those who know the terminology, all the level at which we will describe the ship, crew members are cellular automata. Their behavior can be characterized by a rule that gives a range of possible actions, and specifies an action for each input. Inputs are limited to the state of the immediate environment.

²We can suppose that all or any of the activity that transforms instrument readings into self-locating coordinates on the one end, and commands into motion on the other, are invisible to the map. The most efficient division of labor here will depend on how prone to error those subprocesses are, how penetrable they have to be to the higher level processes, and so on.

³Note that to say that the map is the functional center of the ship is not to say that there must be some place in the ship which is visibly the center of action. The map itself can be concrete and localized or virtual and distributed. It can be positioned on a tabletop, displayed on a screen, or stored in a computer database. What matters only is that it interfaces properly with the rest of the ship’s machinery, including crew members, who function here as part of the machinery.

readings come in, the map is revised and recentered. The new information may prompt reevaluation of destination. Once a destination is settled on, there is instrumental reasoning aimed at determination of means, and each cycle of the process terminates in commands issued to crew, which initiates procedures that guide the movements of the ship. At the next stage, new information comes in—information about changes in the environment, which are themselves in part the causal product of earlier stages in the cycle—and the cycle begins again. The map is revised and newly recentered; fresh commands are issued and translated by the crew into actions that affect the movements of the ship. The computation can take any form. If it occurs in a linguistic or quasi-linguistic medium, the transitions will be inferential and this will bring with it whatever norms govern such transitions. But it can also consist in rule-governed transformations of visual images or transitions between patterns of activation across a transistor or neural net. What makes the process a *loop* rather than a cycle is the causal link between the commands that the subsystem issues at one stage and the readings it receives at the next; information about its own output is fed back into the subsystem at the next cycle of computation in the form of perceptual information about the changes it effects in the surrounding landscape. The ship not only “acts” on what it “sees,” but “sees” how it “acts,” and this creates a feedback between input and output.

The process need not occur in discrete stages. We can start with a process in which the movements of a ship effect changes in the environment that prompt revision of a map that in its turn affects the movements of the ship, and then let the temporal separation between the stages get vanishingly small, so that the result is a map and ship locked in a relationship of continuous reciprocal causation with the casual loop of the temporally staggered procedure compressed along the temporal dimension.

The new dynamics for the joint system is then given by a pair of coupled differential equations in which the state of each is represented by an expression that includes a parameter that represents that of the other. This kind of coupling relationship is familiar in physics. A simple model is provided by a pair of pendulums affixed to a wall in such a way that vibrations produced by each affect the movements of the other. Loops like this, however, very quickly generate mathematically prohibitive complexity and we have a grasp on the detailed dynamics in only the very simplest cases.

I take it that this sort of system is a paradigm of one whose movements are coordinated by a central model that contains a vision or plan for the whole. We can recognize here a natural model for the relationship between the conscious, deliberative component of the human mind and the body. The body feeds information from different sensory pathways into the mind, which in its turn passes it through a self-representational loop where it is plotted jointly on a unified

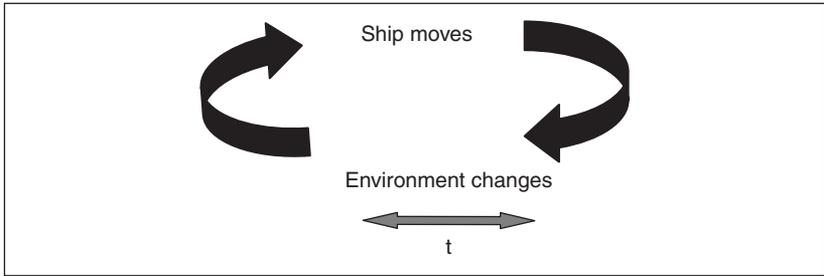


Figure 1

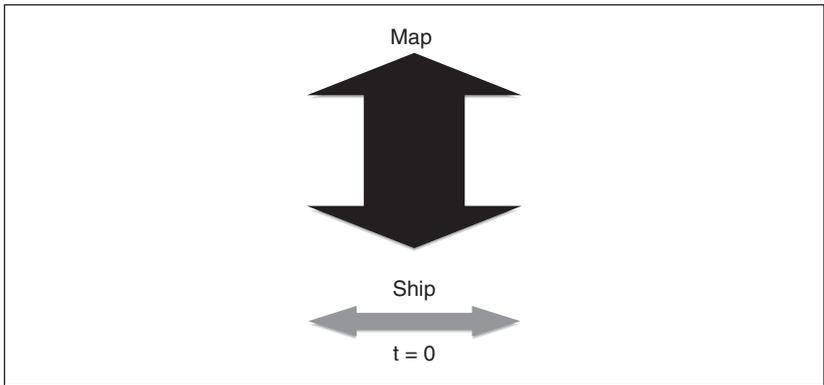


Figure 2

model of the body in its environment. All of the information embodied in the model—which can be anything from a simple map-like representation of the current state of the environment to a complete autobiographical history—is then brought to bear on the movements of fingers, toes, legs, and limbs. Not all parts of the body have their movements regulated by the self-representational loop; the beating of the heart and the activity of the immune and circulatory systems, for example, do not. But the parts of the body whose movements are regulated by the loop are coupled to the mind, and coupled to one another in virtue of their mutual coupling to the mind. These form a dynamical unit in a very strong sense.⁴ Each has a continuing role in the intrinsic dynamics of the other that makes them *effectively* inseparable from one another.

⁴This is not to deny that sensorimotor subsystems control some basic movements, or to hold that motor responses are, as a rule, centrally coordinated. It is to hold only that to the extent that they are, it is in virtue of their mutual involvement with the unified spatial model.

Self-Organization

Letting the ship serve as a model of a self-governance, let us turn to self-organization. There is dispute over the proper explicit characterization of self-organization or even whether there is single characterization that covers all intuitive cases. Informally, however, a self-organizing system is one in which there is no central locus of information and control.⁵ Information and control are both thoroughly distributed, and collective behavior is emergent from the individualistic dynamics of components in a manner that produces the illusion of coordinated effort. Social insects are the most familiar example. A colony of termites will build elaborate structures, explore its territory, store food for the winter, organize foraging expeditions, and so on. And it can seem irresistible to suppose that there is something inside the system coordinating the activity, whereas in fact, the group wisdom result is the product of a large number of individual termites, responding in a programmed ways to their immediate environment, each unaware of what most of the others are doing and without any collective goal or plan.

Biological examples of self-organization can be equally striking. Hydras, for example, are asymmetric freshwater creatures with a “head” on one end and a “tail” on the other. When cut in half, the upper end regenerates by growing a tail and the lower end regenerates by growing a head. What tells the individual cells on the bottom end of the hydra whether to form a head or a tail? It would seem that it has to be something in each of the severed halves with a global view and plan, something that knows that it wants both a head and tail and that can organize the regrowth on the damaged end in light of what it knows about the undamaged one. Slime molds are even weirder. A slime mold is a fungus that usually exists in form of individual cells. When food becomes scarce, the cells move toward one another, congregating, and then differentiating to form a mushroom-like structure with a stalk and cap. The new structure spreads its spores and the cycle begins again. The mechanisms for self-assembly and cell differentiation demand explanation. Again, it can seem irresistible to suppose that there is some central intelligence that

⁵There is no generally accepted definition of self-organization. The mechanisms that underpin the emergent behavior of self-organizing systems are complex and in many cases, not well understood. We have an intuitive working model of a class of closely studied examples. It is contested whether there is a general characterization self-organization, whether there is some dynamical essence that can be distilled out of these examples, or just a cluster of cases, exhibiting a syndrome of properties, remains to be seen.

calls the cells of the slime mold together, tells them where to congregate, and directs the formation of the new structure. But there is not. In both of these cases, the processes are explained in a manner that involves no centralized coordination or control.

Equally arresting examples of self-organization can be found in physics. Lasers, for example, work as follows: atoms embedded in a crystal emit individual light wave trains when excited. At first only a superposition of uncorrelated, amplified wave trains is observed but when the field amplitude is sufficiently high, the atoms begin to oscillate coherently. The transition is as striking to an observer as the transition from the warm-up to a musical performance in which musicians individually tuning their instruments emit random, uncorrelated sounds to the order of the performance after the conductor appears on stage. At the outset, there is apparently random, uncoordinated activity, and then suddenly synchronized behavior.⁶ Turbulent fluids, traffic systems, and market economies provide additional examples. In an orchestra, the transition occurs with the appearance of an orchestrator. In a laser, it happens spontaneously.

We have self-organization wherever we have a system composed of a collection of parts each following the beat of its own dynamical drum that somehow arranges itself into an ordered whole under random external pressure. A little more precisely, we have self-organization when we have emergence of order on the global level from the individualistic dynamics of components without any central coordination and without specific action from outside (Haken 2000).^{7,8} To the untutored dynamical imagination, self-organization is surprising because it is not obvious *how* random influence on (in many cases, undifferentiated) components could lead to internal differentiation via laws that relate components directly to local stimuli.

⁶The suddenness of the transition is inessential. Although it is a feature of many examples of self-organization, there are others in which the transition occurs gradually.

⁷Haken's theoretical orientation is more mathematical and heavily influenced by physical examples than other treatments. See, for example, Flake (2000); Nicolis and Prigogine (1989); Waldrop (1992); Johnson (2001); Cowan, Pines, and Meltzer (1999); and Bar-Yam (1997). The proper formal characterization of self-governance is a matter of controversy. I employ Haken because of the information-theoretic focus. Nothing I say should hinge on peculiarities of his views.

⁸For our purposes, it will do to think of any kind of nonrandom internal differentiation as order.

We know in detail how it works only in some very simple cases, but these are instructive. In these cases, self-organization involves interactions in which locally defined variables that carry information about the system's global state constrain or control the behavior of components. Clever networks of feedback and feed forward loops give rise to a global state represented by collective variables known as "order parameters." These then capture, or "enslave" the system's components, setting up the interaction between levels that is needed to generate internal order. It is convenient to think of the order parameters as characterizing a field that covers the space in which the components operate; local values of the field determine the behavior of components in a manner prescribed by their own intrinsic dynamics. Consider how coherent wave fronts are formed in a laser. Laser-active atoms embedded in a crystal—for example, a ruby—emit individual light wave trains when excited. These hit other excited atoms in the laser cavity, causing the original wave to be amplified and when the amplitude gets high enough, the atoms begin to oscillate coherently and the field is represented by an effectively infinite sinusoidal wave. Hermann Haken, in a discussion of the process, writes:

We have here a typical example of self-organization where the temporal structure of the coherent wave emerges without interference from the outside. . . . The detailed mathematical theory shows that the emerging coherent light wave serves as order parameter which forces the atoms to oscillate coherently, or in other words it *enslaves* the atoms. Note that we are dealing here with circular causality: On the one hand the order parameter enslaves the atoms, but on the other hand it is itself generated by the joint action of the atoms. (Haken 2006, 25)

The heart of the process is a transfer of information between hierarchically organized levels of organization, the level of the collective (the "macro"-level), on the one hand, and that of the individual (the "micro"-level), on the other.⁹ We are familiar with this sort of inter-level transfer of information. When we carry out measurements on microscopic systems, we transfer information from the microscopic to the macroscopic. When we look at a cell under a microscope, we transfer it from the mesoscopic to the macroscopic. When we encode the image on a TV screen in light waves, we transfer it in

⁹The distinction carries no imputation of absolute size, and a system that is microscopic in the context of one discussion may be treated as macroscopic in the context of another.

the other direction from the macroscopic to the microscopic. Such transfer occurs *within a single system* when we divide a system into parts and put mechanisms in place that make information about the whole, information embodied in its global configuration, available to the parts for the regulation of their behavior. In such cases, parameters that contain information about collective variables are placed in control of the states of constituents. Again, in Haken's words:

There is a hierarchy of informational levels. At the lowest level, the individual parts can emit information that hits other parts of the system. . . . Although, in all these cases the exchange of information may initially occur at random, a competition or cooperation between different kinds of signals sets in, and eventually a new collective state is reached which differs qualitatively from the disordered or correlated state present before. Thus, a new state is described by an order parameter or a set of order parameters. The states of the individual parts are determined by means of the slaving principle. (Haken 2006, 25)

Cell differentiation works this way in the hydra. A modulated chemical field is established by the production of diffusion of chemicals inside the organism. These chemicals are present in different concentrations in different parts and have the effect of switching on or turning off genes that cause differentiation. A similar process explains self-assembly in slime molds. Scarcity of food causes mold cells to emit molecules of a substance known as cAMP. When these molecules hit other cells, they increase their production. In time, a gradient field of cAMP concentration is produced with a spiral pattern. Cells are programmed to move toward the point of highest cAMP concentration and assemble at the center of the spiral. In both of these cases, a field produced by the joint activity of a system's components guides their individual behavior. Information about its collective state is trivially present in any system.¹⁰ Order emerges in a self-organizing system, however, because order parameters arising from the interaction between components and

¹⁰We can make do with a very thin notion of information here. A contains information about B relative to a range of contexts C, just in case A covaries with B within C. The configuration of planets carries information about the configuration of planets relative to all possible contexts; we can not, however, use information about the former to resolve ignorance about the latter.

containing information about collective variables are put in control of the states of components, setting up a synergetic relationship of continuous reciprocal causation between state of whole and states of parts.

In these cases, there is no pooling and integration of information, no inner locus of control or centralized computation that forms a master plan and coordinates the joint activity. We are tempted to think that there is something in the each half of the divided hydra that knows what it has on its undamaged end and administers the growth on the other. And we are tempted to think that there is something in the ant colony that is pooling the information collected by its minions and orchestrating the collective activity. But there is not.

Information and control in both systems are distributed. In a self-governing system, by contrast, there is a process that collects information distributed across the parts of the system,¹¹ re-represents it, combines with stored information, and subjects it to new forms of computation. In our ship, the information gathered by the crew in instrument readings and sightings does not directly regulate the ship's movements. Instead, it is used to transform a body of information already stored in the form of a map, and the computation that occurs in the space of the map results in a coordinated plan of action for the whole. In what follows, I shall argue that this difference has a point. Self-governance is less efficient than self-organization, but the extra layer of representational mediation carries an enhancement of dynamical capabilities that makes it worth the trouble.

Self-Governance vs. Self-Organization

First, however, we need an explicit characterization of the intuitive contrast between self-organization and self-government. We have seen examples of self-organization and a few more examples of self-government are helpful. Aside from the ship, there are missiles that navigate by onboard GPS, centrally controlled economies, robots with a centralized architecture, and armies that operate under central command. In each of these cases, the system steers by a model of the environment and in each case, the model serves as a kind of dynamical bridge between the components. The components are coupled to one another in virtue of being mutually coupled to the model. Note the importance of the coupling relationship: the model has to have a *continuing* role in the life of the system, it has to be bound up with the intrinsic dynamics of the system's parts in the physically inextricable way described above to be self-governing in an

¹¹And indeed, collects it so that it can be integrated. It is the additional power of computation over an integrated body of information that gives a self-governing system its advantage.

interesting way. Otherwise, we could just as easily give a description that separates controlling agent from the system under control and it would be a pragmatic matter whether we regard it as self-governed or governed from without. Think, here, of a clock and the clock-maker, or compare a world built in a single creative act and left to run on its own steam with one whose creator has a hand in the motion of every particle throughout its history, and whose own state is in its turn affected by those motions. It is only when the two are coupled in a continuing, reciprocal, causal relationship that separation of controller and controlled is impossible. In interesting examples of self-government, there *is* no way of decomposing the system into dynamically separable units.¹²

We could also view the components of a self-governing system as generating a kind of distributed self-model in the form of a field (material, chemical, or electromagnetic) whose local values contain information about the collective disposition of the system and control the behavior of components located in the region they cover. But this arrangement differs from self-government in that in the case of self-government, the information is collected, re-represented in a format that separates objective and self-locating information, and paves the way for new, more powerful forms of computation.¹³ Self-organization and self-government are two different ways of determining behavior. In both cases, information about the whole affects the behavior of parts. In the first case, the joint activity of parts generates a field that guides the behavior of components in a manner that is not mediated by centralized processing. In the second, information distributed across the components is collected and re-represented in a way that separates objective and self-locating information, allowing the system to make use of information stored in a context-independent format and exercise flexibility in choice of means. It is really the extra layer of processing that distinguishes self-governing from self-organizing systems. The arrangement is more complicated, but it allows the system to store and use information and gives it some flexibility in responses to the environment.¹⁴

¹²The new enactive theories of cognition criticize traditional theories of mind for failing to recognize this kind of inextricability between the human mind and body.

¹³For a more detailed account of self-location and the use of information stored in an objective form, see Ismael (2007).

¹⁴We can say all of this without using notions like “information,” “representation,” “decision,” or “reasoning” in anything but the sense in which they are standardly employed in computer science and which equals to computers, biological systems, word processors, and persons. These notions are indispensable in describing the operation of such systems, though they have proven contentious under analysis.

The difference between self-organization and self-governance is not one that appears if we look at the level of individual components of a system. It is a difference in how the system as a whole manages its activity. This high-level functional organization is stabilized out of a huge number of low-level microinteractions. One of the things we have come to appreciate in the study of complex systems is that this mid-level of description is where we get interesting differentiation. At the bottom level, all complex systems look the same. They are built of the same stuff, obeying the same laws. From too great a distance we see responses to circumstance but we do not see the interesting differences in internal processing that underwrite those responses. The difference between self-organization and self-governance is a mid-level difference in emergent organization for complex systems with large numbers of components.

The benefits of self-governance show up if we look at the second order dynamical properties of a system, specifically, at how its dispositions to respond to various stimuli from one moment to the next as its conception of its progress through the world evolves. Coordination among the components of a self-organizing system results from mutual involvement with a jointly generated field whose local values contain information about collective variables.¹⁵ Since both the field generated by the system and its effect on the behavior of components are fixed features of the system's design, and since the effect of stimulus on behavior is not mediated by combination with stored information or input from elsewhere, the equilibrium behavior is regular and predictable.¹⁶ From one to the next and over time, ant colonies, schools of fish, and traffic systems subject to the same external conditions behave in (more or less) the same way. The mediating computational cycle in a self-governing system, however, has the effect of decoupling behavior from the stimulus. Instead of varying directly with the stimulus, behavior is responsive to the system's assessment of its progress through an objective rendered landscape toward its goals. Since this is something that varies from one system to the next and from one moment to the next in the history of a single system, self-governing systems do not exhibit the regularity and predictability of self-organizing ones. What any self-governing system does from one moment to the next depends on a great deal more than the occurrent stimulus.

¹⁵The field metaphor works better in some cases than others; it is a picturesque way of expressing that information about order parameters is built into the local environment of the components of a system.

¹⁶It can take some time for a self-organizing system to settle into a stable state, so it is only the equilibrium behavior that is predictable.

It depends on what it has already done, what its goals are, and how it might reach them. It depends on everything the system believes about itself and the big wide world outside the current sensory horizon. No such thing is possible for a self-organizing system.

Making It Precise

We represent the first order dynamical properties of a system by a mapping from stimulus to response, which I shall call a response function. It is not that the response functions of self-organizing systems are perfectly fixed. The response functions of indeterministic systems, systems whose responses simply degrade because parts get worn out, and systems whose response functions change in accord with an innate program,¹⁷ for example, all change over time. What distinguishes these from self-governing systems is that the changes are not attuned to changes in circumstance.¹⁸ In addition to these, there is the more interesting class of nonself-governing systems whose response functions respond to conditioning. Let us call these adaptive systems. What distinguishes adaptive systems from those with genuinely *flexible* response functions is one's behavioral speed. The first order dynamical properties of a system whose behavior is regulated by a self-representational loop vary in real time as the system moves through the landscape, whereas the first order dynamical properties of a system that simply responds to conditioning evolve more slowly. We can say something about the conditions under which the first will hold an advantage over the second. If response functions are to be kept attuned to a particular type of contingency, they have to be regulated by soft structure, structure that can be adjusted as quickly as the contingency in question can change. Let us say that system's response function is attuned to a parameter P if it varies with variation in P. In systems in which adaptation is achieved by conditioning, it takes a period of specific, sustained pressure from the stimulus to induce changes in response function. This works fine if the lag time is tolerable and P changes slowly relative to the lag time, but it will not work if P changes quickly and adaptation needs to be spontaneous. For systems like ourselves who move without constraint, spatial location changes too rapidly and unpredictably to allow adaptation by conditioning. We could manage adaptation to slowly changing features of our situation,

¹⁷Think here of a frog zygote that starts out as a very simple system, but whose responses become increasingly complex at each stage of development.

¹⁸Except in the degenerate sense that they are attuned to the age of the system. In general, I use "circumstance" to refer to external parameters.

but adaptation to location itself, or features of the world that are tied closely to spatial location, demands explicit representation.

Summing up, then, self-governing systems are a subclass of complex, open systems characterized by an internal dynamics that incorporates a self-representational loop and supports flexibility of response function. They store information about themselves in an explicit form and combine that information with new input to compute the values of self-locating parameters. They form goals, which are used in conjunction with the values of self-locating parameters to regulate responses to stimuli, making fluid change in first-order dynamical properties possible in real time.¹⁹ They bear a heavy computational load, but there is a payoff. Stored information can be brought to bear on behavior in a manner that is regulated by the values of self-locating parameters. In cases of self-organization by contrast, although information about collective variables is made dynamically available to components, its bearing on components is fixed. We get coordinated behavior, but no flexibility in response function. The first order dynamical properties of such systems evolve slowly and recalcitrantly. There is a premium on flexibility for mobile systems. But the cost in computational complexity is high. Even in a self-governing system, only a small fraction of behavior is usefully passed through the self-representational loop. All of the examples of self-organization above—the ship, human bodies, centralized economies, or armies—are a complex mix of self-organization and self-governance. The needs of a system, together with the variability of its environment will dictate how fluid and responsive internal dynamics really needs to be. The difficult task in designing complex systems (e.g., economies, traffic systems, political organizations) is to find the balance of self-government and self-organization best suited to its needs and environment.

One very striking way to see that there is a real increase in the dynamical power of self-governing system is to notice that the complexity of the responses of a self-governing system, gauged by the number of degrees of freedom it exhibits, *outruns the complexity of the stimulus*. Herbert Simon had the insight that fueled much of the early research in situated cognition from watching ants. Although ants follow complex trajectories, the complexity comes entirely from outside; their intrinsic dynamics is as simple as can be.

¹⁹The first-order dynamical properties of a system are expressed by its response function, and adaptation has to do with whether and how the response function varies with the situation of the system. It is one thing to get behavior attuned to spatial location; Mother Nature does that by making sure changes in location are accompanied by changes in stimulus. It is quite another to get response functions so attuned. This requires a new dimension of variation over time.

While this kind of arrangement is efficient (it avoids reifying structure inside the ant by making use of structure in the environment), the complexity of the behavior it generates can not exceed that of the environmental input to the system. If the stimulus varies along three dimensions, so does the ant's response.²⁰ Not so for self-governing systems. One of the things that we know about persons, for example, is that they are wild cards. How they react to a given piece of information depends on so many internal variables that there is no effective way of predicting it with either precision or certainty. The existence of internal variables that mediate responses to stimuli is what characterizes flexibility.

From Self-Organization to Self-Government

The study of self-organizing systems has taught us that a collection of autonomous subsystems acting jointly but without supervision can achieve apparently purposeful and coordinated activity, and where we see such activity in the natural world, self-organization should be the default explanation. But self-governance involves a real departure from self-organization and brings with it genuinely new capacities. It brings the sort of flexibility that allows not just quick response, but immediate adaptation of first-order responses to the values of self-locating parameters. And it brings with it the capacity to carry out complex, temporally extended plans. These abilities require an extra layer of representational mediation between stimulus and response. Reflexive responses are excellent (much better than self-governors under many conditions) at guiding rapid motor behavior. Robots developed in Rodney Brooks' lab at MIT, for example, provide remarkable examples of self-organizing collections of reflex-driven subsystems that are extremely fast across a wide range of environments. What they are *not* is flexible at the level of response function and they cannot be programmed to carry out complex temporally extended plans without the addition of a self-representational loop.

A self-governing system represents the world in a form that is decoupled from its own location in it. Its location is represented by a free parameter

²⁰This holds only for deterministic systems. That, and the fact that there is no non-question-begging way of gauging the complexity of the stimulus, defeats behavioral complexity that outruns the complexity of the stimulus as a potential positive test for self-governance. This is a very difficult issue; we can not individuate stimuli simply in physical terms, some differences in patterns of activation of retinal cones, for example, do not correspond to visual differences. One wants to say that some criterion of internal availability should serve here, but the problem of providing a nonquestion-begging way of making out "internal availability" remains unsolved.

whose position relative to an identified goal is used to chart a course of action. The intuitive difference between a system that is designed to act in a manner conducive to the achievement of ends and one that explicitly represents its own ends and exercises choice in an effort to achieve them is reflected in a self-governing system by the extra layer of representational mediation between stimulus and response. This extra layer of representational mediation is what opens up the space for deliberation (or if deliberation suggests inference over propositionally structured, belief-like states, a precursor that captures the idea of instrumental rationality). The self-governing superloop that carries out this computation does not control all behavior. It has only a limited role, and only where flexibility of response function carries an advantage. It is most efficient when nested in a system in which day-to-day activity is regulated by self-organizing processes. Systems that employ self-government in any degree are special enough that they deserve to be distinguished from the general class and contrasted with systems whose behavior is the product of pure self-organization. But “self-government” and “self-organization” are really names for different strategies for behavior management; they can be used in conjunction and combine and interact in a manner that can produce a potent mix of top-down and bottom-up control.²¹ We can build self-government on a self-organizing foundation by adding a superloop of system-wide representation. But the superloop can exercise greater or lesser influence. Termite colonies, schools of fish, the free market, and anarchic societies provide examples of pure self-organization. A bureaucracy that aspires to having every last activity regulated by the central office provides an example of core self-governance at the opposite extreme.²² Along this spectrum there are many different and creative ways of combining self-governance and self-organization. Think of the subtle form of control exercised by the conductor in an orchestra. Or the trial-and-error process by which a dentist’s office or a small business finds a stable organizational structure. We struggle continuously to find the right balance of self-government and self-organization in social organizations, bureaucratic structures, and political institutions of all kinds.

²¹I leave it open whether we should think of systems that incorporate self-government as a special kind of self-organizing system, or reserve the term “self-organizing system” to refer only to those systems involving no admixture of self-government. This is a question about how to adjust the definitions, and the definition of “self-organization” is still up for grabs.

²²The most extreme case may be a doubtful possibility. It is likely that if any behaviors are centrally regulated, some behaviors have to be self-organizing, that is, that self-governance has to be built on a self-organizing foundation.

The self-representational loop carries a difference in functionality, but it also makes a difference to how we interact with a system by providing a localized interface for intervention. If I want to get you to perform a drawn-out action involving multiple body parts (e.g., to act as my *sous* chef or to pack a suitcase for me), I just need to address the central self-governor in the head. It will coordinate the movement for me. If the legislature wants Walmart to remove bullets from shelves, it does not have to go to each location separately. It orders the central office and lets the administration organize the action. There is no such localized interface with a self-organizing system. Influencing the behavior of an ant colony, a free market economy, or the direction of evolutionary development requires manipulating the values of distributed variables, difficult to identify and often impossible to control. The absence of centralized control is what makes guerrilla armies so hard to defeat.

There is a range of types of dynamical organization that combine self-governance with self-organization in different mixtures. Although there are many purely self-organizing systems, virtually all actual examples of self-governing systems incorporate some degree of self-organization. Self-organization is an extremely fast and efficient way of achieving coordinated behavior, but pressures toward centralization come from the benefits that attach to flexibility. Why does all of this matter? There is the ontological problem of finding room for selves in a world that is, at the bottom level, self-less. We could add them by hand, including them among the primitive constituents of the world. But the only real prospect for a *naturalistic, nonreductionist* understanding of selves is to see them as dynamically nondecomposable wholes built out of self-less constituents. But there is something else as well, something that pertains more directly to science; with its increasing understanding of the dynamics of complex systems, science is bumping increasingly up against the transition from self-organization to self-governance. That transition marks the shift from order that emerges from the bottom up and order imposed from the top down. It has mostly been avoided by a division of labor. There are those disciplines that deal with fundamentals (physics, chemistry, biology . . . the “hard sciences”), and for the systems that fall in their domains decomposability and self-organization is the norm. And then there are the humanities and “soft sciences,” which deal with systems of tremendous complexity. These disciplines routinely employ intentional vocabulary; they treat the systems that fall in their domain as deliberating agents, attributing to them knowledge some form of goal-oriented decision making. But it has long been the tacit conviction in the sciences that the humanities and social sciences would ultimately have to be absorbed into, or replaced by, the hard sciences. And to the extent that they were progressive, researchers

in these fields should seek self-organizing dynamical explanations, explanations that derive global order from the individualistic dynamics of components. The success of self-organizing models of certain social phenomena (smart crowds, economic behavior, game-theoretic treatments of social dynamics) reinforced this conviction.

If I am right, however, it would be a mistake to insist on self-organizing explanations of all phenomena at every level of complexity. We should acknowledge that there are two distinct sources of order in Nature, neither reducible to the other, constantly straining against one another. If it wants a complete picture, science has to make room both for order that emerges from the intrinsic dynamics of self-less systems, and order imposed from the top by the governing influence of self-representational loops. It should look more to the complex interplay between the two, in both natural and artificial agents, at every level of description, “from cells to societies.”

Dennett’s Challenge: Real and As-If Intentionality

I want to turn now to philosophical contexts on which the distinction between self-organization and self-governance can shed some light. Daniel Dennett has famously iconoclastic views about cognition. In modeling human intelligence, he switches freely between analogies to termite colonies and chess agents, all the while denying that there is any principled difference. In “Intentional Systems Theory” (Dennett 2009), he presents a taxonomy of explanatory stances. The first stance is the *physical stance*, in which one predicts the behavior of a system by using physical laws to evolve forward a description of its microstate. This is very precise but time-consuming and demands detailed knowledge. The second is the *design stance*, in which one assumes the system was designed to produce behavior with particular ends in mind and adduces predictions by looking for behavior conducive to those ends. So, for example, if you know the thermometer on your desktop was designed to measure temperature, without knowing anything about the physical principles that govern its functioning, you can predict that the pointer on the front will covary with ambient temperature (at least *ceteris paribus* in the circumstances for which it was designed to function). This is less time-consuming, but decidedly fallible. Finally there is the *intentional stance*, which is a subspecies of the design stance in which one treats the system as a deliberator, forming beliefs and making decision aimed at realizing goals of its own selection.

These three can be applied to any system, according to Dennett. We can adopt the physical, design, or intentional stances towards an amoeba, the immune system, or a crowd on a rampage. He regards the intentional stance

as useful (and in some cases, practically indispensable) in interacting complex systems, but denies that there is a distinction between systems to which intentional explanation applies literally and systems to which it applies in a mere metaphorical or *as-if* manner. In his view, the behavioral dispositions of a termite colony are given compact expression by treating it as a decision maker, and the same goes for human beings. To ascribe a person goals and see her behavior as a deliberate attempt to reach them in light of what he knows about the world is a compact and informative way of conveying information about a complicated set of interlocked behavioral dispositions, justified by its predictive leverage and neutral about the real mechanisms that underlie the behavioral dispositions. But it should not be understood as describing processes genuinely responsible for behavior. He writes:

The central epistemological claim of intentional systems theory is that when we treat each other as intentional systems, using attributions of beliefs and desires to govern our interactions and generate our anticipations, we are ...finessing our ignorance of the details of the processes going on in each other's skulls (and in our own!) And relying, unconsciously, on the fact that to a remarkably good first approximation, people are rational.²³ (Dennett 1987, 5)

The fundamental tenet of his view is that there is no dividing line along the spectrum from amoebas to persons at which this sort of description becomes literally applicable as a description of the processes that produce behavior. In his words,²⁴

There is a continuum of cases of legitimate attributions, with no theoretically motivated threshold distinguishing the "literal" from the "metaphorical" or merely "as if" cases. (Dennett 2009, 340)

The difference between my position and Dennett's can be put succinctly. I say that intentional description becomes literally applicable to self-governing

²³By "rational" here, I take it, he means exhibit instrumental rationality in pursuit of goals.

²⁴Dennett sometimes seems to have such a view in mind; it certainly fits the example of the chess agent that figures increasingly in his discussions of agency. But it would make intentional description inapplicable to his other examples (e.g., termite colonies) and go against the insistence that there is no principled difference in applicability of intentional description.

systems as a high-level description of processes occurring in the systems and responsible for the production of behavior. In self-organizing systems, by contrast, the link between stimulus and response is not mediated by anything that has the form of a deliberative process involving explicit representation of goals and means-end reasoning about how to achieve them.²⁵ To refuse to recognize this difference is to refuse to recognize a distinction that has practical as well as theoretical significance. Self-organizing systems do not exhibit the flexibility of deliberators, they do not adapt spontaneously to changes in circumstance, they do not have goals of their own, they cannot form complex temporally extended plans, and they are hard to interact with because there is no localized interface with the environment.

Dennett has wielded analogies with termite colonies heavily and somewhat confusingly in arguments about the explanation of human behavior. Many of his arguments trade on the fact that when we look closely at the brain, we do not find any central office that plays the role of the captain on a ship. But that observation, even if is correct, has no power to undermine the claim that persons are self-governing systems. To say that persons are self-governing systems is to say that the human body is guided by a central governing subsystem located in the head, one that collects information from the sensory surfaces, integrates it into an evolving self-model, and supports high-level computational processes that have the form of deliberation. There is no claim, implicit or otherwise, that the brain is a self-governing system, and no claim at all about how the low-level activity in the brain supports high-level computational processing.²⁶ Ultimately, at the most fine-grained level of description of any system (a government, a corporation, an army), we see only the decentralized activity of stupid particles responding immediately to their environment.

²⁵I am ignoring, here, other elements of intentional explanation, viz., that it has the forms of belief/desire psychology, where that involves propositionally structured states and logical inference. Self-governance is a prelinguistic and preintentional platform for more complex notions of intentionality that come with language. It marks a dividing line that answers only to some of the pretheoretic intuitions about a difference between real and merely metaphorical ascriptions of intentionality.

²⁶The relationship between the conscious mind and the low-level activity of the brain is an ill-understood and hotly contested matter. But it cannot get resolved in a manner that fails to recognize the integration and computation that form our conscious mental life. There is nothing in the neurophysiological evidence or in Dennett's own view of the mind that keeps us from viewing it as a virtual machine running on the hardware of the brain.

Insofar as intentional systems theory is an attempt to capture the cognitive competencies of different kinds of minds, the distinction between self-governance and self-organization is something to add to the arsenal. It is wholly in the spirit of that program, of a piece with a naturalistic attempt to gauge similarities and differences between ourselves and other beings. The distinction allows the intentional systems theorist to recognize pretheoretic intuitions about the distinction between real and as-if intentionality while rejecting magical elements in the explication of that difference. And it has foundational importance in social science, where it grounds intentional explanation for systems in the human and social sciences without making it anything but an inappropriate projection at the subhuman level.²⁷ So in many ways, my difference from Dennett is not that great. But there is one crucial disagreement. He moves from ant colonies to chess agents, drawing on the plausibility of the chess agent as a model for human cognition but appealing to the termite colony to argue that the appearance of deliberate goal-directedness is *mere* appearance. Whereas I say that there is a difference of significance here, one that carries with it differences in functionality (i.e., in the flexibility of behavior and susceptibility to localized intervention) and is directly relevant to our intuitions about the difference between real and as-if intentionality. It does not capture *everything* relevant to those intuitions. Self-governance is at best a precursor to these richer notions of intentionality. The sort of intentionality we possess is not magical, but it is complex and not well understood, embedded in a context in which there is language and culture and social infrastructure. It is linked to notions like intelligence and discursive rationality and often, somewhat mysteriously, to consciousness.²⁸ Self-governance is only one element in this matrix. The way to illuminate this matrix, however, is not by exploring intuitions, but by introducing some explicitly defined naturalistic distinctions whose applicability we can begin to investigate.

²⁷This is not to say that all human behavior should be explained in intentional terms. Much individual behavior is the result of self-organizing processes in the brain, and on the social level, intentional description of the dynamics of a free market economy, traffic system, or audience at a rock concert is only as explanatory as intentional description of termite behavior or the immune system. They may be useful, but should not be interpreted literally and realistically.

²⁸This is connected to the question that has arisen in the literature on group agency and group cognition about whether social collectives have Selves, whether they are Agents, or are properly attributed Beliefs and real intentional states. Self-governance has to be a precursor to a full understanding of all of these notions.

Self-Government and Selfhood

Perhaps the most important application of the distinction between self-organization and self-government is that it sheds some light on the emergence of a self.²⁹ Self-governance is built on top of a functional hierarchy that is self-organizing at lower levels by the addition of a superloop of system-wide representation that provides the setting for self-representation. The superloop, recall, consolidates information distributed across lower-level subsystems, re-represents it in a form that separates objective from self-locating information, and puts that information to use in a deliberative process that articulates goals and identifies means for achieving them. At the human level, this superloop of system-wide representation opens up the psychological space within which a personal point of view appears. In the most minimal sense, this point of view is a perceptual standpoint. Full development of the person requires awareness of self not just as a vantage point in space but as a thing with a history. And there are still richer conceptions of self; self as will, as social agent, as subject of autobiography, as personality and locus of value. The story of how these arise from the first glimmer of selfhood that comes when an agent begins to model her situation in the world in objective terms involves self-monitoring and metacognition, socialization, language, and culture. Somehow, in the hierarchy of self-regulating processes and self-regarding attitudes built on top of the self-governing base, a personality emerges with its hands on the reins, exercising a subtle form of top-down, fully intentional control over the movements of the body.³⁰

Crutchfield, in a recent discussion of emergence, describes the line of development in which emergent patterns in the global configuration of a system become available to an observer whose point of view spans those of the constituents (Crutchfield 2007). Such a man, as he says, has available to him information that is not available to any of the system's parts and could use that information to his own benefit. Someone watching emergent behavior patterns in an ant colony, for example, can apprehend those patterns and incorporate them into her own deliberation. I could let the colony lead me to the spilled sugar, for example, and use it to direct my cleaning activities. The global information about where the sugar lies does not inform the behavior of the individual ants heading toward it, however. They are responding blindly to reinforced pheromone trails left by ants who preceded them. There is no useful sense in which the individual ant possesses information about where the sugar lies. The

²⁹Here I sketch a line of thought that has been developed in more detail in Ismael (2007).

³⁰See Dennett (2003).

next step of this line of development is that in which the observer is incorporated into the system and the information embodied in emergent structure is actually used by the system to improve its own functionality. Self-governance shows us how to express this idea in a non-question-begging way by replacing reference to the “observer” with reference to a self-representational loop, that is, an internal deliberative process carried out under the auspices of an intentional standpoint, a personal-level, “I”-involving computational cycle.³¹

Self-governance involves the introduction of a self-modeler—an “autobiographer”—into a system. In the simplest cases, this is a matter of simple map-keeping and self-location. From there it is a not-in-principle-mysterious step to self-description in intentional terms. If I can describe other people’s behavior employing intentional vocabulary, I can certainly describe my own that way. The same goes for *metacognitive* attitudes. If I write down my thoughts, they can serve as premises in long chains of computation. They can become the intentional objects of further thoughts. I can think about them, compare them with other thoughts, check groups of them for consistency, and wonder whether they are correct. Metacognitive abilities come with an open-ended capacity to ascend to ever higher levels, and are bound up with a complex set of further capacities. Self-governance underwrites that ability; it is the platform on which it and the further capacities bound up with it are built.

The full story of the emergence of a personal point of view is a complicated one. It does not make the mistake that Dennett rightly rails against of building selves into the world as primitive constituents. It tells the story of how subpersonal activity gives rise to and supports the personal level of representation. In this story, a personal point of view emerges from subpersonal activity, but there is no occupant of that point of view in the form of a special object inside the system, seeing through its eyes and controlling its movements. In *Unweaving the Rainbow* Richard Dawkins writes,

The individual organism . . . is not fundamental to life, but something that emerges when genes, which at the beginning of evolution were separate, warring entities, gang together in co-operative groups as “selfish co-operators.” The individual organism is not exactly an illusion. It is too concrete for that. But it is a secondary, derived phenomenon, cobbled together as a consequence of the actions of fundamentally separate, even warring agents. . . . Perhaps the subjective “I,” the

³¹Note that that the global model does not need to take the form of a little “map in the head,” or a specially devoted subsystem. There are different ways in which this structure can be implemented.

person that I feel myself to be, is the same kind of semi-illusion . . . The subjective feeling of “somebody in there” may be a cobbled, emergent, semi-illusion analogous to the individual body emerging in evolution from the uneasy cooperation of genes. (Dawkins 1998, 308)

I think he is right to see the self as a secondary, derived phenomenon, analogous to the individual organism. But I do not think there is any reason to see either of these as illusory. There is not anything that is not in principle explicable in natural terms in this story. There are no free-floating Cartesian substances or internal homunculi, or *reses cogitans*. Persons arise at higher levels of organization through the interaction of huge numbers of unintelligent components. At the level of neurons, there is nothing but nerve cells sending and receiving signals, responding in a programmed ways to input. Higher functions emerge out of integrated patterns of signals. Intelligence is properly attributed to a system of interacting components, rather than to any thing or part.

Once self-governing systems have appeared in the natural landscape, they can band together into self-governing units regulated by rules of their own design. These larger collectives become objects of study in the social sciences, where the distinction between self-organization and self-government acquires a new importance, as marking a divide between different kinds of social organization. In one of his most evocative analogies, Dennett says that human intelligence must feature in the natural world as a crane, rather than a skyhook, meaning that once it has appeared on the landscape, human intelligence speeds up selection, lifting solutions out of design space, creating new forms of complex organization. But it has to have evolved naturally without any deliberate design as part of a continuous line of development by addition of something to a self-organizing foundation. Self-governance gives us a nonmagical account of how this could work.

Conclusion

Here, then, are the conclusions. The difference between self-governance and self-organization is a real, mid-level difference in the dynamical organization of certain kinds of complex system. It has both dynamical significance and philosophical importance. The dynamical significance has to do with the internal mechanisms that mediate stimulus and response, and the second-order dynamical differences between self-organizing and self-governing systems. The philosophical importance has to do with the fact that self-governance involves the creation of an internal point of view on the world and deliberative standpoint, supporting the literal applicability of intentional description and opening up the psychological space for the growth of a self.

Acknowledgment

This paper is dedicated to two social scientists from whom I learned most: Jacqueline and Tareq Ismael. Special thanks also to Phillip Petit. The manuscript was prepared under support from the Australian Research Council, QEII Fellowship. I am very grateful.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interests with respect to the authorship and/or publication of this article.

Funding

The author(s) received no financial support for the research and/or authorship of this article.

References

- Bar-Yam, Y. 1997. *Dynamics of complex systems*. Boulder, CO: Westview Press.
- Cowan, G., D. Pines, and D. Meltzer. 1999. *Complexity: Metaphors, models, and reality*. Boulder, CO: Westview Press.
- Crutchfield, J. P. 2007. Is anything ever new? Considering emergence. In *Emergence: Contemporary readings in philosophy and science*, edited by M. A. Bedau and P. Humphreys. London: MIT Press.
- Dawkins, R. 1998. *Unweaving the rainbow*. London: Penguin.
- Dennett, D. C. 1987. *The intentional stance*. Cambridge: The MIT Press.
- Dennett, D. C. 2003. *Freedom evolves*. London: Penguin.
- Dennett, D. C. 2009. *Intentional Systems Theory*, in Oxford Handbook of the Philosophy of Mind. McLaughlin, A. Beckerman, S. Walter (eds.), New York: Oxford University Press.
- Flake, G. W. 2000. *The computational beauty of nature: Computer explorations of fractals, chaos, complex systems, and adaptation*. Cambridge: The MIT Press.
- Haken, H. 2006. *Information and self-organization: A macroscopic approach to complex systems*. New York: Springer.
- Ismael, J. 2007. *The situated self*. Oxford, UK: Oxford University Press.
- Johnson S. 2001. *Emergence: The connected lives of ants, brains, cities, and software*. New York: Scribner.
- Nicolis, G., and I. Prigogine. 1989. *Exploring complexity: An introduction*. New York: W.H. Freeman & Company.
- Waldrop, M. M. 1992. *Complexity: The emerging science at the edge of order and chaos*. New York: Simon & Schuster.

Bio

J. T. Ismael is an ARC QEII Research Fellow at the Centre for Time at the University of Sydney and Associate Professor of Philosophy at the University of Arizona. He works in the philosophy of physics and philosophy of mind.