

## *Individuality, Subjectivity, and Minimal Cognition*

Peter Godfrey-Smith  
CUNY Graduate Center, and  
University of Sydney

October 2016 [G]. To appear in a special issue of  
*Biology and Philosophy* on biological individuality.

1. *Introduction*
2. *Self-Production*
3. *Multicellularity*
4. *Bodies*
5. *Unitary and Modular*
6. *Nerves*
7. *The Prehistory of Subjectivity*

### *1. Introduction*

The aim of this paper is to connect two projects discussed in different parts of philosophy and biology.<sup>1</sup> One project is understanding *individuality* in a biological sense – the sense seen in the discussion stemming from Leo Buss's book *The Evolution of Individuality* (1987) and related works.<sup>2</sup> The other is the attempt to describe the beginnings of mentality, the simplest forms, and to do so in a way that contributes to an understanding of the most puzzling features of the mind, subjective experience and consciousness.

---

<sup>1</sup> Thanks to Gáspár Jékely, Fred Keijzer, Jean-Francois Moreau, Maureen O'Malley, Thomas Pradeu, and Derek Skillings for comments on this material.

<sup>2</sup> See also Maynard Smith and Szathmáry (1995), and the papers in Bouchard and Huneman (2013). These discussions are much older within botany. My paper in the Bouchard and Huneman collection (2013) charts some of the history.

One link between the two topics is the idea of an *organism*. An organism is one kind of biological unit, one kind of "individual," and the clear cases of mental phenomena are the mental states *of organisms*, such as ourselves. Even if you think there is some contingency in that relationship, because of the possibility of minds in AI systems and nonliving robots, the nature of organisms certainly seems relevant to the problem. I think the link is tighter, too. With or without a biological framing, the right approach to the most elusive issues about the mind is by way of the notion of a *subject*, and subjectivity.<sup>3</sup> If so, part of what we have to understand is a certain kind of unit, a certain kind of constituent of the world: systems that are subjects. Setting aside dualist and panpsychist views, it seems that the special features of subjects are organizational, and this organization will be the product of some sort of evolutionary process. Biological organisms are such important evolutionary products that some connection between the evolution of subjectivity and the evolution of organisms seems inevitable. This does not make it clear *how* organisms and subjects are related – whether all organisms are subjects, whether subjects are a subset of organisms, whether subjectivity and organismality are matters of degree with some association between them... but biology certainly seems primed to do some of the work.

A literature that is especially relevant here is recent discussion of *minimal cognition*. How does the category of cognition relate to living activity generally? Do plants have cognitive capacities? Is all life cognitive to some degree?

I'll approach these questions in a way organized by the evolutionary history, especially by looking at some of the pivotal transitions. This paper – which is very much an exploratory one – will focus on the earlier stages; it won't discuss the special features of vertebrates, for example. Much of the discussion will be about the early evolution of animals, and the ways in which the evolution of animals produced a new kind of biological unit. I'll begin well before animals exist, though, and work my way forward.

---

<sup>3</sup> Thomas Nagel, a pessimist about biological approaches to the mind, concurs. "The main question, how anything in the world can have a subjective point of view, remains unanswered" (Nagel 1986, p. 30).

## 2. *Self-Production*

Some of what matters to the evolution of subjectivity appears in the context of unicellular, rather than multicellular, life. This is not because the same principles are visible at both stages and spatial scales; a message of the early part of this paper is that the opposite is true. But the place to start is with unicellular systems, especially prokaryotes.

*Life* in general I understand as a combination of metabolism-related and reproduction-related properties. These are tied together, but with some separability even in the normal course of events (viruses reproduce without having a metabolism). Corresponding to these two sets of capacities, metabolic and reproductive, are two kinds of units, *organisms* and *Darwinian individuals*. My focus in this paper will be mostly on organisms, living things in the metabolic sense, though the other side will sometimes be relevant as well.

Living things are often said to be "self-maintaining" systems. I think that in a broad sense, self-maintenance is indeed a general feature of organisms, but this notion covers different sorts of things at different scales. It is at the cell level that living things confront thermodynamic constraints in a direct way. Cellular life resists tendencies towards disorder, by using sources of energy and other resources from the cell's environment. Living systems are out of equilibrium with their surroundings, and retain that feature by constraining the movements of matter and energy with boundaries, especially membranes. Metabolic processes continually recreate both a pattern of organization and the system's distinctness from its surrounds. In this paper I use "metabolism" in a somewhat broader sense than is usual in biology; it refers not only to the handling of energy, but to the larger set of processes that maintain and continually rebuild organisms, contributing in turn to the distinct capacities of growth, development, and reproduction.<sup>4</sup> Where metabolism begins and ends will be discussed further below.

---

<sup>4</sup> I said above that I understand life in terms of a combination of metabolism-related and reproduction-related properties, and this passage suggests that growth and development are secondary. That is how I think of it, but this might be revised in the light of further analysis. In my (forthcoming a), reproduction is broken down into a combination of the more basic

A tradition of work that emphasizes these themes, as part of a treatment of both life and cognition, is the family of views based on the concept of *autopoiesis*, initiated by Humberto Maturana and Francisco Varela (1980), and continued by Evan Thompson (2007), Randy Beer (2014), and others. The term "autopoiesis" translates as *self-production*. A self-producing system is comprised of a characteristic set of parts and interactions, which continually re-make that very pattern of activity, in a way that includes the remaking of its boundaries, and typically includes the ongoing reconstruction of many of its other physical parts. Such a system is self-demarcating as well as self-maintaining, and persists in this activity through turnover in the matter that makes it up.

The autopoietic approach is perhaps the most worked-out general framework in this area, though it's part of a family of similar ideas, and a silence on these issues from some parts of biology may reflect a feeling that ordinary mechanistic descriptions of cells and metabolism are sufficient, without an abstract theoretical framing. I think the attempt to give a more abstract description is reasonable. The theoretical developments seen in the autopoietic tradition and its relatives include both good elements and bad ones. The idea I see as most useful is that initial idea of *self-production*, which goes beyond the more general idea of self-maintenance and also supersedes older attempts, with a more idealist and teleological flavor, to describe the special relations between parts and whole in living systems. The term "self-production" does have an initial air of paradox, but it can be understood in a straightforward way. A self-producing system's activities at one time are responsible for the persistence of the system through later times in a stronger sense than the sense applicable to most enduring physical objects (chairs, rocks), and a stronger sense than that applicable to artifacts like modern automobiles, which show some capacity to prevent their own breakdown, but do not *rebuild* themselves. Living systems at each time-step act to regenerate their characteristic and otherwise improbable patterns of activity, building many of the physical resources needed for this pattern to continue.

On the negative side, the literature in and around the autopoietic tradition has mishandled another aspect of life, the to-and-fro traffic characteristic of

---

phenomena of *recurrence* and *production*, and within an approach of that kind the roles of development and metabolism might be reconceived.

organism/environment relations. These writers have been determined to avoid imputing certain sorts of passivity to living systems. This may stem in part from their view that life and mind are strongly continuous, so views about the mind are shaped directly by what is said about life. The result is a view of life developed with much philosophical baggage, especially a kind of internalism. A recognition of the real phenomena of self-definition – boundary making – in living systems tends to be expressed in terms of "autonomy," and often includes denial of the essential role of *traffic* between any self-producing system and an environment distinct from it.<sup>5</sup> Recent developments of the theory have sometimes downplayed or modified this element, but the tradition remains resistant to the role of ecology, in a broad sense – resistant to the fact that is part of the nature of life to be in ongoing interaction with an environment that is *other*. Living systems, as "dissipative" systems, exist by means of transactions with a *milieu* that has different, and complementary, features; it must be a source of free energy and a sink for higher-entropy outputs. Cellular systems need to *have* boundaries, to prevent movement towards thermodynamic equilibrium, but also need to ensure traffic *across* them. I adopt the idea of self-production from the autopoietic approach, then, but reject the un-ecological side of the view, and don't think the way forward is to hold onto a notion of "autonomy." Traffic, input and output, are instead embraced.<sup>6</sup>

The evolution of cellular life itself gave rise to subject-like entities in one sense: bounded and self-maintaining units, engaged in traffic with their environments. What relation does these characteristics of cellular life have to "cognition"? Advocates of the autopoietic framework have often claimed that the relationship is very tight, and familiar

---

<sup>5</sup> "An autonomous system... defined by its endogenous, self-organizing and selfcontrolling dynamics, does not have inputs and outputs in the usual sense" (Thompson 2007, p. 43).

<sup>6</sup> The early discussions of autopoiesis did not make much use of thermodynamic ideas, which are surely immensely important in thinking about the boundedness and self-producing character of living systems. I have heard it suggested that this was due in part to a desire to keep the framework clear of the ideas of Ilya Prigogine, and other very ambitious uses of thermodynamics characteristic of that time. Recent work by those influenced by Maturana and Varela has in several cases moved away from a self-contained, un-ecological conception of organisms, but the retention of the notion of "autonomy" as a guiding concept seems to me to point away from the essential connection with thermodynamic concepts and the importance of organism/environment traffic.

or paradigm cases of cognitive activity are just more elaborate versions of something seen in all life: "Living systems are cognitive systems, and living as a process is a process of cognition" (Maturana and Varela 1980, p. 13). This view might initially seem to be supported by the fact that all known prokaryotic organisms have some capacity to sense external events and respond to them (Lyon 2015), but that is a striking empirical regularity, not something that seems to follow from the nature of life itself. Some recent work has looked for ways to mark more of a divide between living activity and the simplest kinds of cognition. I'll discuss two proposals in this paper. The first is due to Marc van Duijn, Fred Keijzer, and Daan Franken (2006, abbreviated "DKF" below). They see minimal cognition as "sensorimotor coordination." This concept, when unpacked, has two features. One is functional; minimal cognition adapts organisms to the *distribution of* metabolic resources, rather than aiding in the processing of those resources themselves. Second, the output of the system involves motion, either of parts or the whole, rather than biochemical change. Motility or physical manipulation of the environment is achieved. A paradigm case of minimal cognition for DKF is chemotaxis in *E. coli* bacteria.

A view of this kind might also be gestured towards by saying that cognition is the control of *behavior*, motion by unicellular organisms is the beginning of behavior, and in its simple cases this behavior adapts organisms to the distribution of resources (and the distribution of threats and other ecologically important entities). As DKF put it, chemotaxis has a "second-order" character, in relation to metabolism, as it is directed on the dispersal of chemicals in the environment rather than those chemicals themselves. One could also say that capacities like chemotaxis are "meta-metabolic," a term I'll use below.

The two features DKF associate with minimal cognition do pull against each other a little. Some control systems adapt unicellular organisms to the distribution of resources but do not include, as output, locomotion or anything similar. The *lac operon* system in *E. coli* and other bacteria is a control system by which an enzyme is produced only when a particular food source is around. DKF say this is not a system engaging in minimal cognition, as is not a case of sensory-*motor* activity; it is "still a part of the organism's metabolic organization" (p. 164). On the other hand, the *lac operon* system is second-

order or meta-metabolic, enabling adaptation to the environmental distribution of resources. A similar case is "quorum sensing" in *Vibrio* bacteria. This is clearly meta-metabolic, and social, too (on the standard interpretation of why it is present). But the output is chemical rather than motion-related.<sup>7</sup>

It seems that all motion-related capacities in these organisms are meta-metabolic, but not vice versa. So there are two divides: the advent of meta-metabolic activity, and the advent of motion as a means.

A different view of where cognition starts has been defended by Alvaro Moreno and his collaborators.<sup>8</sup> This view holds that cognition ("the cognitive domain") starts much later, with nervous systems. Organisms with nervous systems feature a different kind of relationship between their constituent parts from those without them. A nervous system is a control system that, given its role, must be partly "decoupled" from metabolic goings-on. I'll look at this proposal later, but an initial response that suggests itself, in the light of both this and the DKF proposal, is that a number of divides or steps are significant here, and we should not worry about whether a particular one marks a transition to *cognition*. Both sets of authors cited above do want to arm-wrestle a little: Arnellos and Moreno, defenders of the neural/non-neural divide, say that unicellular chemotaxis is just part of metabolism; it refines and extends the self-reproducing processes of the organism and is ultimately biochemical in nature (as DKF said the *lac operon* system is just metabolic). On the other side, DKF are firm that a nervous system is not needed for sensorimotor control, and suggest that "tissue characteristics" do not provide a plausible criterion for cognition. I'll return to this question below.

---

<sup>7</sup> Some of what counts as *behavior* in animals makes use of glandular secretions. When muscle contraction (coordinated motion of parts) is the means for this behavior, it does not require adding an extra category, but muscle contraction is not always the means (Jékely et al. 2015).

<sup>8</sup> See Arnellos and Moreno (2015, 2016), Moreno and Mossio (2015). Bich and Moreno (2015) sketch a somewhat different approach.

### 3. *Multicellularity*

So far we have bounded self-producing units engaged in traffic with their environments, where this traffic includes sensing and adaptive response. Some of their capacities can be seen as a simple form of cognition, and as units that sense and act, they are also subject-like, or subjects in a first, minimal sense. In the previous section I discussed mostly prokaryotic cells. Richer forms of sensing are seen in eukaryotic protists, along with a more elaborate internal skeleton, enabling manipulation of objects and new lifestyles (Jékely 2009, 2014). The transition I'll discuss in more detail, though, is the advent of multicellularity.

This transition is especially interesting because what results, in at least some cases, is a subject with psychological unity, but the parts that make up the system have subject-like features themselves. What we have to understand is not just how physical parts make up unified subjects, but how this results from the interaction of parts that are subject-like in their own right.

Arnellos and Moreno (2015) discuss a version of this problem, which they call the problem of *multicellular agency*. Problems about subjectivity are not the same as problems about agency, but they are related. Part of the transition to multicellularity is a transition in the locus of action, from the level of cells to the level of the larger organism. The micro-acts of cells come to realize the macro-acts of a larger unit. The same is true of sensing – the coordinated micro-sensing of cells gives rise to sensing at the level of the collective – and also of the processing that takes place between sensing and action. Agency, rather than subjectivity, is the concept Arnellos and Moreno focus on (as in Burge 2010), but agency in this sense does not only involve coordination on the output side. The problem of multicellular agency is akin to other problems of "collective agency," that arise in the case of actions by societies, nations, and committees (Pettit 2014). Before looking at the kinds of coordination that are relevant to multicellular agency and subjectivity, I'll look in a more general way at multicellularity itself.

Multicellularity has arisen on Earth perhaps a few dozen times (Grosberg and Strathman 2007). It may be as much as 2 billion years old, in the form of cyanobacteria (Schirmer et al 2011). Sometimes the results are morphologically simple –

branching filaments or clumps. But there are still several independent origins of "complex" multicellularity.

In the previous section's discussion of unicellular life, I emphasized the boundedness of those units, and the way that self-production includes the building of boundary structures, especially membranes. The transition to multicellularity gives rise to new higher-level units. These are also self-maintaining in a broad sense – they use sources of energy and other resources to maintain their structure, and also to develop and reproduce. But a difference is that multicellular organisms have vague boundaries with respect to which cells are parts of the organism and which are not. I think this is the right message to draw from recent work on symbioses between eukaryotic cells and their microbial partners, especially arguments for the "heterogeneity" of multicellular organisms with respect to the origins of their parts (Pradeu 2010).

An ideal case to illustrate the possibility of "heterogeneity" of this kind is that of *riftia* tubeworms (O'Malley 2014). In these deep-sea annelids, a digestive tract that is present at early stages is lost during development and replaced by a colony of environmentally-acquired symbiotic bacteria. In this case the symbionts are surely *part* of the organism itself, by any reasonable standard. The resulting collectives are organisms without being *Darwinian individuals* – without being reproducing entities. The Darwinian individuals are the eukaryotic part of the collective and the prokaryotic cells, separately.

The *riftia* case is an extreme illustration of something that seems to be common or ubiquitous, but with all degrees of tightness of engagement.<sup>9</sup> One might try to carve off a category of eukaryote-symbiont combinations that are so tight that they comprise organisms themselves, while others are treated as cases of collaboration between distinct organisms. The other option is to accept that it's characteristic of multicellular organisms that their boundaries are somewhat indefinite, because of the varied tightness of these collaborations.

I can only give inconclusive reasons to accept the second of these options. First, as O'Malley notes, theory in this area leads us to expect not only all grades of tightness in

---

<sup>9</sup> See also Skillings (forthcoming) on this issue, with particular reference to corals.

these relationships, but a "fluctuating tension" between host and endosymbiont fitness (2014, p. 131). Even in generally mutualistic cases, symbionts can have bad effects in some environments, and "parasites" (like *Wolbachia*) can be beneficial in some contexts. Pradeu seeks to use an immunological criterion, rather than vaguer ideas about cooperation, to re-sharpen organism boundaries; he thinks the boundaries of a multicellular organism do not lie where they are usually supposed to be, but they do exist (2010, 2012). Whatever is immunologically tolerated is part of the organism, regardless of its origin, and an immune response of some form is seen in all, or nearly all, of life. When microbial symbionts are tolerated, as they often are, they are part of the organism in the most literal sense. One reason I have for taking the other road on this question is that symbionts of this kind are accepted by Pradeu to retain a degree of immunological autonomy of their *own*; they are not entirely subsumed, becoming mere parts of a single larger immunological unit. Is it problematic for something that is part of one organism (in the strong sense) to also be an organism in its own right? There has not been a great deal of discussion of this question, as far as I know, but I think there is some tension in seeing symbionts as both organisms themselves and also parts of larger organisms.<sup>10</sup>

This seems far from a conclusive reason to reject the Pradeu view, but it gives some support to the alternative approach. My response to recent arguments about symbiosis, then, is to accept some revision of the old picture of the boundaries of multicellular organisms, in the way advocated by Pradeu, O'Malley, and others, but rather than looking for new boundaries, based on immunological considerations or a broader conception of a "metabolic whole," instead regard it as a feature of multicellularity that boundaries are typically unclear. I see this as a difference from the unicellular case; the essential thermodynamic work done by boundaries at the cell level does not have to be done again – or done with such sharpness – at the higher level.<sup>11</sup> In relation to the

---

<sup>10</sup> This is discussed in Godfrey-Smith (2013) as a possible "exclusion principle" applicable to organism status in part-whole hierarchies.

<sup>11</sup> Hooker (2009) also emphasizes that within an autopoietic framework, it's a wrong move to see the properties of multicellular organisms as some sort of recapitulation of the same properties that are important at the cellular level.

evolution of subject-like individuals, there is not a simple scaling between the relevant properties of cells and the relevant properties of multicellular collectives.

In thinking about individuality and boundaries in this context, a useful comparison can be made between ordinary multicellular organisms and microbial biofilms (Ereshefsky and Pedroso 2013). Biofilms can apparently show a fair amount of integration and division of labor, and are multi-origin collectives. But they are not *individuals* in the same way – and are only *objects* in a rather low-key sense. This is not just a point about the level of cooperation within them, but a point about what kinds of things they are – or, in a sense, what kind of stuff biofilm is. Biofilms do not have borders (though there can be specialized structures exposed to the external environment), and one biofilm can blend into another. There is, I take it, no definite count of biofilms in the world. They can exhibit coordination and signaling between their parts (Prindle et al. 2015), but those parts do not form determinate individuals. In the case of animals like us, on the other hand, the borders are somewhat vague, because of symbionts, but those vague borders are "centered" on something, and there is a fairly definite count of human beings in the world at any time (this is not true of all animals, as discussed in another section below). The cells in a biofilm interact extensively with their neighbors, and sometimes with more distant cells, but do not form a unit in the same way.

This point can be made semi-formally. In the case of an organism, there are relations between their parts that bind them, roughly speaking, into a single unit, separate from its surrounds. There is some organism-relevant relation between the parts that approximates being an "equivalence relation" (reflexive, symmetrical, and transitive). In the case of things like biofilms, plant-fungus mycorrhizal associations, and other open-ended neighbor-based collaborations, there is a role for *intransitive* relations that is not seen in the other cases – a role for situations where A has relation R to its neighbor B, and B has R to C, without A having R to C. Relations between cells may or may not give rise to a non-arbitrary higher-level unit, and if they *do* yield such a unit, there need not be sharp boundaries. What I'm describing here is not just the idea that organism-status is a gradient matter (as in Queller and Strassman 2009, O'Malley 2014), though I think it is. There are two points here: one is the role of gradients in the determination of which

systems are organism-like, and the other is the fact that even in the *paradigm* cases of organisms, like us, the boundaries are not definite.

I'll also make a brief comment about syncytia, multinucleate protists, heterokaryotic fungal mycelia, and the like.<sup>12</sup> These might appear to be counterexamples to my claim about the importance of boundaries at the unicellular level. However, as I understand these cases (and to the limited extent that I grasp their diversity at all), they are compatible with the principles described here. In these cases a membrane, exists that establishes the basic and thermodynamically important self/other relationships that are seen also in more "standard" cellular organization. That boundary encompasses within it nuclei and other material that would usually be found in cells of their own. Rather than having bounded cells and a more permissive collective of them, the collective (or quasi-collective) entity – the syncytium, or similar – does some of the boundary-related work that is usually done at the level of the parts. Cellular endosymbionts such as the precursors of our mitochondria are another special case. In the early days of the association between endosymbionts and host cells, there was a multicellular unit present – one that does not feature cells adjacent in space, but has some cells contained in others. Should that collective be seen as a multicellular organism? It's initially unclear, but rather than going on to produce more definite multicellularity, evolution instead produces a new kind of *unicellular* organism. (It moves from being a "collective reproducer" towards being a "simple reproducer," in my 2009 framework, and an analogous transition applies on the metabolic side).

#### 4. Bodies

The relations between cells and multicellular organisms are analogous to other cases of "collective action," as discussed above. But a multicellular organism is a special kind of collective agent, as its parts are physically connected, and laid out in space according to highly constrained relationships. Cells can only move far if the organism moves as a whole; it's not like a society or team, with independent movement of parts. This is a

---

<sup>12</sup> For the case of fungi, see Booth (2014).

general feature of multicellularity, but it has differences in form that are relevant to subjectivity and self/other relationships.<sup>13</sup>

In animals like us, there is an *interior milieu* that conditions the activities of the parts. In its original use by Claude Bernard, this phrase referred to a specific chemical environment. Here I broaden the term's meaning somewhat, as do Arnellos and Moreno (2016), to apply to an environment for a cell that is comprised of other cells and their products, an environment conditioned by the collective action of many other cells.

Multicellularity need not produce this feature; situations are possible with a clonally generated and connected collection of cells, where the environment of each is exposed to is the *external* environment, for the most part. The filament-like structures formed by various marine multicellular collectives are examples (Bonner 2000, Schirrmeister et al. 2011). In a collective of that kind, each part also encounters a fairly *similar* environment. In a model developed by Carl Schlichting (2003), also sketched in the 19th century by Herbert Spencer (1872), some degree of phenotypic complexity may arise "for free" in multicellular collectives of a different sort. In a multicellular object that is more "filled in," the environment encountered by the internal parts differs from the environment encountered by external parts. A uniform non-flat reaction norm characterizing each part (a reaction norm associated with their common genotype) then can lead to different morphologies across cells, as their micro-environments differ. That model assumes that the parts of the multicellular unit do face different environments, and filament morphologies don't have that feature. A fungal mycelium is also an exception. We can imagine another such case, an organism with a three-dimensional lattice-like form with the external environment interspersed throughout and no *interior milieu*. That case is imaginary, but it has some relation to the case of sponges (*Porifera*). In sponges, according to Keijzer and Arnellos (forthcoming), the water canal system through the body is so pervasive and generates such high rates of flow that it is more a continuation of the external environment than part of the organism itself. Many of a sponge's cells (not

---

<sup>13</sup> This spatial constraint might have exceptions, but all are in contested ground: ant and bee colonies have been seen as organisms, some modular organisms (see below) can have separated parts (are they still parts of one individual?) and bacteria of various kinds have been seen as multicellular even when they are not forming clumps or connected colonies.

all) are in contact with this canal system, and thus are exposed to a common external environment.

Other organisms have *bodies* in a stronger sense. There is not just a collection of cells that comprises them, but a more definite object with an interior milieu. This condition can be described semi-formally by saying that the parts of our bodies more closely approximate a *convex set* – a set of points is convex if whenever a straight line is drawn from one point in the set to another, the points on the line also fall in the set. Organisms with bodies do not have parts that actually form a convex set – that would require no appendages, and so on, but they more closely approximate it than a sponge or branching filament (on the other hand, a straight unbranched filament *is* convex in this sense). Bodies not only yield differentiation for free, they make possible elaborate developmental regulation, and other kinds of regulation, by the interaction of parts in a controlled medium. The history of the multicellular *body* is not the same as the history of multicellularity itself.

Keijzer and Arnellos emphasize the role of *epithelia* in the history of bodies in animals.<sup>14</sup> Epithelial tissue consists of layers of polarized cells, organized and anchored in an extracellular matrix, readily forming origami-like sheets (Tyler 2003). This organizational feature is unambiguously present in cnidarians. Sponges and cnidarians "arguably straddle the divide," they say, between animals with and without a range of important animal features – epithelia, bodies with an interior milieu, and others to be discussed below. Keijzer and Arnellos recognize that the split is not sharp, as recent work has argued that sponges do have epithelium-like structures, along with more behavioral capacity than had been supposed (Leys et al. 2009). The status of sponges also raises the case of *Placozoa*. These animals are mobile, unlike sponges, but small creepers who use cilia (and somewhat unorganized contraction) to move. They are also intermediate with respect to these body-related features, with some of the organized connectedness of cells seen in epithelia, but not the same bounding role. Sponges and placozoa are also the only

---

<sup>14</sup> The characterization above does not restrict bodies to animals. Keijzer and Arnellos, in contrast, focus entirely on animals; I am not sure whether bodies, in their intended sense, are present outside them.

animals who lack nervous systems (through their entire lifespan); the two non-neural animal phyla also share a partial quality in relation to the evolution of the body.

In the gulf between sponges and placozoa on one side and all other animals on the other, we have a package: full epithelia, muscle, nervous system, and also what Keijzer and Arnellos call a "standardized" body, one with a fixed shape across individuals (though I will return to this, with a complication, in the next section).<sup>15</sup> Keijzer and Arnellos suggest that a standardized body is important in the evolution of behavior, given the pivotal role of contraction: "fast and reversible" movements can only be organized in a reliable way "when they play out across a surface that remains stable in size, shape and extension across many life-time occurrences of these fast movements."<sup>16</sup> This package would have additional interest if it turns out, as has been recently argued, that ctenophores (comb jellies) are a sister group to all other animals, and sponges branched off after the split between ctenophores and the rest (see Figure 1). Then the package of features shared across ctenophores, cnidarians, and bilaterians would either have evolved twice, or been lost, as a package, probably twice, given that sponges and placozoans both lack it and they are not usually seen as sisters within an internal branch of the tree.

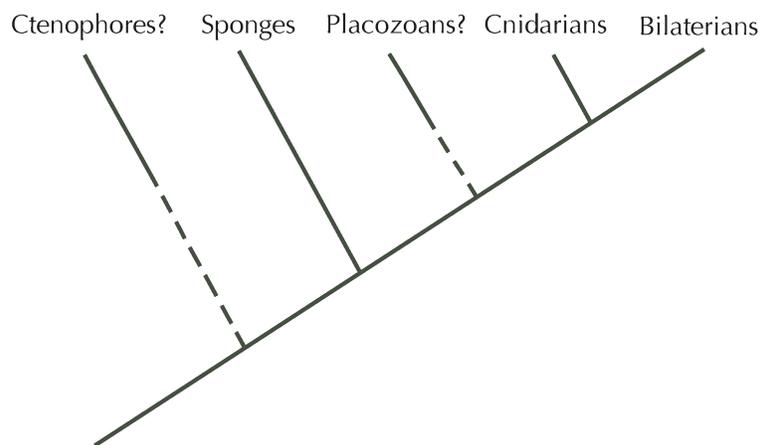


Figure 1: Phylogenetic relationships between some of the animals discussed. The locations of placozoans and ctenophores are uncertain.

---

<sup>15</sup> Sexual dimorphism is another qualification, but not a problem as there are two standard forms.

<sup>16</sup> Octopuses are an interesting special case. They have a fixed form in some senses (a fixed topology), but the overall observable shape is vastly changeable.

### 5. Unitary and Modular

So far we have multicellularity, bodies, and standardization of that form. What may come next is elaboration and larger size, but this can happen along two paths. One path is expansion and elaboration of the single coordinated unit. The other path is not to make the multicellular unit itself larger and more complex, but to bud off *more* units of that multicellular form, in an open-ended way and a way that preserves some degree of autonomy for the parts, metabolic and/or reproductive. This is the path taken by "modular" organisms. The other path, with a standardized adult body shape, a more constrained developmental process, and no asexual budding of similar forms that stay attached, is the "unitary" route.

A modular plan is seen in both animals (including bryozoans, some cnidarians, some ascidians), plants, fungi, and other taxa. The unitary/modular distinction is not entirely clear and sharp (and there is a grey area, as far as I can tell, between modularity and (one kind of) "colonial" organization).<sup>17</sup> The modular strategy has evolved in different versions repeatedly. The result is usually, though not always, sessile as an adult, and there is open-endedness of bodily form. The adult shape is often branched (see the previous section on the "bodylessness" of slender branched forms.) Keijzer and Arnellos made much of the invention of "standardized" bodies in the evolution of behavior; some lineages invented a standardized body and then submerged it, into an open-ended modular form.

In an evolutionary process heading towards more complex individual subjects, modularity of this kind is a step away. It produces instead an unusually cohesive social or colonial object. The results are akin to collective agents such as nations and committees, in roughly the way discussed above for the case of cells, but now the lower-level constituents are more complex. Physical connections between the parts of modular organisms are usually permanent, and coordination between them can be extensive. This is made possible by the way modules are produced – by asexual budding, which generates very high degrees of common interest, in the evolutionary sense. Especially in animals such as bryozoans, the modules have a good deal of individual cohesion

---

<sup>17</sup> Sponges are sometimes said to be modular organisms, though others deny this: Ereskovskii (2003).

themselves – each has a nervous system. Between-module and within-module connections are quite different. In plants, in contrast, fluid transport in a pipe-like system is the way space is bridged, and this includes the flow of hormones such as auxins. Then within-module and between-module traffic are not as different.<sup>18</sup>

In the divide between modular and unitary organisms, we see distinctions in the creation of action of a kind relevant to the evolution of cognition and subjectivity. Stepping back and laying down some generalities: there are *intra-organism* and *inter-organism* patterns of interaction, including those that give rise to collective-level actions. In all collectives, the behavior of the whole is in *some* sense the product of the behaviors of parts; macro-acts are products of many micro-acts. In some cases, though, a definite collective-level behavior comes to exist, as seen in human actions as they relate to the activities of cells. In other cases (a school of fish, a philosophy department) there is a less clear collective-level behavior. What is *done* by these various collectives arises from different relations between part-level and collective-level features. Collectives made of socially interacting organisms can do smart things, but the output has different relations to the coordination of parts from that seen in actions by a single multicellular organism. Rather than tight coordination between parts that each make small contributions, we have looser coordination between "smarter" parts, which each produce individually more complex actions.

If that is the general relationship between intra-organism and inter-organism coordination, modular and colonial animals probe at that relationship, showing features from each side. The zoids or modules have more autonomy than the parts of unitary organisms, but they condition the environment of other modules much more than distinct unitary organisms do to each other. Their degree of common interest from an evolutionary standpoint is higher than members of social group of organisms, but without the subsumed interests (more *joint* than *common*) of the parts of an organism with germ/soma separation.

A phenomenon that would be profitable to think about in detail here, given the focus of recent models, is signaling (Skyrms 2010, Godfrey-Smith 2014). As ordinarily

---

<sup>18</sup> Some think that plants are "post-cellular" in these respects. See the comments on syncytia and similar forms at the end of section 3.

described, signaling is seen in both inter-organism and intra-organism contexts; there is "signaling between people" and "signaling within the brain." They are different in many ways, though. In inter-organism signaling, two or more agents with some alignment of interests come together, and there is a coevolutionary shaping of "sending" and "receiving" rules. The intermediary, the sign or signal itself, will be cheap, passive, and manipulable; it is a go-between with a place in the lives of two distinct agents with their own goals. In intra-organism signaling, on the other hand, there is one agent with coordinated parts and little or no possibility of divergence of interests between those parts. A single evolutionary process shapes behaviors on both sides, sending and receiving, and the intermediate structure is yet another part of the same organism.<sup>19</sup> The intermediate can play an active role, can be much more than a go-between. The separation between sender, sign, and receiver then often tends to be blurred, so much so that the relations between parts do not fit the familiar schema of "signaling" very well. "Signaling in the brain" is not paradigmatically signaling at all, at least in the sense of recent models and accompanying theories (Cao 2012). Put a little over-strongly, if it really was *signaling*, it wouldn't be a *brain*.<sup>20</sup> In relation to these contrasts, communicative interaction between modules is a special case. The traffic between zoids or modules is to some extent akin to signaling between organisms rather than interaction within a single organism, especially when the parts retain significant reproductive autonomy.

More generally, the modular route has its own peculiar properties in relation to the evolution of subjectivity. Arnellos and Moreno (2015) take it to count against a Venus flytrap as "multicellular agent" that its traps operate independently of each other, with no coordination across the whole organism. That is true, but as this is a modular

---

<sup>19</sup> Here I discuss only interactions between spatial parts of an organism, not temporal stages. Things are interestingly different in the case of signaling across temporal gaps (also known as *memory*).

<sup>20</sup> See Levy (2011) for an analysis of information-based and communication-based talk in these parts of biology.

organism, the candidate "multicellular agents" are smaller than the whole plant; each stem is a better candidate for agency, and the stems are multicellular.<sup>21</sup>

Modularity is present in many groups, but unitariness in the clear sense, and in a large organism, seems to be primarily seen in animals (though the boundary is not clear – palms have been said to "behave like unitary organisms," diverging from the modular construction of most land plants).<sup>22</sup> In unitary organisms, there is a different level of interaction and coordination between the parts. Unitary animals also have a special relation to *movement*. Even if there is a lot of communication between modules, coherent macro-actions will be hard to evolve in an open-ended form of this kind (though some colonial salps and siphonophores can actively swim). Motility will move further into view in the final section below.

## 6. *Nerves*

We now reach the most obvious thing people talk about when considering the evolution of minds and behavior: the nervous system. A familiar picture might be summarized like this: animals evolved, in some of them a computer was installed on top (the brain), and neural computation then became the basis for the mental, including subjective experience. I don't say there is no truth in the familiar picture, but in this paper I'll look at nervous systems in a larger biological context, especially how they fit into the evolution of individuality.

The first question to ask is: what *is* a nervous system? This is not a straightforward question. Recent years have seen, for example, a sometimes testy debate about the idea of "plant neurobiology." Many biologists think the idea is absurd, but they

---

<sup>21</sup> Thanks to Derek Skillings for this point – in plants, "multicellular agents" can be smaller than the whole plant.

<sup>22</sup> The reasons palms are special are relevant here, too; see Avalos and Sylvester (2010): being "morphologically constrained by the lack of secondary meristems," "[t]o afford attaining canopy heights, woody palms need to show a high degree of phenotypic integration, shaping their growth and allometric relationships to match spatial and temporal changes in resources. Palms have a specific need for more integrated developmental processes.

don't often say what it would *be* for plants to have a nervous system.<sup>23</sup> A list of standard features of neurons can easily be given: electrical "excitability" (depolarization across a membrane, and transmission of action potentials), and chemical signaling between cells, at the synapse. Synapses are not always present in neurons, though, and it seems we want something a bit more abstract. Suppose the category *neuron* were to be understood purely in terms of a combination of functional properties: a cell is a neuron if and only if it is electrically excitable and can influence another cell's electrical activity, by means of chemical intermediaries or more direct effect.<sup>24</sup> If that is what a neuron is, then various organisms that are usually called "non-neural" do have neurons, including some plants.

Is that result bad? Maybe not. But suppose we tinkered with the definition a little. In Jékely, Keijzer, and Godfrey-Smith (2015), we opt for a definition that combines functional and morphological elements: a neuron is an electrically excitable cell that influences another cell by means of electrical or secretory mechanisms, and whose morphology includes specialized projections. A nervous *system* can be understood as a system made up (in part) of cells of this kind.

Neurons in this sense are restricted to animals, as far as we know, and they are found in all animals except for sponges, placozoa, and adult forms of some parasitic cnidarians (myxozoans, which are so reduced in form that they were once thought to be protists). The *neuron* in this sense seems in some ways a rather arbitrary kind, with its combination of functional and morphological features, but it will perhaps bear some weight in a story like this, because of what it makes possible.

In an animal with a genuine body (in the sense discussed above) but no nervous system, there can be neighbor-to-neighbor interaction between cells ("juxtacrine" influence) and broadcast of chemicals to whoever is listening. Coordinated contraction, the main effector mechanism of animals, can be controlled to some extent with neighbor-to-neighbor interactions, but this sort of coordination is limited. With a nervous system, on the other hand, one cell can influence other another cell at considerable distance, via

---

<sup>23</sup> See Alpi et al. (2007), especially, in which 35 plant biologists object to the idea of "plant neurobiology."

<sup>24</sup> This is a functional characterization that Jékely, Keijzer, and I considered when writing our joint paper (2015).

the axon, rather than by a chain of neighbor-based interactions (Nickel 2010). A cell can also integrate and compute, taking influence from a select set of cells, close or far away, via its dendrites. Proximity is not destiny, as far as influence goes. A new topology of cell-to-cell interaction is created, and the body is tied together in a new way.

To some extent, nervous systems don't *have* to work in this manner. There is an in-principle alternative, and one that has some degree of actual-world reality (Jékely forthcoming). In a nervous system like ours, interaction between cells involves a small number of chemical intermediaries, and a large role for physical targeting, for morphology. The same effects, as far as connectivity goes, might be achieved in the absence of the unusual morphology of neurons, if there was a very large number of chemical intermediaries – "signaling" molecules in a loose sense – that were broadcast widely, paired with very selective receptors in the receiver cells. Then a chemical broadcast from one location might affect only a single cell, some distance away, as only that cell is listening with the right receptor. This would be what Jékely calls a "chemical brain." Some parts of our nervous systems do work this way, he says – we have a bit of chemical brain – but in general we went the other way, with limited chemicals and much role for morphology.

To this I add that the "chemical brain" format is essentially how another family of control systems in our bodies operates – gene regulatory networks. There is broadcast of chemicals, with a great range available, and fine-tuned uptake and response. "Morphology" plays some role – packaging of DNA by histone molecules is a sort of molecular morphology that influences gene expression – but morphology plays nothing like the role it plays in the nervous system, with its elaborately specific connectivity and small number of intermediary molecules. Our way also enables more speed at large spatial scales; especially with myelinated axons, an action potential can cover significant space before an inter-cell junction must be confronted.

Our morpho-chemo-electrical brains achieve fine-grained targeting of influence. By bringing about targeted rapid influence of one cell on another, the nervous system generates a different sort of topology of interaction within a body. The nervous system is not just a biological computer, but also part of how the animal body becomes the distinctive kind of thing it is.

Thinking semi-formally once more, through a nervous system the patterns of influence between one cell and another, especially on rapid time-scales, become less dependent on neighborhoods. A nervous system makes possible short paths in a big body, and networks that freely link diverse regions.

With nervous systems we also reach the second of those views about the transition from metabolic activity to a "cognitive domain." Moreno (with various co-authors) argues that the cognitive domain begins with nervous systems, because this kind of control system has a novel relationship to the rest of what the body does. Neural activity controls what the body does while being partially "decoupled" from general metabolic goings-on. It could not perform its function if it were not decoupled in this way, if instead it was "governed by the general metabolic organization." It needs to have a kind of autonomy, needs to constitute "a new dynamic domain."<sup>25</sup> The "distinctness" of cognition from metabolism, the sense in which it is additional, comes from this decoupling.

Some pressure might be put on this view from recent work on the integration of neural activity with fine-grained features of the blood supply, the diffusion of small molecules in a way conditioned by a large range of cells, and other ways in which the "decoupling" seems less than may have once been envisaged (see, for example, Moore and Cao 2008). I don't know how those empirical discussions will turn out, and Moreno's view can certainly accommodate some influences of this kind, but I suggest that the contribution of nervous systems looks different in principle. What is special about nervous systems is not their partial remove from metabolism, but their changing the topology of (rapid) interaction within the body, through their unique combination of electrochemical and morphological features.

I've looked at two attempts to mark out a divide between the metabolic and the cognitive. The goal of marking such a divide seems reasonable, but the biology seems not to be cooperating very much. As one presses on the proposals, the tendency is not the

---

<sup>25</sup> The quotes in this paragraph are from Barandiaran and Moreno (2006), p. 176. Moreno and Mossio (2015) do discuss some of the features of nervous systems I emphasize here – the changes they make to patterns of connection within a body. They see these features as secondary to the "decoupling" that is treated as central, though.

emergence of a clear split, but another familiar outcome: a lot of relevant-looking steps, none that is quite what was envisaged. What makes this kind of distinction between metabolic and cognitive end up looking inapposite is, I think, the following combination of facts. First, basic metabolism involves a lot of quasi-computational control, especially in gene regulation. Those capacities are *also* the basis for early forms of adaptive flexibility based on sensing and responding to environmental opportunities and threats. Initially they do so in a context where "behavior" of a motion-based kind is very limited, and most output is chemical. But contraction-based behavior follows, in protists and then in animals. The evolution of those larger units create new challenges of coordination and control. Neurons then arise, and what makes them special is the combination of excitability and chemical signaling with a morphology that enables a new kind of integration of the body's activities.

### *7. The Prehistory of Subjectivity*

If the aim is to understand subjectivity in a rich, experiential sense, then this paper has been concerned with very early events – prehistory, most likely, rather than history. In this final section I'll summarize some of the ideas above and also look briefly ahead.<sup>26</sup>

The initial evolution of cellular life gave rise to subject-like entities in one sense – bounded and self-maintaining systems, engaged in traffic with their environment, and sensing and responding to events in adaptive ways. Some unicellular organisms also have genuine sensorimotor arcs, though those mechanisms have close relatives in forms of plasticity more directly involved in metabolism. From there, the evolution of the eukaryotic cell and then multicellular organisms created more complex units. The sensorimotor arc was "rediscovered" at the multicellular level. This requires the creation of coherent *action* in the new units, the coordination of the micro-acts of cells into macro-acts of a whole organism.

The creation of multicellular motion is a paradigm case. As emphasized by Jean-François Moreau, motion is a paradigmatic form of behavior that is *lost* at the cell level, in organisms like us, in order to make possible coherent motion by the collective. Most

---

<sup>26</sup> For further steps, see Godfrey-Smith (forthcoming b and c).

cells become fixed in place (in the adult), in order to achieve the fine coordination of contraction that motion in an animal like us requires. (Some cells of the immune system remain mobile, and are not restricted to the circulatory system but can actively climb through the extracellular matrix; they treat the rest of the body as environment: see Moreau et al., under review).

Once active motion by a multicellular organism is achieved, a combination of motility and good senses brings with it a relationship to the environment quite different from that seen in other organisms. This is especially because of *reafference* (von Holst and Mittelstaedt 1950, Merker 2005). "Reafference" is a general term for the effects of an organism's own actions on its senses. One role of reafference is to introduce ambiguity into sensory input: is this stimulus due to an external event, or to a change in relation to the environment stemming from my own actions? Reafference exists, in principle, in sessile organisms, but is more marked in mobile ones – as Björn Merker puts it, refferent effects are "liabilities of mobility" (2005). Animals have evolved a range of means for internal compensation for these effects, by tracking the likely sensory consequences of their own actions. These mechanisms are subjectivity-relevant because they give the organism the beginning of an internal registration of the self/other divide. Some animals achieve this through simple means, but even when simple, this set of relations between sensory and motor parts gives sensing a *shape* that is new, and is characteristic of unitary mobile animals.<sup>27</sup>

There are also other ways that a multicellular organization makes possible different kinds of relationships between an organism and its surrounds. An image-forming eye that can present objects, as opposed to a light-sensitive patch, must be spatially organized in a way that is almost (not quite) impossible without multicellularity.<sup>28</sup>

---

<sup>27</sup> Keijzer (2015) also argues that reafference, in early animals, may have been a resource, rather than a liability, in the task of sensing environmental structure. An array of active contractile tissue might function as a large-scale sensor, given that the consequences of its activity are conditioned by what is present in the animal's environment, in a way that may register in the ongoing activity of the array.

<sup>28</sup> *Almost impossible*: some protists seem to have image-forming eyes, and there is even a candidate in a cyanobacterium: Nilsson and Colley (2016).

I'll now start to look ahead to the consequences of these ideas for the explanation of subjective *experience*. Using the term "cognition" to refer very broadly to the side of the mind concerned with sensorimotor arcs, control of behavior, learning, and so on, the history of cognition can be drawn, in very simple form, in the way seen in Figure 2a below. That history probably begins early in the evolution of cellular life. From these simple beginnings, as time passes there is a spread in the range of values present. Steps or steepenings might be marked for the evolution of eukaryotes, nervous systems, and Cambrian animals, but I simplify the figure with a smoother function. Viruses and highly reduced cellular parasites are represented with a dip from the initial value. The figure is a first approximation, a cartoon. It represents cognitive complexity as ordered in a single dimension, though that is certainly not right. A figure like this can be useful, though, as way to mark contrasts with other possible shapes.

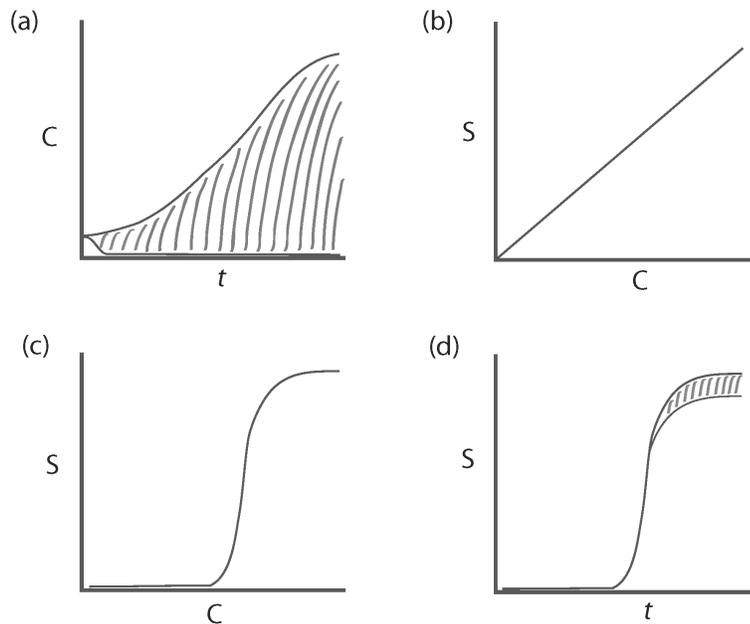


Figure 2: Possible functions relating evolutionary time to complexity of cognition (2a), complexity of cognition to sentience (2b and 2c), and time to sentience (2d, which assumes the applicability of 2c). See the main text for details.

What relation is there between this historical pattern and the evolution of subjective experience? By what function do these changes in cognitive complexity map

to changes in some sort of "complexity of subjectivity," or degree of consciousness? It is easy to think of cognition in a graded way, fading off into marginal forms seen in bacteria and the like, but tempting to think of subjective experience with an on/off model, as a yes-or-no matter. Many philosophical habits of discussion have that feature, but it may be quite misleading. The right model to use is unclear; does subjectivity fade out, with the same easing into marginal forms seen on the cognitive side? I don't know, but let's start with a simple option. Suppose that what we refer to as subjective experience, or the qualitative side of the mind, is just the cognitive side "from the inside." Then we can picture this with a simple linear function from cognitive complexity to the complexity of subjective experience (Figure 2b). For brevity I will use the term *sentience* (symbolized *S*) in the charts and the text below, to refer to this second side of the mind, and that term also encourages thinking in terms of degrees.

The history of sentience is then represented with a combination of those two charts, the first filtered through the second. We have a function for the history of cognitive complexity (2a), a function mapping cognitive complexity to sentience (2b), and those functions compose to yield the history of sentience or subjective experience. The "filter" in this case is trivial, though, and the history of sentience has the same shape as the history of cognition (2a). The first chart could be relabeled with "S" on the vertical axis.

Other views imply a less simple relation between cognition and sentience. They hold that within what can be broadly called "cognition," there is a subset of sentience-relevant capacities and features, perhaps a small subset (Dehaene 2014). A lot of cognitive processing has no subjectively experienced side, and in evolutionary terms, some late-arriving cognitive developments have special importance for the explanation of subjective experience. We can picture this with a very different function relating cognition to sentience, seen in Figure 2c. Then when we compose the two, mapping time to cognitive complexity and cognitive complexity to subjectivity, we see a threshold-crossing event, followed by an extended sliver (Figure 2d).

The two scenarios I sketched here are extreme options, and I'd expect the truth to be somewhere in between: only some cognitive capacities are associated with subjective experience, but these include some old inventions, such as the reafference compensation I

mentioned above. Another good candidate is the internal processing of *valence*, seen in internal reward systems and instrumental learning. Various other kinds of cognitive capacity are probably less subjectivity-relevant.

I'll make two quite different points about subjects to close. First, subjects and organisms have another link to questions about "individuality," now in a much broader sense. In recent metaphysics, some people have wondered which familiar macroscopic things around us are genuine *material objects*, and which do not, strictly speaking, have that status. Peter van Inwagen (1990) argued that only two sorts of things are genuine material beings, (i) the elementary, indivisible physical objects that make up the universe, and (ii) living organisms. He thinks that most "objects" we talk about, such as tables and chairs, are arbitrary with respect to their separation from their surroundings, and with respect to their beginnings and ends in time, and this makes them so deficient as *objects* that we should not accept that they are real. I think the negative claims van Inwagen makes about non-living macroscopic objects are excessive, and I don't think there's much point in denying that the moon is a material being. But *if* one is concerned with questions about boundaries as marks of reality, then it is true that living things are special. The categories "living things" and "organism" are a bit broad for these purposes, though, as we've seen in this paper. The origin of cellular life did bring a new kind of entity into the world, things that have non-arbitrary boundaries because they make and maintain them, with turnover in materials as time passes. One result of the origin of life was the first appearance of objects with a different kind of distinctness *as* objects. Multicellular organisms are more permissive about which cells comprise them, and that's a partial step "back" with respect to definiteness at the macro-scale. Animals with nervous systems are unified and coherent entities in another way; they comport themselves *as* wholes (just as cells "insist" on their reality as objects by self-production). The point could be put by saying that subjects are among the better objects, too.

Lastly, the bulk of this paper has been about how subjects came to exist – how subjects figure as effects of evolution. But subjects also make things happen. Some systems have effects of a certain kind *because* they are subjects; subjectivity has a causal role. This includes effects on large scales in space and time, including effects on what evolution produces. This is seen especially in the Cambrian. That period saw the

evolution of new kinds of bodies, due, at least in part, to the coevolution of bodies, nervous systems, and behavior. New kinds of animal bodies arose because of the subject-relevant behaviors of other animals – their ability to sense you, track you, and manipulate you, especially as prey. Once actions of those kinds are part of the environment, new evolutionary paths are taken. Such innovations come, in time, to have effects not only on the other organisms they immediately interact with, but on the larger ecology – on how resources are distributed, on what food webs look like, on how carbon and nitrogen are cycled around. Before long, the products of these coevolutionary processes came onto land, and began doing all the things animals like us do there. People have often worried that "consciousness" is epiphenomenal, lacking any role as cause. This question can seem acute when consciousness is understood in terms of "qualia," the "raw feels" that seem to populate our minds. But qualia are an aspect of subjectivity, of what goes on in a subject, and subjectivity is not epiphenomenal. On Earth, it is about the least epiphenomenal thing there is.

---

### *References*

- Alpi A and 34 others (2007). Plant neurobiology: no brain, no gain? *Trends in Plant Science* 12: 135– 136.
- Arnellos A, Moreno A (2015). Multicellular agency: an organizational view. *Biology and Philosophy* 30: 333–357
- Arnellos A, Moreno A (2016). Integrating Constitution and Interaction in the Transition from Unicellular to Multicellular Organisms. In *Multicellularity: Origins and Evolution*, edited by K. Niklas, S. Newman, J. Bonner. Cambridge MA: MIT Press.
- Avalos G, Sylvester O (2010). Allometric estimation of total leaf area in the neotropical palm *Euterpe oleracea* at La Selva, Costa Rica. *Trees* 24: 969-974.
- Barandiaran X, Moreno A (2006). On what makes certain dynamical systems cognitive: A minimally cognitive organization program. *Adaptive Behavior* 14: 171–185.

- Beer, R (2014). The cognitive domain of a glider in the game of life. *Artificial Life* 20: 183–206.
- Bich L, Moreno A (2015). The role of regulation in the origin and synthetic modelling of minimal cognition. *BioSystems*, <http://dx.doi.org/10.1016/j.biosystems.2015.08.002>
- Bonner, J (2000). *First Signals: The Evolution of Multicellular Development*. Princeton: Princeton University Press.
- Booth A (2014). Populations and individuals in heterokaryotic fungi: A multilevel perspective. *Philosophy of Science* 81: 612-632.
- Bouchard, F and Huneman, P (eds.) (2013). *From Groups to Individuals: Perspectives on Biological Associations and Emerging Individuality*. Cambridge MA: MIT Press.
- Burge T (2010). *The Origins of Objectivity*. Oxford: Oxford University Press.
- Buss L (1987). *The Evolution of Individuality*. Princeton: Princeton University Press.
- Cao R (2012). A teleosemantic approach to information in the brain. *Biology and Philosophy* 27:49–71.
- Dehaene S. (2014). *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts*. New York: Penguin Random House.
- Ereshefsky M, Pedroso M (2013). Biological individuality: the case of biofilms. *Biology and Philosophy* 28: 331–349
- Ereskovskii A (2003). Problems of coloniality, modularity, and individuality in sponges and special features of their morphogeneses during growth and asexual reproduction. *Russian Journal of Marine Biology* 29, Suppl. 1: S46–S56.
- Godfrey-Smith P (2013) Darwinian individuals. In F. Bouchard and P. Huneman (eds.), *From Groups to Individuals: Perspectives on Biological Associations and Emerging Individuality*. Cambridge MA: MIT Press, 2013, pp. 17-36.
- Godfrey-Smith P (2014). Sender-receiver systems within and between organisms. *Philosophy of Science* 81: 866-878.
- Godfrey-Smith P (forthcoming a). Complex life cycles and the evolutionary process. *Philosophy of Science*.
- Godfrey-Smith P (forthcoming b). Evolving across the explanatory gap.
- Godfrey-Smith P (forthcoming c). The subject as cause and effect of evolution.

- Grosberg R, Strathman R (2007). The evolution of multicellularity: A minor major transition? *Annual Review of Ecology, Evolution, and Systematics* 38: 621–54.
- Holst E von, Mittelstaedt H (1950). The reafference principle (interaction between the central nervous system and the periphery. Reprinted in *The Behavioural Physiology of Animals and Man: The Collected Papers of Erich von Holst*, vol. 1, trans. R. Martin. Coral Gables, FL: University of Miami Press, 1973, pp. 139–73.
- Hooker CA (2009). Interaction and bio-cognitive order. *Synthese* 166: 513–546.
- Jékely G (2009). Evolution of phototaxis. *Phil. Trans. R. Soc. B* 364: 2795–2808.
- Jékely G (2014). Origin and evolution of the self-organizing cytoskeleton in the network of eukaryotic organelles. *Cold Spring Harbour Perspectives in Biology* 2014;6:a016030.
- Jékely G (forthcoming). The chemical brain.
- Jékely G, Keijzer F, Godfrey-Smith P (2015). An option space for early neural evolution. *Philosophical Transactions of the Royal Society B* 370(1684), 1–10.
- Keijzer F (2015). Moving and sensing without input and output: early nervous systems and the origins of the animal sensorimotor organization. *Biology and Philosophy* 30: 311–331
- Keijzer F, Arnellos A (forthcoming). The animal sensorimotor organization: A challenge for the environmental complexity thesis.
- Levy A (2011). Information in biology: A fictionalist account. *Noûs* 45: 640–657.
- Leys SP, Nichols SA, Adams ED (2009). Epithelia and integration in sponges. *Integrative and Comparative Biology* 49: 167–177.
- Lyon P (2015). The cognitive cell: bacterial behavior reconsidered. *Frontiers in Microbiology* 6:264. doi: 10.3389/fmicb.2015.00264
- Maturana H, Varela F (1980). *Autopoiesis and Cognition: The Realization of the Living*. Boston Studies in the Philosophy of Science, vol. 43. Dordrecht: Reidel.
- Maynard Smith J, Szathmáry E (1995). *The Major Transitions in Evolution*. Oxford: Oxford University Press.
- Merker B (2005). The liabilities of mobility: A selection pressure for the transition to consciousness in animal evolution. *Consciousness and Cognition* 14: 89–114.

- Moore, C and Cao, R (2008). The hemo-neural hypothesis: On the role of blood flow in information processing. *Journal of Neurophysiology* 99: 2035-2047.
- Moreau J-F, Pradeu T, Grignolio C, Castiglione F, Tieri P, Capri M, Salviolo S, Taupin J-L, Garagnani O, Franceschi C (forthcoming). The re-emerging role of ECM crosslinking in T cell mobility as a hallmark of immunosenescence (under review).
- Moreno A, Mossio M (2015). *Biological Autonomy: A Philosophical and Theoretical Enquiry*. Dordrecht: Springer
- Nagel T (1986). *The View from Nowhere*. Oxford: Oxford University Press.
- Nickel M (2010). Evolutionary emergence of synaptic nervous systems: what can we learn from the non-synaptic, nerveless Porifera? *Invertebrate Biology* 129(1): 1–16.
- Nilsson, D-E, Colley N (2016). Comparative Vision: Can Bacteria Really See? *Current Biology* 26: R369–R371
- O'Malley M (2014). *Philosophy of Microbiology*. Cambridge: Cambridge University Press.
- Pettit P (2014). Group agents are not expressive, pragmatic or theoretical fictions. *Erkenntnis* 79 (Supplement 9): 1641–1662
- Pradeu T (2010). A mixed self: The role of symbiosis in development. *Biological Theory* 6(1): 80-88.
- Pradeu T (2012). *The Limits of the Self: Immunology and Biological Identity*. New York: Oxford University Press
- Prindle A, Liu J, Asally M, Ly S, Garcia-Ojalvo J, Süel G (2015). Ion channels enable electrical communication in bacterial communities. *Nature* 527: 59–63  
doi:10.1038/nature15709.
- Queller, DC, Strassman J (2009). Beyond society: the evolution of organismality. *Philosophical Transactions of the Royal Society B* 364: 3143–3155.
- Schirrmester B, Antonelli A, Bagheri H (2011). The origin of multicellularity in cyanobacteria. *BMC Evolutionary Biology* 11:45.
- Schlichting CA (2003). Origins of differentiation via phenotypic plasticity. *Evolution and Development* 5: 98–105.
- Skillings D (forthcoming). Holobionts and the ecology of organisms – From communities to integrated individuals. To appear in *Biology and Philosophy*.

- Spencer H (1872). *First Principles of a New System of Philosophy*, 2nd ed. New York: Appleton.
- Skyrms, B (2010). *Signals: Evolution, Learning, and Information*. Oxford: Oxford University Press.
- Thompson E (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge, MA: Belknap Press.
- Tyler S (2003). Epithelium—the primary building block for metazoan complexity. *Integrative and Comparative Biology* 43(1):55-63
- Van Duijn, M, Keijzer, F, and Franken, D (2006). Principles of minimal cognition: Casting cognition as sensorimotor coordination. *Adaptive Behavior* 14: 157–170.
- Van Inwagen P (1990). *Material Beings*. Ithaca: Cornell University Press.