MINIMAL MINDS

Bryce Huebner, Department of Philosophy, Georgetown University

> All the virus wants to do is break inside a healthy cell, steal its genetic machinery, and start profiting from the intrusion. To stop a thief, you need to throw a monkey wrench—or several—into his plans. That's exactly what anti-HIV drugs, known as antiretrovirals (ARVs) are designed to do.[1]

We sometimes speak, in a "loose and popular" way, as though viruses have human-like beliefs and desires. It comes as no surprise when such intentional idioms occur in the hyperbolic prose of the mainstream media. However, it is more surprising when scientific discourse relies on anthropomorphizing claims when discussing the nature of viruses. In recent papers, it has been suggested that if the HIV "virus wants to resist AZT, it needs to make a specific mutation at codon 215 and at another position, such as position 41"[2] and that the Hepatitis C virus must try to coexist with its host by reducing its visibility.[3] Yet, philosophical reflection suggests that although we *can* use intentional idioms whenever we find an entity that exhibits robust and systematic patterns of behavior, a plausible theory of mentality should not treat viruses as even minimally minded. After all, viruses are nothing more than packets of RNA (or sometimes DNA) encased in protein. So, although they are likely to be mechanical and complex, the type of complexity that we find in viruses is unlikely to implement the computational and representational capacities that are necessary for being in cognitive states or carrying out cognitive processes. Viruses are not the kind of entities that can *want to* threaten our wellbeing; they do not *make plans* and *adopt deceptive strategies* for navigating our immune systems; and they have not *learned* to outsmart our best anti-virals. Although many viruses exhibit systematic and predictable patterns of behavior, we can and should treat them as complex biological robots.[4] However, difficult theoretical issues arise in providing a more general strategy for distinguishing genuine mentality from witless mechanical behavior.

In this chapter, my goal is to establish the *minimal conditions* that must be satisfied for an entity to be in cognitive states or carry out cognitive processes. I begin with an examination of the philosophical assumptions that underlie the Cartesian assertion that linguistic capacities are necessary for genuine mentality. I argue that the explanatory resources provided by contemporary cognitive science undercut the force of these assumptions and provide compelling reasons for thinking that even invertebrates (e.g., crickets, cockroaches, termites, and honeybees) possess genuine cognitive capacities. While this claim may seem familiar, I extend the arguments for treating invertebrates as possessing *minimal minds* in two ways. First, I argue that minimal mindedness is likely to be realized by computations that operate over non-linguistic representations; thus, minimal cognitive states and processes are unlikely to be beliefs or desires. Second, I argue that the theoretical and empirical evidence that supports treating invertebrates as minimally minded also warrants the more intriguing conclusion that some sorts of *collective behavior* are best explained in terms of collective mental states and processes;

---

[1] Alice Park, "Beefing Up the Arsenal Against AIDS," *Time Magazine* (2007), http://www.time.com/time/health/article/0,8599,1595377,00.html (accessed August 9, 2010)

[2] David Ho, "Therapy of HIV Infections: Problems and Prospects," Bulletin of the New York. Academy of Medicine, 73 (1996), pp. 37.

[3] Vito Racanelli and Barbara Reherman, "Hepatitis C Virus: Infection When Silence is Deception," *Trends in Imunology* 24 (2003): 456-64.

[4] Cf. Daniel C. Dennett, *Kinds of Minds* (New York: Basic Books, 1996), pp. 21ff.

hence, I argue that if invertebrates have minimal minds, then so do some *groups* of invertebrates.

1. MINDS LIKE THE HUMAN MIND

Our commonsense strategies for making claims about the mental lives of other entities are a heterogeneous mixture of genuine ascriptions of mental states, ungrounded anthropomorphic assertions, and idiosyncratic expressions of unjustified theories. This makes it difficult to know how to proceed in addressing ontological and epistemological questions about other kinds of minds. As Bertrand Russell rightly notes:

> We observe in ourselves such occurrences as remembering, reasoning, feeling pleasure, and feeling pain. We think that sticks and stones do not have these experiences, but that other people do. Most of us have no doubt that the higher animals feel pleasure and pain, though I was once assured by a fisherman that "Fish have no sense nor feeling." I failed to find out how he had acquired this knowledge. Most people would disagree with him, but would be doubtful about oysters and starfish. However this may be, common sense admits an increasing doubtfulness as we descend in the animal kingdom, but as regards human beings it admits no doubt.[5]

However, the problem is not merely that we become less sure about the existence of other kinds of minds "as we descend in the animal kingdom." More importantly, the diverse range of strategies that we deploy in ascribing mental states and processes also yields deep theoretical disagreements across philosophical traditions and scientific methodologies. Indeed, there is no consensus about how to draw the distinction between genuine mentality and witless mechanical behavior.

Even if we must begin from commonsense presumptions about what minds really are, a philosophically respectable theory of the nature of the mind must also triangulate these commonsense presumptions against a wide variety of scientific data and philosophical theories about the nature of mental states and processes.[6] Where commonsense intuitions about some cognitive states or processes conflict with a philosophical or scientific theory, these intuitions provide defeasible evidence against the truth of that theory. However, commonsense ascriptions of mental states and processes are also anthropomorphic in a way that licenses treating any entity that behaves in an apparently intentional manner "as if they were just like us—which of course they are not."[7] So, developing an adequate understanding of what it means to be minded, and taking the results of scientific inquiry seriously, may sometimes require deep revisions to our commonsense understanding of mentality.

To make this point clear, consider a case from cognitive ethology.[8] Marc Hauser once observed a dominant male macaque who attacked a subordinate male for mating with a high-ranking female. Hauser reports that the subordinate male remained quiet as the dominant monkey ripped off one of his testicles, never even exhibiting a facial expression indicative of pain. More surprisingly, the subordinate was again attempting to mate with a high-ranking female after little more than an hour. Although we readily assume that

[5] Bertrand Russell, *Human Knowledge: Its Scope and Limits* (London: Allen & Unwin, 1948), p. 486.
[6] Cf., José Luis Bermúdez, *Philosophy of Psychology: A Contemporary Introduction* (London, Routledge, 2005), chapter 1.
[7] Dennett, *Kinds of Minds*, p. 33
[8] Marc D. Hauser, *Wild Minds* (New York: Henry Holt, 2000), p. 9.

monkeys experience pain in some way that is relevantly similar to the pain felt by humans, this case suggests that we should not assume that this is the case without further evidence to support this intuition. In short, a plausible account of macaque pain must draw upon commonsense intuitions about what pain is, but it must then utilize careful ethological observations, neurological data, and evolutionarily informed theories about the relationship of these neural structures to those responsible for implementing the human experience of pain.

Still, Russell is right to note that our minds are the only ones about which we have knowledge first-hand. Whenever we ask whether an invertebrate feels the sort of pain that warrants our compassion, or remembers where it has stored food, we are asking—even if only implicitly—whether it has memories or feels pains that are like our own in some significant respect; and this fact opens up a philosophical puzzle about the nature of minimal minds: in what respect can an invertebrate (e.g., a fly, a honeybee, or a spider) be in mental states or carry out mental processes that are enough like ours to justify the claiming that it has a mind?

One answer to this question adopts the conservative suggestion that *genuine* mentality requires the capacity for linguistic representation. On this view, invertebrates cannot be in mental states or carry out mental processes, and neither can cats or dogs. This may initially seem implausible. However, the presence of robust linguistic capacities in Homo sapiens yields an enormous increase in behavioral flexibility and cognitive sophistication well beyond what can be exhibited by any non-human animal. Although there is little consensus about the precise role that is played by language in human cognition, maximal minds like our own exhibit numerous cognitive capacities that are deeply embedded in linguistic and social practices. At bare minimum, the capacity to use language augments and transforms the cognitive strategies that are employed to remember the past, coordinate future actions (both individually and collectively), deliberate about counterfactual possibilities, and form fine-grained beliefs and desires about the world in which we live.[9] Moreover, language underwrites our capacities for reasoning together and engaging in collaborative activities in ways that are not likely to be present among the members of any other species.

For these reasons, our ascriptions of mental states and processes to other humans also rely on a practice of cooperative mind-shaping that allows us to establish and maintain strategies for cooperative engagements, a practice that we cannot adopt in making sense of non-linguistic minds.[10] Consider the mental states and processes that are ascribed during a collaborative writing project. Such writing often takes place without considering what a collaborator believes or desires; but, where a branchpoint suggests no clear agreement about how to proceed, it is often necessary to project the sorts of beliefs and desires that a collaborator is likely to have formed in light of the way that the world has been disclosed to her. For example, in the face of a slight conflict over theoretical commitments, it is often necessary to draw inferences about the judgments that a co-author is likely to express regarding a certain lab's results, or about the theoretical tools

[9] For arguments about the role of language in human cognition, see: José Luis Bermúdez *Thinking Without Words* (New York: Oxford University Press, 2003); Carruthers, *The Architecture of The Mind* (Oxford: Oxford University Press, 2006) and "How We Know Our Own Minds," *Behavioral and Brain Sciences* 32 (2009): 121-38; Donald Davidson, "Thought and Talk," in *Inquiries into Truth and Interpretation* (Oxford: Clarendon Press, 2001); Daniel C. Dennett, *Brainstorms* (Cambridge, MA: Bradford Books, 1978), *Consciousness Explained* (New York: Penguin, 1992), and *Kinds of Minds*; and Ray Jackendoff, *Language, Consciousness, Culture: Essays on Mental Structure* (Cambridge, MA: MIT Press, 2007).

[10] Tadeusz Zawidzki, "The Function of Folk Psychology: Mind Reading or Mind Shaping?" *Philosophical Explorations* 11 (2008): 193-210.

that should be applied in addressing an issue. However, adopting this interpretive strategy is rarely a matter of attempting to predict or explain a collaborator's future behavior.[11] Instead, this strategy is typically adopted to recalibrate beliefs and desires by giving and asking for reasons, in hopes of retrieving a shared set of background assumptions that can underwrite further collaboration.

We can and do stretch these interpretive strategies to include human beings who live in radically different situations from our own. For example, when I have a conversation with a homeless person in my neighborhood, I ascribe a variety of mental states and processes over the course of the conversation in an attempt to negotiate a shared understanding of his world—for the world has been disclosed to him in a radically different way as a result of his living situation. However, as we move even further away from our own situation, it becomes harder to adopt this strategy. Even highly trained dogs who spend every day with a human being acquire only a narrowly constrained ability to communicate and never develop the capacity to flexibly respond to the *meaning* of what is said to them. (I return to this worry in Section 2.1.)

Ascribing mental states and processes to even the most familiar non-human animals, thus, requires a different kind of mentalizing strategy. In attempting to understand the behavior of a non-human, non-linguistic animal, we adopt an interpretive and explanatory strategy that treats their mental states and processes as functionally specified "black boxes." The existence of these black boxes is inferred from robust and predictable behavioral dispositions. While the supposition that an entity utilizes such states and processes is often useful for explaining its dispositions to behave in predictable ways, this thin strategy for ascribing mental states and processes also ushers in a host of theoretical difficulties. Everything from self-replicating macromolecules, to pine trees, thermostats, amoebas, plants, rats, bats, people, and chess-playing computers *can* be treated as an intentional system that has mental states and processes of this sort.[12] However, little philosophical ground can be gained by demonstrating that flies, honeybees, and spiders are minded in the same way as a pine tree that can be tricked into "believing that it is summer" by lighting a fire beneath its boughs. So we must look deeper for find a theoretically and empirically plausible foundation for distinguishing between genuine cognition and witless mechanical behavior.

2. MAXIMAL CARTESIAN MINDS

In the opening pages of the *Principles of Psychology*, William James argues that although the boundary between genuine cognition and witless mechanical behavior seems vague on first blush, there is a plausible intuitive test for determining whether a non-human entity has genuinely cognitive capacities. James extracts the details of this test by contrasting the behavior of soap bubbles blown through a straw into a pail of water and the behavior of a frog who is placed at the bottom of the same pail. Although the bubbles rise to the surface in a way that can be "poetically interpreted as a longing to recombine with the mother-atmosphere above," they continue to display this behavior even when a glass dome stands in the way of their "goal."[13] However, although the frog will also swim upward out of a desire to reach "the mother-atmosphere above," if a jar full of water be inverted over him, he will not, like the bubbles, perpetually press his nose

[11] Cf. Adam Morton, "Folk Psychology is not a Predictive Devise," *Mind* 105 (1996): 119-37; and Tadeusz Zawidzki, "The Function of Folk Psychology," *Philosophical Explorations* 11 (2008): 193-210.
[12] Dennett, *Kinds of Minds,* p. 34.
[13] William James, *The Principles of Psychology* (Cambridge, MA: Harvard University Press, 1890), p. 7.

against its unyielding roof, but will restlessly explore the neighborhood until by re-descending again he has discovered a path around its brim to the goal of his desires.[14]

James contends that for the frog the goal remains fixed and the means of achieving this goal are varied in light of the salient environmental contingencies of a particular situation. This fact, he claims, justifies the claim that the frog mentally represents its goal in some way. Building from this case, James argues that the "pursuance of future ends and the choice of means for their attainment, are thus the mark and criterion of the presence of mentality."[15] However, in spite of James's protestations to the contrary, it is not easy to know whether frogs, flies, honeybees, and spiders possess the representational capacities that would allow them to satisfy these criteria, or whether they merely behave as-if they do. If mental states and processes must pick out "real, intervening internal states or events, in causal interaction, subsumed under covering laws of causal stripe," then something could hop like a frog and croak like a frog, without really thinking like a frog.[16] In short, adopting methodological principles that are too permissive risks leading us to posit mental states and processes where nobody is home. So, how can we draw a sharp line between minded and non-minded entities?

2.1. Cartesian Arguments

René Descartes famously argues for a fundamental and unbridgeable gap between genuine mentality and the merely mechanical activity of non-human entities. He claims that the capacity to use language provides the only evidence that an organism's *goal-directed* behavior is not fully determined by its design; and, he argues that the ability to use a public language cannot be explained merely by appeal to the operation of complex physical mechanisms. In this section, I briefly examine each of these claims.

Descartes' argument relies on a version of the Principle of Sufficient Reason—the claim that for everything that is the case, it must be possible to provide a causal explanation or reason *why* it is the case.[17] Descartes' version of this principle suggests that an explanation of why a sandbar occurs at a particular point in a river is provided in terms of the structural properties of the grains of sand and the mechanistic principles governing their composition and aggregation.[18] Because Descartes also views biological bodies as complex machines, he advances a similar strategy for describing the behavior of biological organisms.[19] Briefly, to give an adequate explanation of a biological organism's behavior, it is sufficient to appeal to the organization of complex biological mechanisms, which are composed of bones, muscles, nerves, arteries, veins, and "animal spirits," and to explain how they are designed to carry out particular tasks in particular environments. With this model of explanation in hand, Descartes claims that the "high degree of perfection displayed in some of their actions makes us suspect that animals do not have free will."[20]

---

[14] James, *The Principles of Psychology*, p. 7
[15] James, *The Principles of Psychology*, p. 8
[16] Jerry Fodor, as quoted by Dennett, *The Intentional Stance* (Cambridge, MA: MIT, 1987), p. 53.
[17] As Descartes, suggests in the *Second Set of Replies* (p. 135) "there is nothing in the effect which was not previously present in the cause, either in a similar or in a higher form is a primary notion which is as clear as any that we have;" he says that this is just the claim that "Nothing comes from nothing." All references to Descartes' work refer to *The Philosophical Writings of Descartes,* 3 vols, translated by John Cottingham, Robert Stoothoff, and Dugald Murdoch (Cambridge: Cambridge University Press, 1985).
[18] Janet Broughton, *Descartes Method of Doubt* (Princeton: Princeton University Press, 2003), p. 155.
[19] Descartes, *Discourse on the Method*, vol. 1, p. 56.
[20] Descartes, *Early Writings*, vol. 1, p. 219.

According to this Cartesian view, we can exhaustively explain why it is difficult to swat a fly in mid air by treating the fly as a complex machine that relies on a fixed set of action patterns to rapidly change directions when it is faced with a looming object. Ascribing free will or attempting to work out what the fly believes about the looming object would seem to provide no additional explanatory advantage. Similarly, we can explain why cockroaches are efficient at scuttling away when the lights come on by appeal to the fixed action patterns that allow them to monitor the structure of their immediate environment and rapidly change directions without re-presenting anything in a cognitively significant way. These invertebrates are "designed" to navigate particular environmental contingencies. Where such design-based explanations prove successful, they seem to preempt appeals to intentional explanations. So, even if they often seem to outwit us in their native environments, we should not take this to be a matter of thought.

Descartes argues that while such mechanistic explanations always suffice for claims about the behavior and structure of complex physical bodies, they leave something to be desired in the case of human behavior. Here, we must adduce an additional reason or cause to explain the capacities that derive from will and judgment. Descartes contends that, as rational agents, we possess irreducible capacities for deliberating and withholding judgment, and although we do not have control over every behavior we exhibit, we experience ourselves as possessing the capacity to freely will and consciously control at least some of our actions. These facts call for an explanation; and since mechanistic principles could only explain human behavior by explaining away our capacity to freely will and consciously control our own actions, Descartes contends that a sufficient explanation of these capacities must advert to the existence of a thinking substance that obeys rational, as opposed to physical laws. In light of this argument, Descartes raises well-known epistemic worries about the existence of other human minds by suggesting that the bodies crossing a square *could be* mere automata.[21] Since we cannot tell from the outside whether these bodies possess minds, we must find some behavior that can only be explained by adverting to the presence of a thinking substance.

To resolve this worry, Descartes proposes a proto-Turing test, suggesting that the capacity to flexibly respond, in language, to the meaning of any unexpected question provides incontrovertible evidence that a mind is associated with a body.[22] Of course, Descartes acknowledges that an incredibly complex biological machine could be constructed so as to exhibit a finite and situationally constrained capacity to use particular words and expressions in response to a specified range of stimuli. However, he claims, it is inconceivable "that such a machine should produce different arrangements of words so as to give an appropriately meaningful answer to whatever is said in its presence, as even the dullest of men can do."[23] After all, over the course of an ordinary conversation, we encounter everything from jokes and sarcasm, to unexpected communicative breakdowns, absurd suggestions that require further elaboration, and even semantically ambiguous sentences that make sense only in light of shared background assumptions. Our flexible and open-ended capacity for thought allows us to seamlessly deploy a variety of cognitive resources that allow us to reason through such situations; and the fact that we can reestablish a shared ground for communicating in light of such breakdowns does

---

[21] Descartes, *Meditations,* Meditation II, vol. 2, p. 32.

[22] Alan Turing suggests an imitation game where a judge asks any question that she wishes to a sophisticated computer and an ordinary person; both the computer and the human attempt to convince the judge that they are the person. Turing argues that any machine that could regularly convince a discerning judge that they were a person by responding to the sense of her questions would display any plausible mark of genuine intelligence. Turing, "Computing Machinery and Intelligence," *Mind* 59 (1950): 433–60.

[23] Descartes, *Discourse on the Method*, Part 5, vol. 1, pp. 57ff.

seem to provide us with incontrovertible evidence that other people have context-sensitive capacities for thought just like our own. Descartes contends that the capacity to produce context-sensitive representations in a public language suggests an unbridgeable gap between genuine cognition and the fixed action-patterns displayed by non-human animals; and on this basis concludes that language provides the only clear way of distinguishing intentional action from witless and mechanical behavior that is exhaustively explained in terms of the design of some machine.[24]

Of course, this does not establish that these capacities for genuine mentality cannot be constructed from simpler, less sophisticated parts. However, Descartes suggests that any attempt to articulate similarities between the capacities of human beings and those of non-human animals will find that "the actions of the brutes resemble only those which occur in us without any assistance from the mind." He continues:

> when the swallows come in spring, they operate like clocks. The actions of honeybees are of the same nature; so also is the discipline of cranes in flight, and of apes in fighting, if it is true that they keep discipline. Their instinct to bury their dead is no stranger than that of dogs and cats which scratch the earth for the purpose of burying their excrement; they hardly ever actually bury it, which shows that they act only by instinct and without thinking. The most that one can say is that though the animals do not perform any action which shows us that they think, still, since the organs of their bodies are not very different from ours, it may be conjectured that there is attached to these organs some thought such as we experience in ourselves, but of a very much less perfect kind. To this I have nothing to reply except that if they thought as we do, they would have an immortal soul like us. This is unlikely, because there is no reason to believe it of some animals without believing it of all, and many of them such as oysters and sponges are too imperfect for this to be credible.[25]

In other words, since simple invertebrates like sponges and oysters possess mechanistic capacities to successfully navigate their environments, we must either suppose 1) that they have minds of a sort that cannot be explained in purely mechanistic terms, 2) that some high degree of mechanical complexity is sufficient for genuine mentality, or 3) that all non-human behavior can be exhaustively explained in mechanistic terms. Descartes rules out (1) by claiming that it is implausible to suppose that sponges and oysters have the capacities for will and judgment that are indicative of minds—as organisms, they are too simple to exercise conscious control over their behavior. He rules out (2), if only implicitly, by suggesting that any moderate criterion for ascribing mentality to non-human entities will be a matter of arbitrary line drawing. He sees no principled reason why we must draw the boundary between genuinely cognitive and merely mechanical systems in one place rather than another; however, his commitment to the Principle of Sufficient Reason entails that there must be a principled reason why some behavior can be explained mechanistically while other behavior cannot. The only distinction that

---

[24] In the *Discourse on the Method* (Part 5, vol. 1, p. 58), Descartes claims that it is "a very remarkable fact that although many animals show more skill than we do in some of their actions, yet the same animals show none at all in many others; so what they do better doesn't prove that they have intelligence, for if it did then they would have more intelligence than any of us and would excel us in everything. It proves rather that they have no intelligence at all, and that it is nature which acts in them according to the disposition of their organs. In the same way a clock, consisting only of wheels and springs, can count the hours and measure time more accurately than we can with all our wisdom."

[25] Descartes, *Fourth Set of Replies*, vol. 2, p. 230.

Descartes sees as tenable is the capacity for fluidly and flexibly responding, in language, to the meaning of what is said. Put simply, we must either claim that the behavior of oysters, sponges, and human beings can all be exhaustively explained in mechanistic terms, or we must adopt the restrictive claim that treats our linguistic capacities as providing the only incontrovertible evidence that an entity is minded.

2.2 Where Does the Cartesian Argument Go Wrong?

Descartes' avowedly materialist contemporaries saw that his arguments rested on problematic metaphysical presuppositions. Thomas Hobbes argued that *deliberation* and *will* were nothing more than *capacities* to manipulate internal representations; and because the capacity to represent an "alternate Succession of Appetites, Aversions, Hopes and Fears, is no lesse in other living Creatures then [*sic*] in Man," he thought it was obvious that non-human animals had these capacities as well.[26] Similarly, Baruch Spinoza argued that people fall prey to Cartesian assumptions about the mind because they have not "learned from experience what a body can and cannot do, without being determined by mind."[27] He argued that although commonsense psychology searches for purpose in the world, "final causes are but figments of the human imagination."[28] Causal and natural forces govern the behavior of all bodies; so, human minds are governed by the same laws that govern the mind of the snail.[29]

The naturalistic worldview of the cognitive sciences takes up this anti-Cartesian perspective, suggesting that the complexity and sophistication of an entity's behavior must be explicable in terms of interaction and aggregation of simpler, mechanistic systems. These presumptions derive from two "inversions of reason" that emerged in the 19th and 20th Centuries. Charles Darwin's opponents aptly captured the first inversion of reason, noting that "in order to make a perfect and beautiful machine it is not requisite to know how to make it."[30] Indeed, the mechanisms of natural selection explain how evolution can witlessly produce complex organisms out of uncomprehending parts. But mental states and processes should not be seen as a "lone exception" to the paradigm of mechanistic explanation.[31] Behavioral and cognitive capacities must also be explained by demonstrating that entities that possessed this capacity were more likely to survive and reproduce in their natural environments.

This brings us to the second inversion of reason, Alan Turing's argument that a mechanistic explanation of information processing can be given that does not advert to mechanisms of top-down or intentional control. Just as evolution by natural selection replaces the trickle-down theory of design by divine will, the theory of computation displaces the assumption that a computational device must understand the meaning of the symbols it manipulates to compute a function. In short, Turing provided a theoretical apparatus that can explain how computations can be carried out by a system that consists exclusively of uncomprehending and mechanistic components.

---

[26] Thomas Hobbes, *Leviathan* (Indianapolis: Hackett Publishing, 1994), p. 33; Similarly, David Hume, *A Treatise of Human Nature* (Oxford: Oxford University Press, 1978), p. 176: "no truth appears to be more evident, than that beast are endow"d with thought and reason as well as men."

[27] Spinoza, *Ethics*, trans. Samuel Shirley (Indianapolis, Indiana: Hackett, 1991), p. 105.

[28] Spinoza, *Ethics*, p. 59.

[29] Yitzhak Melamed, "Spinoza's Anti-Humanism," in *The Rationalists,* edited by Carlos Fraenkel, Dario Perinetti & Justin E.H. Smith (Amsterdam: Kluwer, in press), p. 17.

[30] Robert MacKenzie, as quoted in Dennett, "Darwin's 'Strange Inversion of Reasoning'" *Proceedings of the National Academy of Science* 106 (2009), p. 10062.

[31] "Darwin's 'Strange Inversion of Reasoning,'" p. 10062.

Together, these inversions of reason suggest plausible resources for resisting Descartes' arguments. Rather than assuming discontinuity between the physiological capacities exhibited by human and non-human animals, we should begin from the methodological assumption that every behavioral or cognitive *capacity* can be explained mechanistically. Complex physical structures emerge from simpler structures though a process of modification by natural selection; and as Turing argues, even the *capacity* for flexible linguistic engagement can be given a mechanistic explanation. Thus, psychological explanation, like all mechanistic explanation must begin by decomposing complex cognitive systems into simpler parts, and repeating this process until a level of explanation is reached that requires nothing more than the equivalent of on-off switches. Such explanations look "backward" to an organism's evolutionary history, and "downward" to the underlying computational architecture that implements its capacities to behave in various ways.

Thus, genuine mentality is always *implemented* by some mechanically complex computational system, and simpler minds are merely simpler mechanistic systems. So, extending mentality to invertebrates is not as dire a response as Descartes supposed, at least so long as a plausible boundary can be drawn between genuine cognition and witless mechanical behavior. Turing's suggestion is that the capacity to utilize representations in the production of goal-directed behavior serves to mark this boundary. However, to undercut the force of Descartes' arguments, it is necessary to offer a fuller theory of minimal mentality and to demonstrate that representational capacities provide a plausible strategy for distinguishing genuine cognition and mere mechanistic activity. I turn to these tasks in section 3.

3. A THEORY OF MINIMAL MINDS

As I noted above, the interpretive strategy that we adopt in describing intentional behavior can be readily deployed even where a simple entity displays nothing more than a fixed-action pattern (e.g., in describing the behavior of a virus). With this in mind, both cognitive ethologists and philosophers of cognitive science have tended to adopt a conservative perspective on non-human minds. Specifically, they have tended to adopt "Morgan's Canon," a special application of Ockham's razor according to which the behavior of a non-human animal is not "to be interpreted in terms of higher psychological processes, if it can be fairly interpreted in terms of processes which stand lower in the scale of psychological evolution and development."[32] I too agree that we should attempt to avoid the empty anthropomorphism that arises in a careless deployment of the intentional stance; treating viruses and thermometers as believers is at best crude behaviorism, and at worst flatly misleading. Thus, we need a more careful philosophical analysis of the problem of minimal minds.

Starting near the bottom of the phylogenetic tree, we find that all eukaryotic cells (i.e., cells that contain a distinct membrane-bound nucleus) possess the capacity to flexibly arrange and rearrange their internal structure in response to the current state of their environment.[33] Tecumseh Fitch has recently argued that eukaryotes, including single-celled bacteria, have the capacity to flexibly respond to novel circumstances in ways that are not explicitly encoded in their DNA, using trial-and-error methods to discover new responses to novel difficulties, and recording these discoveries for future use. A simple example of this biological capacity is the healing response: "a damaged organism can

[32] C. Lloyd Morgan, *An Introduction to Comparative Psychology,* 2nd edn.. (London: W. Scott, 1903), p. 59.
[33] W. Tecumseh Fitch, "Nanointentionality," *Biology and Philosophy* 23 (2008), p.158.

often stem the loss of precious bodily fluids, stitch itself up, and (with some scar perhaps) continue living."[34] However, the possession of this capacity is insufficient to license a claim of genuine mentality. It only shows that simple organisms are capable of engaging in some degree of semi-autonomous functional reorganization. Because there is no "dedicated information-processing machinery even vaguely equivalent to a vertebrate nervous system," it doesn't make sense to say of bacteria that they have mental circuits of the sort that would allow them to represent the world in any way whatsoever.[35] In short, bacterial cells possess the capacity to record the strategies that they have used in responding to their environment; however, these recordings are fixed by particular interactions with particular bodies. Every interaction requires a new attempt at functional reorganization, because bacteria are not internally complex enough to re-deploy previously honed skills in coping with novel situations.

From a Darwinian perspective, such capacities are likely to be necessary preconditions for the emergence of the flexible "strategies" that are required to achieve even the simplest of goals (e.g., feeding and mating). However, simple bacteria do nothing more than take in sensory inputs, and witlessly and mechanically translate them into the sorts of internal changes that allow them to survive (another day, another hour, or another minute). Given the Darwinian history that structures their pattern of responses, we can truly say that these bacteria are not fully designed from birth. There are elements of their design that can be adjusted by the events that occur in over the course of interactions with their environment.[36] However, these adjustments do not *mean* anything to the bacteria themselves. Genuine cognition requires enough mechanical complexity to represent the world in a way that can facilitate being in cognitive states and carrying out cognitive processes.

Crude behaviorism and empty instrumentalism do little more than operationalize "belief" and "desire", radically redefining these terms to the point that they share little in common with the familiar states and processes of other human minds. Empty anthropomorphism, by contrast, threatens to deploy our familiar mental terminology in cases where the use of such terms is completely unwarranted. To avoid these worries, discussions of other kinds of minds often begin by providing relatively conservative criteria for discriminating genuine cognition from witless mechanistic behavior. For example, even the most strongly naturalist positions on the existence of invertebrate minds, such as that of Peter Carruthers, depend on the claim that a minimally minded entity must have "distinct belief states and desire states that are discrete, structured, and causally efficacious in virtue of their structural properties"; these beliefs and desires must be "construed realistically," and because these states must "be both discrete, and structured in a way that reflects their semantic contents," it has been claimed that genuine mental states and processes must be implemented in a system of symbols that obey systematic transformational rules (e.g., in a language of thought).[37]

I believe that there is a great deal of value to Carruthers' proposal. However, there is something strange about the claim that the states and processes that must be present for an entity to count as minded must be beliefs and desires. Carruthers attempts to defuse this worry by adverting to *spatial* beliefs and desires as opposed to *causal* beliefs and desires. He holds that many kinds of invertebrate beliefs are likely to represent by encoding the spatial relationships between salient features of their environment rather

---

[34] Fitch, "Nanointentionality," p. 162.
[35] Fitch, "Nanointentionality," p. 163.
[36] Dennett, *Kinds of minds*, p. 83.
[37] Carruthers, *The Architecture of Mind,* po. 67-68; and, this volume.

than representing the causal relations that obtain between various objects and entities in their environment. However, this seems to miss the crucial worry about positing invertebrate beliefs and desires: Why should the high-level computations that underlie the capacities for believing and desiring play such a central role in philosophical and scientific theorizing about the capacity for genuinely intelligent action? Our cognitive capacities depend to a great extent on our linguistic capacities; and, I agree with Carruthers that any entity that can be in cognitive states or carry out cognitive processes must be able to "act with flexibility and forethought, choosing between different courses of action and anticipating future consequences. These abilities seem to demand representations that stand in for external objects."[38] However, it is not obvious what it takes to be able to represent in this way.

As my discussion of bacteria suggests, genuine cognition requires internal states and processes that *represent* rather than *merely record* the way that the world is. Put differently, genuine cognition requires internal states or processes that have the *function* of *conveying information* about the way that the world is, in a way that can provide the necessary resources for guiding the behavior of that organism. As John Haugeland puts the point, an internal state need only be the result of particular processes in order to count as a recording, but "representing is a functional status or role of a certain sort, and to be a representation is to have that status or role."[39] Here we must ask, what sort of functional role has to be filled if something is to count as a representation? I propose to follow Haugeland in accepting the following rough-and-ready desiderata on the sort of functional organization required for representation.[40] These are only intended as rough-and-ready criteria, but they should seem highly intuitive and it should be clear that they play a dominant role in structuring philosophical and empirical approaches to questions about representations.[41] I suggest that the possession of a minimal mind requires that an entity:

1. Possess internal states and processes that have the function of adjusting the entity's behavior in ways that allow it to cope with features of its environment in ways that are not fully determined by the design of the system;

2. These states and processes must be capable of standing-in for various features of the environment that are significant for the system, and they must be capable of doing so even in the absence of immediate environmental stimuli;

3. These states and processes must be part of a larger representational scheme that allows for a variety of possible contents to be represented (in a systematic way) by a corresponding variety of possible representations; and,

4. There must be norms governing the proper production, maintainance, and use of these representations under various environmental conditions.[42]

[38] Jesse J. Prinz, *Furnishing the Mind* (Cambridge: MIT Press, 2004), p. 4.
[39] John Haugeland, "Representational Genera," in *Having Thought* (Cambridge, MA: Harvard University Press), p. 180.
[40] Haugeland, "Representational Genera," p. 172.
[41] Andy Clark, *Being there: Putting Brain, Body and World Together Again* (Cambridge, MA: MIT Press, 1997), pp. 143-77.
[42] This does not require misrepresentation in the fullest sense of the term. Familiarly, you can fill a frog with buckshot by shooting BBs past it in the lab. This is a representational failure, but it is not a genuine

I suggest that any entity that possesses such capacities will have sufficient representational complexity to be in genuinely cognitive states and to carry out genuinely cognitive processes. However, it is important to acknowledge that there are a number of ways in which the second clause of (2) can be construed.

On the most robust construal, such states and processes would have to be able to represent things that do not exist. For example, my representation of a DEMOCRATIC SOCIETY regulates much of my behavior, in ways that allow me to cope with my environment. It stands-in for a nonexistent structure of political organization, plays an important role in my overall understanding of the world, and is embedded in a system of social norms that regulate my claims about whether some society can reasonably be said to be a genuine democracy.

On a much weaker construal, however, these criteria might license the claim there are some sorts of entities that possess these sorts of representations without having the kinds of discrete beliefs and desires that Carruthers requires. Indeed, it may be that for the most minimal minds, all of the relevant representational capacities will are implemented in perception-action circuits that preclude the complete decoupling of the indicative and imperative components of their representations. That is, minimal minds might *only have* pushmi-pullyu representations.[43] Such states and processes both represent the world as being a particular way, and function to direct immediate action. However, they are not mere couplings of belief-like states and desire-states that are filtered through a general-purpose practical reasoning mechanism. These are a more primitive type of representation, which immediately yield changes in behavior as a direct function of changes in the environment.

According to Morgan's Canon, "we should not trust psychological explanations of behavior unless we are convinced that those explanations are indispensable—that is to say, unless we are convinced that the behavior in question cannot be explained in nonpsychological terms."[44] Indeed, we *should be* dubious about theories that accord minimal minds the capacity to entertain complex thoughts that are "linguistically vehicled," or capacities for reasoning that require the manipulation of abstract propositions.[45] Much more than this, I hold that without the capacity to use linguistic representations, it is overwhelmingly unlikely that minimal minds will possess capacities to form episodic memories, deliberate about counterfactual possibilities, or form anything resembling a fine-grained belief or desire. This is why José Luis Bermúdez is right to argue that non-human animals are likely to possess only capacities for thought and reasoning that are tightly coupled to the perceptible features of their environment.[46] Yet, explanations in terms of the cognitive capacities of fairly simple invertebrates might still be indispensable from the perspective of cognitive science, even if such entities are not complex enough to have states or processes that are like beliefs and desires in any interesting sense. Examining this point, however, requires a turn to the empirical data that support the claim that some invertebrates can be in cognitive states and carry out cognitive processes.

3.1 Darwinian Minimal Minds

misrepresentation. The frog's visual system is just too impoverished to reliably discriminate between food and near-non-food.

[43] Ruth Millikan, *Language, Thought, and Other Biological Categories* (Cambridge, MA: MIT Press, 1988).

[44] Bermúdez, *Thinking without Words*, p. 7.

[45] Bermúdez, *Thinking without Words*, p. ix.

[46] Bermúdez, *Thinking without Words*.

In his final book, *The Formation of Vegetable Mould, through the Actions of Worms with Observations on Their Habits*, Darwin turns his ethological eye to the behavior of earthworms. While the majority of this work is dedicated to examining the impact of worms on their physical environment, Darwin also examines the possibility that worms are intelligent. His observations begin with the fact that worms tend to plug the openings of their burrows with leaves, sticks, and petioles of varying sizes. Darwin found that these objects were pulled into the openings of burrows in a way that was too uniform to be attributed to mere chance, but at the same time they seemed to be carried out in a way that is not unvarying or inevitable enough to be called a mere instinct.[47] By carefully examining the diverse strategies for manipulating these objects, Darwin finds that there are significant differences in the techniques used for pulling various leaves (e.g., Rhododendron leaves, pine needles, and artificial leaves made from triangles of folded paper) into the opening of the burrow; and with unfamiliar artificial leaves, he found that worms often attempted several distinct strategies for pulling the leaf, using a methodology of trial and error before settling on a preferred strategy for handling these new objects.

With these observations in hand, Darwin argues that worms should be seen as possessing a capacity to learn from experience to discriminate between various shapes of leaves, twigs, and petioles. So, adopting an argument from parity, he argues that "if worms have the power of acquiring some notion, however rude, of the shape of an object and of their burrows, as seems to be the case, they deserve to be called intelligent; for they can act in a manner as would a man under similar conditions."[48] Darwin thus attempts to show that earthworms adjust their behavior in ways that allow them to cope with novel features of their environment; and he argues that these states are not fixed by, though they are clearly structured by, the evolutionary history of these organisms. Darwin is keenly aware of the initial oddity of this claim and he carefully articulates a series of reasons for why his observations of worm intelligence cannot be generalized to explain the behavior of other invertebrates without further empirical investigation. However, through the careful examination of worm behavior, and by ruling out all of the available alternative explanations, Darwin argues that his views "will strike everyone as improbable; but it may be doubted whether we know enough about the nervous system of the lower animals to justify our natural distrust of such a conclusion."

Darwin's groundbreaking examination of the mental life of worms is important not for its substantive conclusions—which fail to demonstrate a capacity for mental representation, even if worms might have such a capacity—but for its methodological conclusion. Instead of arguing that there is no reason to deny the possibility of insect minds, Darwin set out to examine the evidence in favor of this claim. He demonstrates that claims about invertebrate intelligence are

> not ontologically problematic because it is not rational to presume, prior to inquiry, that the existence of conscious action is unlikely, even among invertebrates; not epistemologically problematic because once the question of conscious action is allowed to be posed, the scientific imagination finds fascinating ways to address it; and finally, not semantically problematic because

[47] Charles Darwin, *The Formation of Vegetable Mould Through the Action of Worms* (Chicago: University of Chicago Press, 1985), p. 93.
[48] Darwin, *The Formation of Vegetable Mould Through the Action of Worms*, p. 97.

writing off “conscious action” as anthropomorphism commits the deeper fallacy of anthropocentricism.[49]

Darwin demonstrates that careful observation and experimentation can be tuned on the behavior of invertebrates; and, he shows that it is possible to investigate the question of invertebrate minds scientifically. While he recognizes that this claim will seem questionable from the standpoint of commonsense psychology, he hopes to displace our unfounded prejudices about the nature of cognition with sound scientific inquiry so that the commonsense understanding of mentality can be revised to construct a plausible scientific understanding of what it takes to have a minimal mind.

3.2. The Mind of the Honeybee

To further develop these Darwinian assumptions about minimal mentality, it will be useful to consider the research on the intelligence of honeybees. It has long been known that honeybees possess the capacity to communicate the distance and direction of food sources by way of a set of ritualized dancing movements.[50] When a high-quality food source is found near the hive, honeybee scouts use a “round” dance that informs the other bees that they should search for the odor associated with the scout upon leaving the hive. However, for more distant food sources a more elaborate “waggle dance” is used to convey information about the direction, distance, and quality of the food source. Direction is indicated by the angle of the movements of the abdomen across the center of a figure eight (which represents the angle from the position of the sun); distance is indicated by the duration of the dance (and sounds); and quality is indicated by the vivacity of the dance.

Honeybees have also been shown to possess capacities for navigating and encoding new memories acquired during explorations. Bees that travel long distances from their nest site seem to construct cognitive maps that rely on simple geometric representations of space, and vector information that can be recovered using “dead reckoning” strategies.[51] They also appear to integrate landmark information that is “experienced en route” to re-calibrate distances, reduce the potential navigational errors, and to encode procedural information about what to do next.[52] Moreover, Randolf Menzel argues that honeybees discriminate between different colors, shapes, patterns, odors, and textures while foraging to construct accurate mappings of landmarks.[53] Finally, Shaowu Zhang, James Bartsch, and Mandyam Srinivasan have shown that bees can navigate complex and unfamiliar mazes by learning to see particular colored disks as meaning either “turn left” or “turn right.”[54]

[49] Ellen Crist, “The Inner Life of Earthworms,” in *The Cognitive Animal*, eds. Marc Bekoff, Collin Allen, and Gordon Burghardt (Cambridge: MIT Press, 2002), pp. 7-8.

[50] Karl Von Frisch, *The Dance Language and Orientation of Bees* (Cambridge: Harvard University Press, 1967); James Gould and Carol Gould, *The Honey Bee* (New York: Scientific American Library, 1988).

[51] “Dead reckoning” is the process of integrating direction, velocity, and time traveled to produce a representation of current location relative to initial location. See Randy Galistel, *The Organization of Learning* (Cambridge: MIT Press, 1990); and James Gould, “The Locale Map of Honey Bees,” *Science* 232 (1986): 861-63. A skeptical response to these data is provided by Margaret Wray, Barrett Klein, Heather Mattila, and Thomas Seeley, “Honeybees do not Reject Dances for “‘Implausible’ Locations—Reconsidering the Evidence for Cognitive Maps in Insects,” *Animal behavior* 76 (2008): 261-69.

[52] Randolf Menzel and Martin Giurfa, “Dimensions of Cognition in an Insect, the Honeybee,” *Behavioral and Cognitive Neuroscience Reviews* 5 (2006), p. 26.

[53] Menzel, “Searching for the Memory Trace in a Mini-Brain,” *Learning and Memory* 8 (2001): 53-62.

[54] Shaowu Zhang, James Bartsch, and Mandyam Srinivasan, “Maze Learning by Honeybees,” *Neurobiology of Learning and Memory* 66 (1996): 267-82.

Honeybees also develop long-term expectations that cannot be derived from the strength of an association with a predicting signal, using these expectations to guide behavior even after a relatively long delay following an initial reinforcement.[55] This capacity plays a critical role in determining the quality of a food source; but it is also deployed to evaluate the threat of predation by crab spiders that change color to match their environment. Foraging bees that face a simulated risk of predation by robotic crab spiders that are camouflaged (yellow spiders on yellow flowers) or easily detectable (white spiders on yellow flowers) initially land randomly on flowers that are inhabited by predators randomly. However, after a simulated attack, the likelihood that they will land on an inhabited flower falls off rapidly, at the same rate for camouflaged and conspicuous spiders. Using tracking software, however, Thomas Ings and Lars Chittka have shown that bees that are exposed to camouflaged threats trade off speed for accuracy in threat detection, increasing the duration of inspection flights prior to landing on potentially occupied flowers. This levels out predation risk by increasing the amount of time that they spend foraging. However, bees do not alter their foraging speed when they expect to be able to detect predators easily.[56]

Finally, honeybees also extract within category similarities and between category differences for various kinds of stimuli. For example, bees that are presented with a variety of gratings that differ in both the distances between lines and in edge-orientation can extract edge-orientation as a category that indicates the presence of a reward. When 45-degree gratings indicate a high reward sucrose solution in a training condition, bees generalize to predict that 45-degree gratings will indicate the presence of a reward in an experimental condition that contains novel stimuli that differ in all structural properties other than edge-orientation.[57] Additionally, honeybees who are trained to fly down the corridor in a Y maze marked with the same (or a different) color than an entry room reflexively transfer this skill to a novel domain in which corridors and entry rooms are marked with smells instead of colors.[58] Indeed, numerous experiments confirm that honeybees can successfully learn match-to-sample and don't-match-to-sample rules—suggesting a capacity to encode differences and not just similarities.[59]

The communicative activity displayed in the waggle dance, as well as the individual representations of location, landmarks, predators, and the like require a set of internal states and processes that have the function of carrying information about salient features of their environment. Some of these internal states and processes are then expressed as communicative signals that are encoded in the precise structure of the "waggles" in the waggle dance. These states and processes carry information about the world that is not completely determined by the design of the system, but encoded in the interaction of the bee with its environments. The internal states and processes that are expressed in the "waggles" are stand-in for various features of the environment that are significant for the bees, and they stand-in in this way even for the bees who have not yet foraged in the relevant environment. The "waggles" also express internal states and processes that are part of a larger representational scheme that allows a variety of possible contents (e.g., different distances, directions, and food sources) to be systematically represented.

---

[55] Mariana Gil de Marco and Menzel, "Learning Reward Expectations in Honeybees," *Learning and Memory* 14 (2007): 491-96.

[56] Thomas Ings and Lars Chittka, "Speed-Accuracy Tradeoffs and False Alarms in Bee Responses to Cryptic Predators," *Current Biology* 18 (2008): 1520-24.

[57] Uwe Homberg and John Hildebrand, "Serotonin-Immunoreactive Neurons in the Median Protocerebrum and Suboesophageal Ganglion of the Sphinx Moth *Manduca sexta*," *Cell and Tissue Research* 258 (1989): 1-24.

[58] Menzel and Giurfa, "The Concepts of 'Sameness' and 'Difference' in an Insect," *Nature* 410 (2001): 930-33.

[59] Reviewed in Menzel and Giurfa, "Dimensions of Cognition in an Insect, the Honeybee," pp. 24-40.

Finally, the experiment on predation demonstrates that the representational system can misrepresent a robotic spider as a genuine threat. This much is sufficient to establish the existence of honeybee minds that include "a suite of information generating systems that construct representations of the relative directions and distances between a variety of substances and properties and the hive, as well as a number of goal-generating systems taking as inputs body states and a variety of kinds of contextual information, and generating a current goal as output."[60]

It is important to note, however, that these capacities do not involve arbitrary signals that can be *robustly decoupled* from the salient features of the environment. For example, because the waggle dance is an iconic representational system, the content of a particular waggle is necessarily tied to the particular qualities of the environment that are at issue in a particular dance. For each dimension of variation in the environment (e.g., distance or direction), there is a single transformation rule that maps the variation in that parameter onto a parallel variation in the signal. This being the case, the consumer of the signal needs only apply an inverse mapping to decode the signal.[61]

This iconic symbol system still allows for a variety of foraging behaviors that take honeybees incredible distances away from their hives.[62] Yet, many of the inferential capacities that are present in language-using animals like us are wholly lacking in honeybees. Specifically, the representational states and processes that we find in honeybees do not appear to satisfy the *generality constraint*. In explicating what the generality constraint comes to, Gareth Evans notes that there is an important sense in which thoughts like ours are complex and structured representations; this being the case, anyone who can have the thoughts SASHA IS HAPPY and RAMON IS SAD will, thereby, be able to think RAMON IS HAPPY and SASHA IS SAD.[63] Put more formally, "If a subject can be credited with the thought that A is F, then he must have the conceptual resources for entertaining the thought that A is G, for every property of being G of which he has a conception."[64] The problem, as Carruthers aptly notes, is that it appears as though bees can be in mental states with the following content: THERE IS SOME HIGH QUALITY NECTAR 115 METERS DUE EAST OF THE HIVE and THERE IS SOME POLLEN 45 METERS DUE WEST OF THE HIVE. However, it seems highly unlikely that possessing these capacities can allow the bee to think THERE IS SOME NECTAR 190 METERS DUE EAST OF THE POLLEN.[65] This being the case, we might suppose that the honeybee representations are not belief-like in a robust, conceptual sense. Carruthers dismisses this worry by noting that because ecologically valid situations in which such a complex representation as this is likely to be required are highly unlikely, the fact that bees never do produce such a representation does not imply that they cannot.[66] However, there are further empirical data suggesting that bees cannot draw the sorts of inferences that we would expect from an entity that had robust conceptual abilities like our own.

Bees do not possess the inferential capacities that allow for transitive inferences.[67] Although they can be taught that the reward at point A is greater than the reward at point B, and that reward at B is greater than the reward at C, they cannot infer that the reward

[60] Carruthers, *The Architecture of the Mind,* p. 74.
[61] Bermúdez, *Thinking Without Words*, p. 288.
[62] Fred Dyer and Seeley, "Distance Dialects and Foraging Range in Three Asian Honey Bee Species," *Behavioral Ecology and Sociobiology* 28 (1991): 227-34.
[63] Gareth Evans, *Varieties of Reference* (Oxford: Oxford University Press, 1982), p. 101.
[64] Gareth Evans, *Varieties of Reference*, p. 104.
[65] Caruthers, *The Architecture of Mind*, pp. 78ff.
[66] Carruthers, *The Architecture of Mind*, p. 79.
[67] Menzel and Giurfa, "Dimensions of Cognition in an Insect, the Honeybee."

at A is greater than the reward at C; "bees do not establish transitive inferences between stimuli but rather guide their choices by the joint action of a recency effect … and by an evaluation of the associative strength of the stimuli" (i.e., they rely on the most recent item encountered and the strongest conditioned associations).[68] This fact cuts against the claim that bees have representational states and processes that resemble the conceptual representations possessed by language users; and it seems plausible that because bees do not possess the capacities using words in a natural language, their mental states and processes will be both more restricted and less inferentially promiscuous than human beliefs and desires.

I contend that it is a mistake to treat the states and processes of honeybees as beliefs and desires, as Carruthers suggests they must be. Yet, Carruthers is right to note that an argument grounded in the generality constraint fails to undercut the claim that invertebrates have simple mental states and processes. Carruthers' concerns derive from a desire to resist empty instrumentalism. However, this argument neglects the fact that there are many ways of being in mental states and carrying out mental processes that do not depend on having beliefs and desires like our own. Honeybees do possess cognitive capacities that cannot be explained unless we advert to representational capacities.[69] However, the mistake is to assume that mental representations must have a proto-linguistic structure; in fact, many mental representations are embodied and skill-based representations that underwrite the fast-and-frugal coping behavior in rapidly changing environments.[70]

The quasi-Cartesian assumption that genuine mentality requires conceptual representations in a language of thought ignores the representational capacities that are most likely to be implemented in biological systems that must rapidly cope with changes in their environments. For example, sensory systems have been selected for the way in which they provide information to motoric systems to facilitate rapid coping behavior in the face of danger, food, or potential mates. Important relationships may well obtain between such representations and the properties, relations, and things in the world; however, because sensory systems have been selected to yield fast-and-frugal action, it is unlikely that they will depend on symbolic relations whose content is isomorphic with the content of the semantic representations that we find in language. To put this point another way, the most "primitive" representations in sensory systems are likely to be constructed in a contextually sensitive and narcissistic fashion (*sensu* Akins). More complex representations in language and thought are then likely to be the result of triangulating these low-level representations on the basis of competitive or quasi-competitive algorithms (see the concluding section of this chapter).

With this fact in mind, I suggest that there are likely to be many kinds of mental states and processes that are not completely decouplable from their immediate causes, and that may not satisfy Evans' generality constraint. However, such states and processes can be genuinely representational and can play an important role in the mental lives of human and non-human animals alike. To clarify, consider the population of neurons in the rat's parietal cortex that *represent* (in a rich mental sense) the direction of the rat's head. As Andy Clark notes, we gain a great deal of explanatory power by acknowledging that these neurons represent the position of the rat's head; unless we do so, we cannot understand how information flows through the rat's cognitive system as a whole.[71] Yet, it

---

[68] Menzel and Giurfa, "Dimensions of Cognition in an Insect, the Honeybee," p. 36.
[69] Carruthers, *The Architecture of Mind*, p. 73.
[70] Kathleen Akins, "Of Sensory Systems and the 'Aboutness' of Mental States," *The Journal of Philosophy* 93, (1996): 337-72; Clark, *Being There.*
[71] Clark, *Being There*, p. 145.

would be a mistake to treat these neurons as-if they implement beliefs and desires; doing so would be nothing short of revisionary semantics given that they are nothing more than information-bearing structures that represent the direction of the rat's head. However, provided that they are integrated into a larger cognitive system that allows for the representation of a variety of facts about the world (in a variety of ways), they ought to be treated as mental states that are empirically tractable from the perspective of the cognitive sciences.

The internal states and processes that we find in the neural architecture of a honeybee are likely to be carried out independently of anything like central cognition. Many of these states and processes will be simple sensory-motor loops that reflexively yield particular behaviors depending on the current state of the environment. Others may be more decouplable from immediate presentations, allowing for more generalized inferences on the basis of category membership and cross-modal integration. However, the minimal minds of these invertebrates are unlikely to have anything like robustly conceptual, amodal, and quasi-linguistic representations that can be completely decoupled from their environmental triggers. In fact, the fact that these representations cannot be decoupled from environmental triggers is the crucial thing that distinguishes relatively maximal from relatively minimal mental representations. In short, different kinds of mental states and processes are likely to lie along a continuum, with pushmi-pullyu representations lying at one pole, and the richly amodal representations of linguistically structured thought at the other. While it is unlikely that we share our capacity to represent amodally with any non-human animals, many of our capacities to represent the world by way of pushmi-pullyu representations are likely to be shared far down the phylogenetic tree.

4. COLLECTIVE MINIMAL MINDS

In this final section, I wish to redeem the promissory note that I offered at the end of the introduction. There, I suggested that an adequate understanding of minimal mindedness would also provide evidence for the existence of collective mental states and processes in eusocial insects such as bees. Thomas Seeley has argued that honeybee colonies should be seen as unified systems, and that these systems rely on the iconic representations of the waggle dance to propagate information in a way that allows the colony to respond to stimuli that are salient to the colony *as such*.[72] There is a growing consensus that colonies of eusocial insects should be treated as single units of selection for the purposes of biological research. Even Richard Dawkins, a rabid "smallist" about explanation, agrees.[73] However, more is required to establish that honeybee colonies have a cognitive life beyond that of the individual bees.

Seeley contends that observations of foraging bees suggest that monitoring the location and richness of food sources, and evaluating the relative quality of various food sites can only take place in the distributed representations of the colony.[74] Seeley argues that foragers act as a diffuse sensory extension of the colony. Each bee begins with a random search for foraging sites. These initial foragers map the surrounding

[72] Thomas Seeley, *The Wisdom of the Hive* (Cambridge: Harvard University Press, 1995).
[73] Richard Dawkins, *The Selfish Gene* (Oxford: Oxford University Press, 1989).
[74] Seeley, "Social Foraging by Honeybees: How Colonies Allocate Foragers Among Patches of Flowers," *Behavioral Ecology and Sociobiology* 19 (1986): 343-54; "The Tremble Dance of the Honey Bee: Message and Meanings," *Behavioral Ecology and Sociobiology* 31 (1992): 375-83; and, "Honey Bee Colonies are Group-Level Adaptive Units," *The American Naturalist* 150 (1997): 22-41.

environment, finding patches of food as far as 10 km away.[75] However, for foraging sites within 2 km, it seems that bees can engage in comparisons of richness in just the way that should not be possible given the results reported by Menzel and Guirfa.[76] This process takes place by way of an aggregation of information at the level of the colony that cannot be carried out by any of the individual bees.[77] When employed foragers return to the hive, they advertise the distance, direction, and quality of a foraging site with their waggle dance. However, while individual bees follow only one other bee's dance, the likelihood of being recruited to a foraging site is determined by the duration and vivacity of a forager's waggle dance.[78] Those foragers who have visited desirable worksite dance longer and more vivaciously than bees who have visited less desirable foraging sites. The information about quality is distributed across the employed foragers in a way that does not require a centralized decision making structure to allocate unemployed foragers to new foraging sites.

By modulating the quality and quantity of a pair of artificial food sources, Seeley demonstrated that honeybee colonies become more selective when food sources are abundant.[79] However, under conditions of scarcity, foragers are allocated to even low-profit nectar sources. The mechanism of this selectivity can be specified in terms of the modulation of the threshold at which waggle dances occur. When a forager returns to the hive, its first task is to find a receiver bee who will accept nectar for storage. When resources are scarce, returning foragers rapidly find receiver bees, and so even a short dance will find an audience. This being the case, bees are recruited to less profitable foraging sites. When food is abundant, by contrast, the search for a receiver bee takes a longer amount of time; so only a longer and more vivacious dance will find an audience. Only high profit food sources are then exploited. Although there is no central cognition dedicated to monitoring the abundance or scarcity of food, the colony is capable of evaluating the relative abundance of food sources at various foraging sites even though none of the individual foragers or receivers is capable of representing this.

Still, there are complications. Nectar collection and processing sometimes fall out of synch. When this happens, foragers who have found incredibly rich food supplies need to help boost nectar collection rates to collect as much nectar as possible, but they also have to increase the rate at which nectar is processed to allow the bees who are returning from a high quality foraging site to find receivers for their pollen. When a forager returns from an incredibly high quality foraging site and finds that it has an extensive search time for finding a receiver, it executes a "tremble dance" that carries the information that unemployed bees should immediately begin processing nectar. For bees who have been foraging, this dance also carries the information that they should refrain from recruiting additional foragers, thereby acting as a suppression signal to inhibit the waggling of other bees. The execution of a tremble dance thus updates the rate at which nectar is processed so that the quantity and quality of nectar can be recalibrated.

---

[75] Seeley, "Honey Bee Colonies are Group-Level Adaptive Units."
[76] Menzel and Giurfa, "Dimensions of Cognition in an Insect, the Honeybee," pp. 24-40.
[77] Seeley, "Division of Labor Between Scouts and Recruits in Honeybee Foraging," *Behavioral Ecology and Sociobiology* 12 (1983): 253-59; "Social Foraging by Honeybees;" "Honey Bee Colonies are Group-Level Adaptive Units;" and Corinna Thom, Seeley, and Jürgen Tautz, "Dynamics of Labor Devoted to Nectar Foraging in a Honey Bee Colony: Number of Foragers Versus Individual Foraging Activity." *Apidologie* 31 (2000): 737-38.
[78] Seeley, Scott Camazine and James Sneyd, "Collective Decision-Making in Honey Bees: How Colonies Choose Among Nectar Sources," *Behavioral Ecology and Sociobiology* 28 (1991): 277-90; and Seeley and William Towne, "Tactics of Dance Choice in Honey Bees: Do Foragers Compare Dances?" *Behavioral Ecology and Sociobiology* 30 (1992): 59-69.
[79] Seeley, "Honey Bee Colonies are Group-Level Adaptive Units," pp. 28ff.

Finally, Seeley and his colleagues turn to the process by which new nest sites are selected.[80] When a colony outgrows its hive, it splits in two. One colony swarms around a tree branch and sends out scouts (approximately five percent of the swarm) to find a new nest site. During the initial search, as many as a dozen potential nest sites are selected, and each is evaluated by a scout according to six desiderata: cavity volume; entrance size, height, direction, and proximity to the cavity floor; and presence of combs in the cavity.[81] As the scouts return, they waggle dance to indicate the presence and quality of these features of the potential nest site, and although each scout only dances for one site (rarely, if ever, dancing for another site after having made their initial selection), a collective decision emerges and there is a consensus on one site (I address this claim more fully below). Strikingly, the swarm reliably chooses the site that best satisfies the six desiderata listed above (rather than settling on the first adequate site, for example), and the swarm only moves where there is complete consensus on that site. Seeley and his colleagues demonstrate that this consensus emerges because after the initial scouts dance for their chosen site, those bees that have found a mediocre or passable nest site dance less vigorously than those bees that have found a high-quality site.[82] Heavier recruitment of additional scouts occurs for higher-quality nest sites, and eventually this leads to the cessation of dancing for lower-quality nest sites. Lower-quality sites lose support until only the highest quality site is being danced for, leading eventually to the reliable selection of the highest quality nest site without requiring any of the individual bees to have a broad knowledge of all of the alternative possible nest sites that are under consideration by the swarm.

Seeley's data provide good reason for thinking that the specialization of function in a honeybee colony can facilitate the propagation of representational states (e.g., states that represent the location of nectar, the quality of a foraging site, and the location of a nest site) between bees with very different functionally specified tasks. As these representations are propagated between the members of a colony, a complex comparative evaluation emerges that cannot be made by the individual bees on their own. Thus, he suggests that the comparative judgments are carried out by distributed computational architecture that is realized by the colony as a whole, rather than by any of the computational nodes (i.e., the individual bees). By positing cognitive states and processes that are properly attributable to the honeybee colonies as such, Seeley can explain such diverse phenomena as the decision to build a nest in one site rather than another and the decision to allocate more resources to collecting or storing nectar. Such predictions are only possible on the assumption that there are cognitive states and process that properly attributable to the collectivity. The choice of a nest site is a striking demonstration of this fact. The colony chooses the best nest site possible even though

---

[80] Madeline Beekman, Robert Fathke, and Seeley, "How Does an Informed Minority of Scouts Guide a Honeybee Swarm as it Flies to its New Home?" *Animal Behaviour* 71 (2006), 161-71; Kevin Passino, and Seeley, "Modeling and Analysis of Nest-Site Selection by Honey Bee Swarms: The Speed and Accuracy Trade-Off," *Behavioral Ecology and Sociobiology* 59 (2006): 427-42; Seeley and Susannah Buhrman, "Nest-Site Selection in Honey Bees: How Well do Swarms Implement the 'Best-of-N' Decision Rule?" *Behavioral Ecology and Sociobiology* 49 (2001): 416-27; and, Seeley and P. Kirk Visscher, "Choosing a Home: How the Scouts in a Honey Bee Swarm Perceive the Completion of Their Group Decision Making," *Behavioral Ecology and Sociobiology* 54 (2003): 511-20.

[81] Seeley and Buhrman, "Nest-Site Selection in Honey Bees."

[82] Beekman et al, "How Does an Informed Minority of Scouts Guide a Honeybee Swarm as it Flies to its New Home?"; Passino and Seeley, "Modeling and Analysis of Nest-Site Selection by Honey Bee Swarms"; Seeley, "Consensus Building During Nest-Site Selection in Honey Bee Swarms: The Expiration of Dissent," *Behavioral Ecology and Sociobiology* 53 (2003): 417-24; Seeley and Buhrman, "Nest-Site Selection in Honey Bees;" and Seeley and Visscher, "Choosing a Home."

none of the individuals has the capacity to choose or even represent any of the nest sites as better or worse than any other. It is only through the coordinated activity of a number of bees, and only through the representation of particular facts about particular nest sites across various bees that this capacity can emerge. This coordination gives us good reason to think that there is a sufficient amount of emergent phenomena here to give the collectivity a rich life of its own. This research suggests two questions. First: Do these states and processes satisfy the four rough-and-ready desiderata on representation? And if they do: Should the relevant states and processes really be called "decisions" or "judgments"?

I suggest that the colony seems to have internal states and processes with the function of adjusting the hives behavior to facilitate skillful coping with changes in the environment. Consider the mechanisms that implement a decrease in foraging when too much food is coming into a hive too quickly. Here, none of the individual bees represents a need for a decrease in foraging, but the colony is designed to be sensitive to the relation between incoming nectar and nectar storage. When the rate at which nectar is being returned to the hive exceeds the rate at which it is stored, the system is designed to decrease the amount of nectar that is coming into the system. Moreover, this evaluation it is not a matter of absolute quantity of input or output; rather, it is measured by examining the relationship between the current state of a honeybee colony and the current state of the foraging sites in the area. Seeley clearly demonstrates that it is only by way of such internal states and processes that this sort of behavior, which is sensitive to changes in the environment, can be produced. Of course, there is still an important sense in which the behavior of the honeybee colony is fully a function of its evolutionary design. Unlike a human mind, a colony honeybee does not possess the sorts of representations that can be used to preselect behaviors on the basis of internal models. Indeed, the states and processes that we find in a honeybee colony share far more in common with the context-bound, and action-oriented, pushmi-pullyu representations that we find in the case of individual bees. So, along the spectrum of mental states and processes that I mentioned at the end of the last section, honeybee colonies are likely to have minimal minds, even though they are quite large in size.

Finally, let me close on a more technical note by suggesting that these colony-level representations share much in common with the computational structures that we find in a human brain. Each neuron in a human brain constantly updates its structure in light of the behavior of the other neurons to which it is connected (e.g., by modulating the production of neurotransmitters, extending and pruning dendritic branches to increase connectivity with preferred neighbors, adjusting firing-patterns to compensate the flow of neurotrophins, and dying when unable to integrate into the local environment). As Dennett (personal correspondence) has recently argued, neurons are likely to have taken on much of the Darwinian "research-and-development" carried out by the eukaryotic cells from which they are descended. The individual mind is a pandemonium that consists of numerous layers of demons, sub-demons, and sub-sub-demons, all competing for control of neural resources (e.g., by modulating the amount of neurotransmitter that it is available, extending and retracting dendritic branches, and by adjusting firing patterns). However, these competitions also modulate the structure of the brain's overall neural architecture—and unwittingly drive the computations that are carried out at higher levels. Because neurons inhabit a highly integrated, hierarchically organized, and massively parallel computational system, their competitive interactions yield computational outputs that can be consumed by the computational structure that they compose.

Research in the neurobiological sciences has suggested the presence of attentional mechanisms in the parietal cortex that depend on extreme winner-take-all computations.

For example, Christof Koch and his colleagues have argued that the allocation of "bottom-up" attention is the result of visual features (e.g., color, orientation, and movement) being encoded in separate feature maps that, when subjected to a competitive algorithm, can be integrated into a unified representation that encodes the strength of each feature in a topographically oriented saliency map that represents multiple values in a single multi-dimensional space.[83] Similarly, Robert Desimone and John Duncan offer a winner-take-all model of visual attention to explain how multiple competitions can eventually result in consciousness and subsequent reportability.[84] As Dennett has often argued, higher-level information is often processed in a way that "contributes to the creation of a relatively long-lasting Executive, not a place in the brain but a sort of political coalition that can be seen to be in control over the subsequent competitions for some period of time."[85] In each of these cases, we find competitive algorithms that integrate various sources of information to create unified representational structures that can readily be deployed in the production of action. The relevant representations might eventually come to be categorized in map-like or language-like structures in a semantically transparent language of thought; however, what is important for understanding their role in cognition has very little to do with representational genera, and a whole lot to do with the way that they can be rapidly deployed in skillful coping behavior.

I suggest that the competitive algorithms that are operative in honeybee colonies should not be seen as decisions, as decisions are the things that have to be aggregated out of beliefs and other cognitive sophisticated representations. However, as representational structures they do share much in common with the perceptual-motor and attentional structures in a human brain. While waggle and tremble dances, as well as search times, can stand-in for features of the environment (specifically the location of a food source and the rate or consumption by the system), they do so only when the system is immediately presented with raw data about the natural environment. The dance times as well as the vigorousness of an individual bee's dance are fully determined by features of the world, and the behavior of unemployed bees and collectors are fully determined by the dances of the returning forager bees. However, honeybee colonies are incapable of engaging in behavior that is anywhere close to being as rich as our own cognitive behavior. This being the case, even though I hold that we should recognize that there are collective mental states and processes in a honeybee colony, the mental life of a honeybee colony is far more impoverished than the mental life of a human being. This collective mind must itself be seen as a minimal mind, along some dimensions even more impoverished than the minimal mind of a single honeybee. The key point is that these emergent phenomena suggest an interesting range of cognitive phenomena that can only be studied by examining the behavior of honeybee colonies *as such*. These states and processes are cognitive to the same extent as the states of the neurons in my parietal cortex are cognitive states—they are not beliefs or desires, but they play a crucial role in structuring the way that the world is disclosed to me.

The upshot is that the range of explanatory projects that ought to be studied within the cognitive sciences is likely to outstrip the commonsense understanding of the mind;

---

[83] Christof Koch and Shimon Ullman, "Shifts in Selective Visual Attention: Towards the Underlying Neural Circuitry," *Human Neurobiology* 4 (1985): 219-27; and, Ernst Niebur and Koch, "Control of Selective Visual Attention: Modeling the 'Where' Pathway," *Neural Information Processing Systems* 8 (1996): 802-08.

[84] Robert Desimone and John Duncan, "Neural Mechanisms of Selective Visual Attention," *Annual Review of Neuroscience* 18 (1995): 193-22.

[85] Dennett*, Sweet Dreams: Philosophical Obstacles to a Science of Consciousness* (Cambridge, MA: MIT Press, 2005), p. 141.

and this is why psychology is best served by dissociating cognitive states at the subpersonal level from the core cases of cognition that might be present only in language users like us. There are many sorts of states that are important for explaining behavior, but we must be sure not to get carried away in ascribing states such as beliefs and desires to honeybee colonies—positing structures of collective mentality does not require assuming that honeybee colonies have beliefs, desires, hopes, wishes, and dreams.

5. CONCLUSION

If the arguments in this chapter are successful, they demonstrate that there is at least one respect in which commonsense assumptions about mentality should be revised: some groups of organisms—as such—can be in genuinely cognitive states and carry out genuinely cognitive processes. This conclusion stands in stark contrast to the commonsense assumption that cognitive systems are always organism bound. However, I have attempted to show that this conclusion is not nearly as counterintuitive as it might seem.

I first argued that adopting the computational approach to cognition that is advanced within the cognitive sciences provides us with strong reason to reject the assertion that the capacity for using language yields an unbridgeable gap between genuine cognition and witless mechanical activity. Contemporary approaches to the study of the mind begin from the assumption that even the most cognitively sophisticated capacities should receive a mechanistic explanation in terms of the representational and computational mechanisms by which they are implemented. On the basis of this argument, I suggested that different kinds of minds are likely to lie along a continuum running from minds that consist exclusively of pushmi-pullyu representations to human minds that deploy the richly amodal representations that are indicative of beliefs and desires of linguistically structured thought at the other end of the continuum. On the basis of these arguments, I suggest that the decision to treat some collectivities as minimally minded is no more and no less reasonable than the decision to treat an ordinary invertebrate as minimally minded. In both cases, these minimal minds are likely to represent the world by way of pushmi-pullyu representations. They are not likely to be populated by beliefs, desires, hopes, dreams and wishes.

Unfortunately, this argument leaves open difficult philosophical questions. Given that minimal minds like these are likely to be quite different from our own, my arguments do not have straightforward ethical implications regarding the use of poisons to kill cockroaches, the consumption of honey, or the use of silk. To answer such questions, one would have to move beyond these questions about minimal mentality and ask difficult empirical questions about the range of mental states and processes that populate the minds of invertebrates. With these data in hand, we would still face hard philosophical questions about the kinds of minds that warrant our compassion. Even offering cursory answers to such questions, however, would take me far beyond the scope of this chapter.