

Daniel D. Hutto and Erik Myin *Radicalizing Enactivism: Basic Minds Without Content*
Cambridge MA: MIT Press. \$35.

Daniel Hutto and Erik Myin (henceforth H&M) set out to establish their doctrine of “radical enactive cognition” by attempting to show that theories of content are utterly hopeless. The result is a rollicking campaign through the battlefields of content. The rhetoric is martial throughout, intimating a great insurrection. “Revolution is, yet again, in the air.” (1) They sound triumphant: “Far from being merely at the gates, the barbarians are, it seems, now occupying cafés and wine bars in the heart of the city,” (3) they say, casting themselves, presumably, in that role. And sometimes invasively colonizing: “Everybody imposes his own system as far as his army can reach,” says Stalin in the epigraph to chapter 3, which extolls the power of radical enactivism.¹ (Spoiler: The Red Army sweeps the field!) H&M devote chapters to the “retreat” of the theory of content (“Honorable retreats are no ways inferior to brave charges,” says William O. Douglas in the epigraph) and to its “last stand” (and in this chapter there are sections devoted to the phenomenal and the factual cavalries). To close, there is a chapter on “regaining consciousness” after the hard campaign. Ah, the high spirits of youth. The book has the verve and the panache, as well as the subtlety, of a hard-fought game of Paintball.

What is “radical enactive content” (REC)? It is, first of all, the view that “there can be intentionally directed cognition and, even, perceptual experience without content” (x), where content is understood as “conditions of satisfaction”. In other words, there can be cognition that is somehow *about* the world, not merely *in* the world (for this is what it means to be “intentionally directed”) but which nonetheless has no conditions of truth (or satisfaction).

REC is secondly the view that cognition is dynamic interaction with the world.

¹ The chapter epigraphs are totally in your face. Chapter 2: “He who does not see things in their depth should not call himself a radical—José Martí”. A hint of self-congratulation in a book titled “Radicalizing Enactivism.”

Mentality—with the possible exception of phenomenality . . .—is in *all* cases concretely constituted by, and thus literally consists in, the extensive ways in which organisms interact with their environments, where the relevant ways of interacting involve, but are not exclusively restricted to, what goes on in the brain. (7)

REC . . . uncompromisingly maintains that basic cognition is literally constituted by, and [is] to be understood in terms of, concrete patterns of environmental situated organismic activity, nothing more or less. (11)

Mentality is interaction; it is activity. Cognition lies as much in the world as in the brain.

Now, interaction is a relation; the H&M paradigm is that cognition is a relation between subject and world. Here, I mean ‘relation’ in its technical sense in logic: it takes *objects* as its terms, not proxies for objects, such as descriptions. Relations in this sense are not sensitive to the modes under which their objects are presented. Suppose that I *wrestle* with you. When we do this, I dynamically interact with you. Considering this activity by itself, and putting aside any mental process in the background, I am engaged with *you* when I wrestle with you, not with a mode of presentation of you. Nor is my wrestling with you “under a description”: if you are the tallest man in Regina SK, then I am wrestling with the tallest man in Regina SK, whether I know it or not. Wrestling is not a form of “intentional direction;” it is simply interaction. The question is this: In what way is cognition intentionally directed if it is a way that “organisms interact with their environments?” How, in other words, is cognition “intentionally directed”?

This leads to a more fundamental question: Is cognition somehow different in kind from wrestling? Or does wrestling count as “mentality” too? If there is a difference, what is it? This is an updated version of Aristotle’s question in Book II of the *De Anima*: how is perception different from a mere causal effect of the environment on the perceiver? Aristotle’s question was posed in the context of theories that made perception a unidirectional causal influence; it is equally relevant in the context of H&M’s bidirectional interaction. Aristotle thought that there had better be a difference between perception and processes such as my hair

being warmed up by the hot rays of the Sun. My guess is that H&M are not much bothered by this sort of question: on their account, cognition *is* very much like wrestling. (Maybe they'd allow that wrestling is one form of cognition.) "Ways of experiencing" are, for them, nothing but "ways of responding" (21). To assume otherwise is to fall into the "Hard Problem of Content." ("Hard Problem" . . . get it?) To suppose that experience is always aspectual (or, as Dretske had it, that "the objects I see *look a certain way to me*") runs into this Second Hard Problem. An aspectual account of experience is precisely to construe experience as "under a description." H&M say that any account of aspectuality or content runs into the Hard Problem of Content. So once again: how is cognition "intentionally directed?"

REC rejects representation. The *intellect* represents, H&M concede, but "it hardly follows that this type of cognitive engagement is the basis of, is required for, or is suitable for all sorts of tasks, always and everywhere" (41). This encapsulates one line of thought in the book. The static, linguistically structured representations of intellectual content are unsuitable for dealing with domains that are in a state of flux. (Is the theory of content committed to this kind of content? I'll come back to this later.) Positive argument for this thesis is lacking, however. Instead, the authors rely on Rodney Brooks, Randall Beer, and Barbara Webb. In the 1990s—the citations are all to papers in that decade—these three were champions of interactive, analogue approaches to cognition. The burden of understanding REC shifts to their work, and to philosophers such as Andy Clark (1996) and Evan Thompson (Varela, Thompson, and Rosch 1991), who built on these and similar ideas (also during the 90s). H&M's book is mainly negative in thrust: get the content theorists out of the way, and the city will fall.

To get a better idea of what enactivism is, let's take a look at Brooks's famous paper, "Intelligence Without Representation" (1991). Brooks kicks off with an argument so daft that one can only surmise that he meant it rhetorically. Human progress has been very fast, he says—"He invented agriculture a mere 19,000 years ago, writing less than 5000 years ago and "expert" knowledge only over the last few hundred years" (141). It follows that these accomplishments are "pretty simple."

Once dynamic environmental interactions have evolved, he writes, “problem solving behaviour, language, expert knowledge and application, and reason” are relatively small add-ons. Language is a just a gewgaw, Brooks suggests: to focus on it in the attempt to understand cognition is like time-travelling engineers of the 1890s being allowed to travel on a Boeing 747, and duplicating its seats and windows in an attempt to recreate “artificial flight.”

Now, Brooks is a great pioneer of a certain style of robotics. I’ll come back to this in a moment. But as a philosopher of mind and theoretical psychologist, he is prone to certain blind spots. He argues, as we have just seen, that since intellectual creativity emerged pretty recently—let’s say only 100,000 years ago—it must be a “pretty simple” supplement to systems for sensorimotor engagement, which have been evolving for billions of years. But the premise of this argument is dead wrong. The evolution of learning is very ancient indeed (in evolutionary terms). For example, relatively simple invertebrates sense and learn—Eric Kandel found there to be conditioning in the sea hare, a very simple and ancient organism possessing only about 20,000 neurons in total. (Actually, Kandel discovered the cellular realization of conditioning in the sea hare, thereby making the complication of the brain somewhat moot as far as learning is concerned.)

Now, conditioning is obviously behaviourally relevant. (That’s why it’s an essential part of the Behaviourist tool-kit.) When animals associate A with B, they attach to B a slightly different or weaker version of the unconditioned behavioural response appropriate to A. What Kandel found in the sea hare is the possibility of conditioning its gill withdrawal behaviour. This modification of behaviour is essentially an *enduring* change in the organism’s behavioural dispositions that allows it to adapt to an *enduring* environmental association. It *may* not be a discrete representation akin to those that the intellect produces (though see Gallistel and King 2010 for a pretty powerful argument that this is just what it is). But whether or not it is a discrete representation such as those that we find in von Neumann machines, it is emphatically not a dynamic interaction between organism and environment either. It is a stored representation of an environmental regularity.

Now, I have allowed that the conditioning of the sea hare might not result in discrete representation in the style of, say, your belief that democracy is on the rise in Africa (or your belief that it isn't, as the case may be). It is a very interesting question when and why discrete representations arose. My guess is that it has something to do with *versatility*. The conditioned response of the sea hare is the modification of a single behavioural disposition. But in higher animals, environmental learning can have multiple consequences for action. A dog that learns that to expect its food at five pm will do all sorts of things if it doesn't get it. Some of these are voluntary; it will sit by its bowl, bother its mistress, whine, etc. Some are involuntary; it will slobber, feel hunger pains, etc. A single learned association controls all of these behaviours. This argues for discrete representation. The whole organism cannot just be rewired so that it incorporates all of these context-dependent behaviours.²

However that might be, the evolutionary history of conditioning and learning shows that there is a very long history of cognitive evolution that is independent of sensorimotor evolution. Language and intellect are the products of that evolutionary stream as much as it is of any other. It is neither discontinuous with what went before nor a simple add-on to dynamic sensorimotor engagement.

Brooks also says a number of misleading things about perception. For example: "Certainly, there is no identifiable place where the "output" of perception can be found." I am not sure what this means: perhaps that the stream of consciousness is not to be found in an anatomically localized "Cartesian Theatre." However that might be, there certainly are identifiable places where the results of perceptual processing are to be found. For example, some of the results of face-processing are to be found in the cells of the fusiform face area. Brooks goes on to say, that "processing of the sensor data proceed independently and in parallel." Perhaps this is correct, but it does not gainsay the fact that there are areas of the brain in which certain precise perceptual data reside. When I recognize something

² I am grateful to Evan Thompson for productive correspondence about this issue.

as a face, and even more, when I recognize it as a particular person's face, I am drawing on the resources of a particular brain area. Why should I not see this brain area as encoding certain aspects of faces? H&M deny (64) that cognitive systems (aside, presumably, from "intellectual" ones) manipulate representations. But this seems false of the face recognition system. It employs schematic face-templates, and discriminates among faces on the basis of different values of parametric variables; moreover, it draws information about these parametric values from earlier processing of lines and shapes. Why doesn't this count as manipulating representations? (H&M present themselves as going beyond such conservative enactivists as Varela, Thompson, and Rosch 1991. It is worth noting that these writers seem to acknowledge my reasoning: see p 136, for example.)

Brooks's flagship example (dating back to 1987) is a robot that wanders around avoiding obstacles. In his introductory description of this machine, he says:

It is necessary to build this system by decomposing it into parts, but there need be no distinction between a "perception subsystem," a central system, and an "action system." In fact, there may well be two independent channels connecting sensing to action (one for initiating motion, and one for emergency halts), so there is no single place where "perception" delivers a representation of the world *in the traditional sense*. (147, emphasis added)

The traditional idea of obstacle avoidance relied on an egocentric map of the surrounding area. Brooks found that this was not necessary. He talks repeatedly about "data" and the like, but protests:

Even at a local level we do not have traditional AI representations. We never use tokens that have semantics that can be attached to them. The best that can be said in our implementation is one number is passed from a process to another. (149)

The second sentence above sounds perversely like Fodor's syntactic theory of mind: the machine runs by the interactions of its internal tokens without knowing its own semantics. But this is not the question. The question is really: Does it *have* semantics? Here's a sample of what Brooks says:

The finite state machine labelled *sonar* simply runs the sonar devices and every second emits an instantaneous map with the readings converted to polar coordinates. This map is passed on to the *collide* and *feelforce* finite state machine. The first of these simply watches to see if there is anything dead ahead, and if so sends a *halt* message . . . Simultaneously, the other finite state machine computes a repulsive force on the robot, based on an inverse square law . . . (153)

I am not suggesting that this kind of agential talk should be taken literally. My point is that it provides a design perspective on the machine without you cannot comprehend the setup. In an evolutionary setting, this kind of description shows us *why* an organic system has the external connections that it does. In short, it tells us what environmental significance various state transitions possess. And, as Dretske says, the machine's sensing is aspectual: things are sensed as having a location; they seem a certain way. It is in virtue of this semantics that it is an "intentionally" directed system, not merely an interactive one.

I have described Brooks's paper in some detail because H&M rely on it and a couple of other papers for their understanding of REC. Their positive characterization and argument is surprisingly brief. Yet, their summary is almost totally misleading. Here's the bottom line: Brooks's machines are guided by "continuous, temporally extended interactions . . . rather than being based on represented internal knowledge . . . (knowledge that would presumably be stored somewhere in the robots' innards)" (42). It is certainly true that the machine's knowledge and goals are reflected in its design rather than in what it gathers and stores. But this is the distinction between *innate* and *acquired* knowledge, the former being tacit and implicit in the inherited phenotype.

It is true that Brooks's machines store no information. That is absolutely not a reason to say that we similarly store nothing, except in the language driven "intellectualist" parts of our minds. As I tried to point out earlier, the answer is a clear no. Animals, even very simple and primitive animals, are learning-driven agents. In this respect, they are very different from the machines that are discussed in this book. True, learning can be conceptualized as an architectural modification.

Learned information does not have to be explicitly stored. But this does not mean that we cannot give semantic content to the modification. If an organism finds that a certain kind of fruit is sweet and pleasant to eat, it learns to seek out that fruit by its look and smell. This is a new behavioural disposition, but it is an internalization of the information that the fruit is good to eat.

Finally on this point, Brooks's machine and Barbara Webb's insect robots (which are also inspirational to H&M) are hierarchically constructed modules. Most *organisms* are multimodular. For example, Webb's cricket robots (1996) hone in on auditory signals just as real crickets do; real crickets, however, don't do just this—they also eat and mate and sleep. Real crickets, then, have a multiplicity of modules and, minimally, inter-module signal-passing that prioritizes what the cricket will do. If a cricket is hungry, its feeding module is activated and its mating module is suppressed. All of this demands inter-module communication. On natural assumptions, these communications have a semantics: "I am hungry; this is not the time to make whoopee!" for example.

Further, where there is inter-module communication in perception, there is similarly meaningful communication. For example, Marc Ernst and Martin Banks (2002) argue that when two senses (e.g., vision and touch) deliver inconsistent messages about some feature of the world (e.g., the shape of an object), each sense delivers a percept that depends on the information supplied by both, and by the estimated reliability of each sense. For example, when the visual percept appears degraded, the visual information is discounted and the tactile signal counts for more. This kind of computation requires meta-content (content about the reliability of perceptual content) as well as inter-module communication. It is certainly true that this communication will be analogue in character, and hence it will not correspond to traditional computer architectures. Nevertheless, there is no reason to abstain from interpreting such processes as being directed to a summation of data from different sources.

If there is a theme in this book, it is that causal entanglement is capable of supplanting content. Here's a quote H&M take approvingly from Anthony Chemero (2009):³

The representational description of the system does not add much to our understanding of the system . . . [Thus,] despite the fact that one can cook up a representational story once one has the dynamical explanation, the representational gloss does not predict anything about the system's behaviour that could not be predicted by dynamical explanation alone. (Chemero 2009, 77; quoted on 58-59)

I have tried to show why I disagree with this. First, dynamical explanations are not necessarily transparent with regard to what a system is doing. You may not, for example, understand some routines embedded in the code that controls Brooks's robots until you realize that the machine would be avoiding an obstacle by executing this routine. It may be excessively didactic to formalize this by intoning a practical syllogism, thus: "The machine is designed to avoid obstacles; if routine R is activated, then it will avoid an obstacle; it had no other way of avoiding the said obstacle; so it activated R." Didactic or not, this is, roughly speaking, why the routine R is built into the system, and why it is hierarchically subordinate to other parts of the code. Secondly, and perhaps more substantively, real creatures learn and real creatures contain multiple modules that communicate with each other. In order to balance the competing needs of these modules, messages must be passed about their current state. These internal operations demand representations.

Let us move now to another theme in H&M's critique of the theory of content. They insist that covariance does not constitute content. Their point is just that "anything that deserves to be called content has special properties—e.g., truth, reference, implication . . ." (67), and that covariance does not constitute semantic

³ Here's a quote from Chemero (2009) that I approve of: "Books by philosophers almost always begin by arguing that everyone else is incorrect. I have just explained why nobody believes these arguments. I will, therefore, not be presenting arguments that all approaches other than radical embodied cognitive science are somehow bad" (16). The bulk of H&M's presentation consists of arguments that the theory of content is bad.

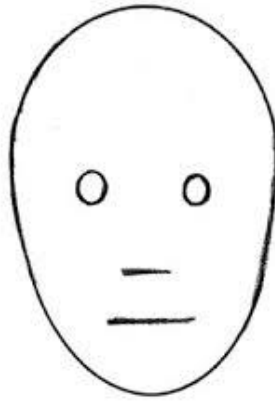
properties such as truth, reference, and implication. They are absolutely correct. A classic intellectualist case illustrates why. Suppose that

1. Sally believes that p , and
2. Sally believes that if p then q .

Then, we can expect that Sally will have a causal disposition to believe that q . Now let Sally's first belief causally covary with the truth of p , and let her second belief similarly covary with the truth of the conditional "If p then q ." Clearly, these covariance relations are no guarantee at all that Sally will have a causal disposition to believe q when she has the two beliefs stated above. That is, there is no guarantee or expectation that states that carry information in this way will interact in ways that reflect the logic of the information they carry.

That much granted, here's an explanation of why Sally's beliefs follow the above pattern. Human beings are so constructed as to undergo belief transitions that reflect the contents of their beliefs. Content is information *provided that* it is carried by states that behave in ways appropriate to the logic of that information (where 'logic' is construed broadly to include not only deductive logic, but all sorts of inductive and pragmatic rules). In short, content is a system property. It is a property of states in a system that treats information carrying states in a certain way.

Now, the example of Sally's belief's is, of course, intellectualist, but my claim is that the visual system (for example the visual face-recognition system) uses other states of the visual system in characteristically content-driven ways. Here's an example that is concerned with image content:



Each of the component shapes above—circles, horizontal lines, oval—might be detected by cells in one area of the brain. The fusiform face area will be active when these shapes are aligned together in the above way. It will be inactive when the shapes are arranged some other way. Here, we have cells that covary with the smaller shapes, with the system so arranged that a particular alignment of these activates a further covariant system state. The covariance of the component states does not guarantee this; it is the operation on information-carrying states that makes them content-bearing—the computational properties of the fusiform face areas is dedicated to detecting patterns to generate further covariance. The Bayesian program in perception is another example: state transitions in perceptual processing obey Bayesian laws of probabilistic inference with regard to their contents.

H&M make a serious misstep with regard to these matters. As I said, they are correct to say that content is not mere covariance. But they infer from this that the theory of content is going to have to adopt “a notion of information [that] is meatier than covariance but . . . nonetheless equally naturalistically respectable” (69). In other words:

3. Content has semantic properties
4. Semantic properties do not accrue in virtue of covariance alone.

5. Therefore, the kind of information that representational states carry is “meatier than” covariational information.

This is wrong. I agree with 3 and 4. But the conclusion does not follow. I argued above that states possess content because they are part of a system in which their causal role is dependent on the logic of the information they carry. This does *not* imply that these states carry information in some way other than by covariance. What it implies is *just* that the system in which they occur is so organized as to respect the logic of the information they carry.

Now, just for the record, I do not think that content is information-as-covariance. As I said above, I agree with 4. I hold rather that the relevant kind of information is teleosemantic (Matthen and Levy 1984, Matthen 1988). But this point is totally tangential to H&M’s critique. My contention is that these “special properties” devolve to states because of system properties—i.e., because the system is so organized as to treat these states in ways that accord with the logic of the content they carry. H&M’s critique does not give us reason to think that the kind of information contained in these states is anything other than covariance.

So much then for the “Hard Problem of Content.” This problem is nothing other than that the notion of content is overly “metaphysical.” It has to be said that the sudden intrusion of metaphysics-critique felt a bit diversionary. Enactivism is usually presented (e.g., by Brooks or by Evan Thompson) as a strategy in cognitive science, not a strategy for keeping the philosophy of mind metaphysically clean. Brooks is not driven to his view about content by worries about supervenience and physicalism; Thompson is not either. Both would surely agree that if they ever became persuaded of representations of the traditional sort, metaphysics would not deter them. Find the right architecture of mind, and the naturalizing strategy will follow. Even dualism can be made compatible with all that matters in physicalism (or so I would argue⁴). This is why I see H&M’s arguments about naturalizing

⁴ <http://tinyurl.com/pdmtaqn>

content almost completely irrelevant at this stage of the inquiry. Never mind. There is no effective argument here against even a crude naturalization of the notion of content.

The book closes with a discussion of phenomenality. One strand of the argument is insightful. You can't understand phenomenal experience without understanding how it relates to the subject's active participation in the surrounding world. "If we want to understand the place of phenomenality in nature—how it originally came to be, and why it has the features it has—we are likewise forced to widen our scope" (163). This seems right to me: you can't understand much about how *red* looks if you didn't say something about colours and surfaces and the like. Of course, this is not going to give us a material account of consciousness as such. H&M recognize the fundamental difficulty for enactivism that Jesse Prinz has so lucidly exposed: "Just as it's hard to understand why brain states feel a certain way, it's hard to understand brain states together with bits of the external environment would feel a certain way" (Prinz 2006, quoted by H&M at 156).

This is where things become difficult. For here is how the Hard Problem (of Consciousness) is supposed to "dissipate." The first step is to "allow that phenomenal experiences admit of true physical descriptions, *without assuming that such descriptions exhaust or fully characterize all aspects of their nature*" (168). Given the italicized rider, Chalmers and other dualists should be happy with this. Perhaps they would even agree to the next rider—that phenomenal descriptions of experiential states are not "ultimately privileged" over physical descriptions. But, they now say, we shouldn't think that "phenomenal descriptions might be inferred *a priori* from physical descriptions."

But isn't this to give the game up? Chalmers has a (much contested) argument from two-dimensional semantics that if phenomenal descriptions *are not* inferred *a priori* from physical descriptions, then there is an "ontological gap" between the two. Perhaps this argument is mistaken, but H&M don't attack it. But if Chalmers' argument stands, then it's hard to see how it helps to "take

phenomenality to be nothing but forms of activities . . . associated with environment-invoking activities” (169). Activities feel a certain way, not experience: we talk of what it is like to savour, taste, and feel but not about “what it is like to experience the ‘redness of red’” (177).

This seems simply to ignore Prinz’s dictum altogether. H&M seem just to assert, baldly and unashamedly, that brain states interacting with bits of the external environment feel a certain way. I don’t understand how this is supposed to yield any kind of approach to the Hard Problem of Consciousness. Characteristically H&M don’t help much. Having laid out their proposal in some three pages, they change direction again and launch an attack on David Papineau.

Radicalizing Enactivism is a fun book. Its high spirits and gleeful bad-boy antics are a much-needed antidote to the cheerlessness of much academic philosophy. But it achieves very little. It mounts a heavy attack on the theory of content, but this attack feels disjointed, a bit like throwing everything they can find at their opponent, with little regard for a coherent positive account. Well-known theories that, on the face of it, demand representational content are ignored. The kind of theory that is invoked is not followed through beyond what is needed for the performance of a single task.

REFERENCES

- Beer, Robert (1995) "A Dynamical Systems Perspective on Agent-Environment Interaction," *Artificial Intelligence* 72: 173-215.
- Brooks, Rodney (1991) "Intelligence without REepresentation," "Intelligence without Representation." *Artificial Intelligence* 47: 139-159.
- Chemero, Anthony (2009) *Radical Embodied Cognitive Science*. Cambridge: MIT Press.
- Clark, Andy (1997) *Being There: Putting Brain, Body, and World Together Again* Cambridge MA: MIT Press.
- Ernst, Marc O. and Banks, Martin S (2002) "Humans Integrate Visual and Haptic Information in a Statistically Optimal Fashion," *Nature* 415: 429-433.
- Gallistel, C. R., & King, A. P. (2009). *Memory and the Computational Brain: Why Cognitive Science Will Transform Neuroscience*. New York: John Wiley & Sons.
- Matthen, Mohan and Levy, Edwin (1984) "Teleology, Error, and the Human Immune System," *Journal of Philosophy* 81: 351-72.
- Matthen, Mohan (1988) "Biological Functions and Perceptual Content," *Journal of Philosophy* 85: 5-27.
- Varela, Francisco J., Thompson, Evan, and Rosch, Eleanor (1991) *The Embodied Mind: Cognitive Science and Human Experience* Cambridge MA: MIT Press.
- Webb, Barbara (1996) "A Cricket Robot," *Scientific American* 275.6: 94-99.