Commentary on Brette (2019): Is coding a relevant metaphor for the brain?
https://doi.org/10.1017/S0140525X19000049

Computational neuroscience needs theoretical ethology: correcting an anthropomorphic bias

Alex Gomez-Marin
Behavior of Organisms Laboratory
Instituto de Neurociencias (CSIC-UMH), Alicante, Spain
agomezmarin@gmail.com
https://behavior-of-organisms.org/

Brains enjoy a bodily life. Therefore animals are subjects with a point of view. Yet, coding betrays an anthropomorphic bias: we can, therefore they must. Here I propose a reformulation of Brette's question that emphasizes organismic perception, cautioning for misinterpretations based on external ideal-observer accounts. Theoretical ethology allows computational neuroscience to understand brains from the perspective of their owners.

An apparently innocuous word in Brette's question is a major source of confusion but also contains a great deal of the answer. Is coding a relevant metaphor for "the" brain? Yes and no. It depends on whose brain we are talking about. For the scientist studying the animal, coding is certainly relevant (at least, as the ubiquity of such figure of speech attests in current neuroscience). But, in so far as we are interested in the animal and its brain, the answer is likely no. The mantra "stimulate, record, correlate" misses the point of the organism. It is for us, by us. That the experimenter's model can decode the signal does not mean that the brain can or does. The information necessary to make sense of the data in terms of coding is seldom available to the organism, upon which coding is predicated. This creates a can-ought problem: a description of what the neuroscientist can do prescribes what the animal must do. Such implicit tension pervades most of the disagreements that Brette's question shall spur. The problem, I believe, is deeper than coding: there is a conflict of interests between the scientist and the laboratory animal.

Biology is the science of living beings. Organisms are centers of action. As such, perspective matters. To be an organism is to have a point of view. All animals share a common world but not all animals have a world in common. Each living organism has its own *Umwelt* (meaningful environment), which is different than its *Umgebung* (physical surroundings): a tree is a tree, but a tree for an ant has little to do with a tree for a carpenter [1]. What is meaningful for an organism —or even what is possibly apprehensible— need not be meaningful for the scientist studying it, and vice-versa (a concrete and pervasive example: stimuli are more the experimenter's output than the animal's input). The use of the definite article ("*the* brain") or the indefinite pronoun ("*one* finds") is so delicate in biology. It easily blurs the subject (I? you? the mouse? what mouse?), unbinding grave conflations and misleading thought and interpretation. Eloquently said: "Hedgehogs as such do not cross roads (…). On the contrary, it is man-made roads that cross the hedgehog's millieu" [2]. Rather than being an exception, coding illustrates such misattribution. Paraphrasing, we could say that cat brains as such do not encode stripes, but it is stripes that we decode from the cat's brain. A clash of *Umwelts* is going on in our laboratories.

The notion of *Umwelt* has no place in physics; it does not violate physics, but it is not reducible to physics either. Living beings inhabit a world of meaning that includes but

exceeds the physical world of masses and forces, and even more so the mathematical world of zeros and ones. The appreciation of the uniqueness of biology discords with a cornerstone of the scientific approach: objectivity. Of course we always observe reality from a viewpoint, explicitly or implicitly chosen. But it is ultimately deemed irrelevant. Objectivity, then, is the pretense of self-exclusion from the phenomenon under study. The observer vanishes in classical physics (also in biology). By means of a representation of things that ultimately does not depend on the reference system, an observer-independent reality is erected. Yet, "[o]n the strength of the immediate testimony of our bodies *we* are able to say what no disembodied onlooker would have a cause for saying: (...) the point of life itself: its being self-centered individuality" [3]. From subjectivity we have prodigiously built an objectivity that can dispense with the former. However, upon inspection, objectivity becomes a particular kind of inter-subjective consensus. This is biology's scotoma: we are subjects whose objects of study are subjects too.

In behavioral neuroscience there is an observer-observed gap. Physiology aspires to study the inner workings (brain) of an organism from the outside (scientist's perspective); ethology strives to understand the outer happenings (behavior) from the inside (animal's perspective). Isn't the neurophysiologist's decentering a covert self-centering? Sticking electrodes is not sufficient to know what it is like to be a rat. But, how to look through the animal's eyes? A cute example is Turtle Geometry: it actually matters if a turtle traces a circle by solving the $x^2+y^2=r^2$ equation, or by iterating a "run and turn" procedure. Both are mathematically equivalent (from an external ideal observer, perhaps indistinguishable, even irrelevant) but biologically they are not the same. There is much to gain from discovering "the range of complicated things a turtle can do in terms of the simplest things it knows" [4]. What is it to make sense from the animal's perspective when it does not do so the way we do? Such is the paradox: the *Umgebung*, the objective world of scientists, can be part of our human *Umwelt* (we do not feel neutrinos crossing our bodies, but we can detect them in bubble chambers), but it collides with the *Umwelt* of the animal, which is never an *Umgebung*. Neuroscientists yearn for neural codes; the animal has no clue.

Neuro-ethology is actually meta-engineering: our problem is to solve how animals solve their problems—to scientifically empathize with each creature. This entails a revision of Bernard's foundational words: the scientist "no longer hears the cry of animals, he no longer sees the blood that flows, he sees only his idea and perceives only organisms concealing problems which he intends to solve" [5]. By reformulating Brette's question, my intention here has been to emphasize that computational neuroscience can benefit from the insights of theoretical ethology to transform its anthropomorphic bias. To crack codes, "it would suffice that we be angels. But to do biology, even with the aid of intelligence, we sometimes need to feel like beasts ourselves" [2]. The question then is not so much whether coding is relevant or wrong, but to what extent it is misleading. We must then ask: Whose brain is the coding metaphor relevant for?

Acknowledgements: I thank Ehud Ahissar, Konrad Kording, Spyridon Koutroufinis, Ibrahim Tastekin, Zach Mainen, and specially Asif Ghazanfar for insightful discussions.

References:

[1] Uexküll JV (1926) Theoretical Biology. Oxford, England: Harcourt, Brace & Co.
[2] Canguilhem G (2008) Knowledge of Life. Fordham University Press.
[3] Jonas H (2001) The Phenomenon of Life. Northwestern University Press.
[4] diSessa A, Abelson H (1981) Turtle Geometry. MIT Press.
[5] Bernard C (1957) An Introduction to the study of Experimental Medicine. Dover Publications.