On the Neurobiology of Truth

Ron Bombardi
Department of Philosophy, Middle Tennessee State University
http://www.mtsu.edu/~rbombard/RB/

1. Introduction.

This talk is a philosophical exercise: it is not a talk about things, but a talk about our talk about things.

(a) Abstract. The concept of truth arises from puzzling over distinctions between the real and the apparent, while the origin of these distinctions lies in the neurobiology of mammalian cerebral lateralization, that is, in the evolution of brains that can address the world both indicatively and subjunctively; brains that represent the world both categorically (*as-is*) and hypothetically (*as-if*).

(b) Elevator Words. Philosophical investigations do not question a different world than the sciences question, but philosophers do question the world differently; typically, not from the vantage of mathematical analysis, but from perspectives afforded by what Canadian philosopher, Ian Hacking, calls "elevator words."[1] We use elevator words not to ask about things-in-the-world (tables, atoms, laughter, taxes), but to ask about things *about* things-in-the-world (facts, truths, principles, proofs, meanings).

(c) Two Flavors of Truth. We elevate our perspective in this way when we notice that there are two general classes of things we call true:

(i) what is or can be said;
(ii) everything else.

In the first category, we have things like propositions, judgments, beliefs, representations; and in the second, things like crosses, hearts, spaniels, and North.

(d) Family Resemblances. Much can be said about the grammar of attributions of the property, *true*, to the various components of the non-discursive world, but philosophical speculation has over the centuries focused largely on how the linguistic predicate, "is true," functions in discourse. (I shall now follow suit, but intend to return to the "everything-else" category towards the end of the talk.)

Of the things that are or can be said, the expression, "is true," is or can be *used* in several semantically distinct, if overlapping, ways, forming a set of what Wittgenstein called *family resemblances*. When we aggregate these uses, we find for expressions of the form "*S* is true," (that is, when "is true" is predicated of a given sentence, *S*) at least three quite distinct functions: to say that:

(i) the elements and relations in *S* sufficiently match objects and relations in the world;
(ii) S is sufficiently consistent with other sentences, S_1, S_2, ... S_n, in a given system;
(iii) holding S is sufficiently rewarded in experience.

Not to be outdone by a mere social construct like grammar, philosophers have long offered to advance their own Theories of Truth.

2. Theories of Truth.

After some 2500 years of thinking about it, the Western philosophical tradition has come up with three major theories of truth: correspondence, coherence, and pragmatist.

(a) The Correspondence Theory. This theory forms the basis of Aristotle's logic, runs rampant through mediaeval scholasticism, was famously held, then rejected, by Wittgenstein, and recently given a complete, consistent formal analysis by Alfred Tarski. Aristotle's version is perhaps the clearest: "to say of what is that it is or of what is not that it is not, is true."[2] G.E. Moore formulated the same view in more modern idiom thusly: "To say that something is true is to say that there is a correspondence between it and a fact."[3]

These are both rather thin formulae; their purport is often thickened with assumptions (e.g., that the relevant correspondences only obtain when the two items enjoy separate and independent existence). Moreover, the effort to specify the meaning of the verb, "to correspond," has yielded little consensual precision. Nonetheless, the theory remains robust, especially among logicians and philosophers of mathematics.

(b) The Coherence Theory. The coherence theory of truth makes its first, documented appearance on the philosophical scene in the teachings of Zeno the Stoic and his successors, notably Philo of Megara and Diodorus Cronus. Later adopted by the rationalists, Spinoza and Leibniz, as well as by the idealists, Hegel and Bradley, and eventually championed by logical positivists Neurath, Hempel, and Quine, this view holds that to predicate "is true" or "is false" of a judgment or proposition is to say that it coheres or fails to cohere with a system of other things judged or said. Systems, in turn, are understood as logical constructs, whose elements are related to one another by logical implication. One of Quine's dicta gets squarely to the heart of the matter: "No statement is immune to revision."[4]

Unlike correspondence theories, it should be noted, coherence theories typically recognize *degrees* of truth (rumors are said to contain less truth than direct observations). The systematic exploration of this idea (that truth comes in degrees) probably began with Descartes' worries over certainty; it has received considerable formal elaboration in recent decades with the advent of fuzzy logic and fuzzy set theory.

(c) The Pragmatist Theory. Classical pragmatism originated with the doctrines of the great Athenian Sophist, Protagoras, and his circle, but, running counter to the epistemological inclinations of Plato, Protagorean pragmatism fell into disrepute, even in antiquity, forming some of the first ranks of radical philosophy in the Western tradition. Modern pragmatism, on the other hand, descends from the more respectable lines of German idealism, although its genome is primarily American (Peirce, James, and Dewey, its principals, with a nod to the Englishman, F.C.S. Schiller).

Truth is a property of ideas, according to the pragmatists, and ideas are instruments or tools; a true idea is an idea that fulfills a function; false ideas to not. Dewey devised a particularly clear rendition of this view in 1920:

> ... if ideas, meanings, conceptions, notions, theories, systems are instrumental to an active reorganization of the given environment, to the removal of some specific trouble and perplexity, then the test of their validity and value lies in accomplishing this work. If they succeed in their office, they are reliable, sound, valid, good, true.[5]

3. An Immodest Proposal.

(a) The Tradition & Its Discontents. So, we have three major theories of truth, all with ancient pedigrees and contemporary proponents. In fact, traditional philosophy has failed to arbitrate much among the three principal views; certainly no clear winner has emerged.

But this should come as no surprise, especially once we trace the origin of each theory to its grammatical roots, that is, to the distinct uses human languages have found for the expression, "is true," and its respective cognates. Later Wittgenstein modestly proposed that we stop exactly here (that's why he had no theory of truth after rejecting his own earlier *Tractatus*).

(b) Postmodernism. Much contemporary thought follows Wittgenstein's lead, and tends accordingly toward disciplinary boundaries, compartmentalization, deconstruction, and other forms of postmodern fragmentation. Like Delilah, with a ring on every finger, and a finger in every pie, nowadays we have a theory for every project and a project for every problem (e.g., we now have logical, empirical, conceptual, self-evident, aesthetic, and moral flavors of truth, just to name a few).

(c) Unification from Neuroscience. My proposal goes against this grain in recent thought. I argue that contemporary neuroscience provides adequate theoretical grounds for a unified theory of truth. More specifically, I shall contend that the correspondence, the coherence, and the pragmatic utility of symbols are each *biological* features our neurophysiological information processing systems—that is to say, our brains. On my view, the traditional trifurcation of philosophical accounts of the predicate, "is true," stems from a trifurcation of focus on the information latent in sensory, motor, and somatosensory cortices of the human brain.

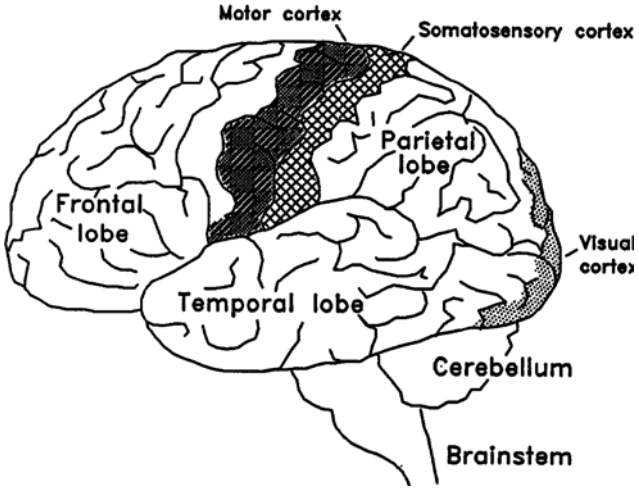

Figure 1. The location of some of the primary and secondary sensory and motor areas within the primate cerebral cortex From Paul M. Churchland, *The Engine of Reason, Seat of the Soul: A Philosophical Journey into the Brain*, p. 129.

4. Before Truth.

(a) Long, Long Ago. Around 150 million years ago, there was no truth on this planet. What made the appearance of truths possible was the evolution of placental and marsupial mammalian brains—when our own distant ancestors parted reproductive company with the monotremes; when hemispheric lateralization spawned a new kind of brain, a brain that can dream.

(b) Evolutionary Feedback. We now recognize that anatomical and physiological differentiation form feedback loops in evolutionary development. From the first swelling of the cerebellum shared with reptiles to the neocortical bulges where humans figure the contours of fugues, altered forms have led to adaptive functions, which have led to further anatomical transformations that have further expanded physiological potentials, and so on.

Cortices and nuclei in mammalian brains eventually developed independent, *specialized* functions; and along with hemispheric lateralization, these multiple, specialized functions have afforded mammalian brains with two salient abilities:

(i) recombinant feature extraction and composition;
(ii) off-line processing (*as-if* loops).

(c) The Invention of Imagination. The net *psychological* result of these evolutionary developments is the *invention of imagination*; that is, the evolution of animals capable of forming hypotheses, which is to say, animals capable of dreaming. This was a monumental step in brain evolution: when the first firing patterns in and between sensory and pre-motor cortices were excited while related circuits in motor cortex were inhibited; that is to say, when the first dream on earth occurred.

Daniel Dennett has ingeniously dubbed the first dreamers and their descendants, "Popperian Creatures," after Karl Popper's celebrated portrait of inductive reasoning as a logical space in which we let "our hypotheses die in our stead."[6] Before the advent of this logical space, courtesy of inhibitory neurotransmitters originating in the pons and projecting to cerebellum, superior colliculus, and cerebral cortex, there was no truth in the world because there was *no functional distinction* in brains between the *as-is* and the *as-if*, between the categorical and the hypothetical, between the real and the apparent, between externally and internally generated stimuli.

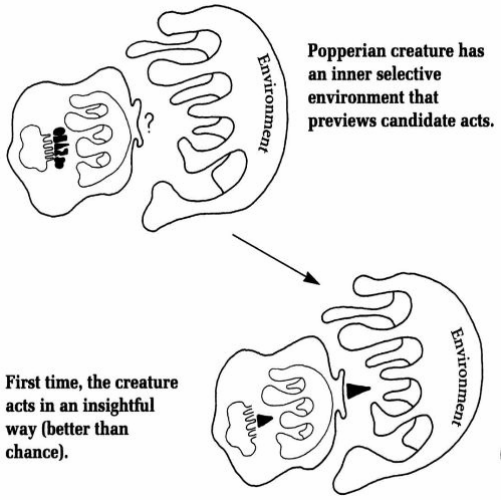

Figure 2. The model implies that "there must be a filter, and any such filter must amount to a sort of inner *environment*, in which tryouts can be safely executed--an inner something-or-other structured in such a way that the surrogate actions it favors are more often than not the very actions the real world would also bless, if they were actually performed. In short, the inner environment, whatever it is, must contain lots of *information* about the outer environment and its regularities." From Daniel Dennett, *Kinds of Minds, pp. 88-89.*

5. A Recipe for Rationality.

(a) The Four Fs. Seen in evolutionary perspective, hypotheses are about survival, endowing organisms that can form them with a *recipe for rationality*, a recipe we can abbreviate in a slogan:

Fiction becomes Fact when Forecasts are Favored.

(b) Cartoon of the Elements. The essential ingredients in this recipe are fourfold; we can sketch a cartoon of their interactions from the following characterizations:

Fiction: circuits active in subjunctive convergence zones; a product of *as-if* loops redistributing primary sensory and pre-motor data through cortico-subcortico processing without excitation of corresponding motor circuits.

Fact: circuits active in indicative convergence zones; a product of *as-is* loops; effectively, on-line coordination of sensory, pre-motor, somatosensory, and motor circuitry, mediated by cerebellum and limbic processing.

Forecast: running active *as-if* loops; may be either conscious (pre-frontal activity) or unconscious; typically involves activation of pre-motor areas.

Favored: reinforced loops, where reinforcement is of two sorts, short- and long-termed. Short-termed favoring is a matter of enjoining sympathetic firing frequencies among neuronal clusters; what William Calvin calls, "winning a cloning competition."[7] Long-term favoring occurs at the subcellular level, via increases in receptor sites proportional to the period of stimulation.

(c) What the Cartoon Tells Us. The cartoon tells us three important things:

(i) the brain is a self-referential system.

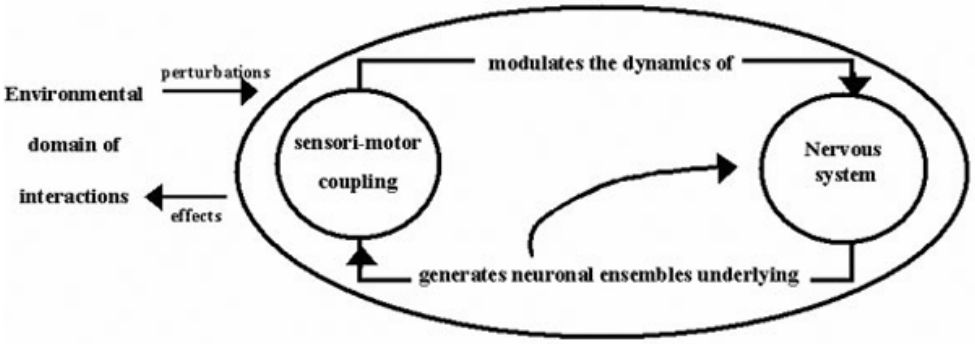

Figure 3. The operational closure of the embodied system. As a circular process, an individual is engaged in the continuous cycles of operation defined by its eigenbehaviors. Three levels of circular causality are distinguished in the figure: (i) the level of the central nervous system as a closed dynamical system; (ii) the level of the sensory-motor mutual definition of the state of the brain and of the body; (iii) and the level of the ongoing coupling between the autonomous system and its surroundings, including potential inter-individual interactions. From David Rudrauf , et al., "Autopoiesis to neurophenomenology: Francisco Varela's exploration of the biophysics of being." Laboratoire de Neurosciences Cognitives et Imagerie Cérébrale, Paris, France.

http://www.scielo.cl/scielo.php?pid=S0716-97602003000100005&script=sci_arttext

As pioneering neuroscientist Rodolfo Llinás has it, "the brain is a closed system in which the role of sensory input appears to be weighted more towards the specification of ongoing cognitive states

than toward the supply of information—context over content. This is no different than input modulating a pattern of neural activity generated in the spinal cord to produce walking, except that here we are talking about a cognitive state generated by the brain and how sensory input modulates such a state. The principle is the same."[8]

(ii) the brain is essentially a dream machine that constructs virtual models of the world, including itself.

(iii) that arriving at truths is always and necessarily an inductive affair, a dynamic process of building images (of both the body and the world) from massively redundant cortical and subcortical cross-talk.

This much of the story Hume predicted. But notice that the cartoon doesn't tell us precisely *how* brains emulate reality. Hume famously despaired of our ever having such an account.

6. The Neurobiology of Truth.

(a) The Argument. I want to argue next that Hume's despair was premature, which is to say that we can in fact now assemble a thoroughly mechanical model of how brains make worlds.

(b) Two Keys. The key to this model is the single neuron; the excitable cell. Likewise, but an octave higher, the key to the connection between neurons and facts is *animal motion*, that is, the generation of a physical response to a physical stimulus.

The simplest sort of animal motion is seen in reflex arcs (such as those responsible for appropriate muscle contraction and relaxation as you grip a pen to write this idea down on paper, or support the weight of a beer bottle after the talk). In mammalian brains, neuronal assemblies negotiate the transduction of multiform information from sensory peripheries to bodily motion via multiple, re-entrant, recombinant, intermediary stages.

(c) Representations. Localized to convergence zones, representations occur during these interim stages, as patterns of neuronal activity from disparate cortices and nuclei are assembled into the objects of perception, imagination, and abstract thought. Representations from the semantic to the kinetic are effectively amplified, enhanced, synthesized, and compressed loops of sensory/motor replay. This much, we can say, Schopenhauer predicted.

(d) Desire Drives Reason. But we need not invent transcendental forces to account for how and why desire drives reason. We need only recognize that the *province of proof* (operating on abstract symbols) is marked by the *method of measure* (operating on concrete stuff); measure, as the ancient Greeks well, knew, mediates between sensation and knowledge.

By "measure" I mean the neuronal activity of coordinating sensory/pre-motor representations (*as-if* loops) against the background of somatosensory representations, such that the current body-state serves as an evaluation function, collapsing possible (hypothesized) outcomes of an action into an actual motion. This view of the nervous system has been pioneered and extensively developed by Hannah and Antonio Damasio.[9]

(e) What Symbols Do. Symbols arise from the persistent activation of representations; and they come to serve as stand-ins or surrogates for bodily events, not things. (Aristotle and Locke were

right; the symbol "5" does not name an abstract entity but stands for a counting event.) Proof, or reasoning, can thus be understood as a *search* for symbols to function as measurements (that is, to initiate somatosensory *as-if* loops sufficient to collapse cortico-cortico variation).

What symbols do is enjoin the convergence of disparate bits of cortical replay (a proper name might stimulate the simultaneous activation of a scent, a look, and a tone-of-voice, for example). Symbols can piggy-back too, so that symbols for collections of symbols arise as an organism adds cortico-cortico circuits to its existing repertoire.

In sum, neurobiology points to the overwhelming conclusion that thinking ultimately represents movement.

7. Philosophical Implications.

(a) A Unified Theory of Truth. A neurobiological perspective on how brains think additionally shows how the three traditional theories of truth can be unified into a single account:

Truths are correspondences all right, but these correspondences are primarily temporal, not spatial, and appear between neocortical maps and sensory/motor maps, mediated by somatosensory loops. Achieving such a correspondence is what we call coherence, degrees of which are most likely modulated by dopamine.

Sentences do not map the world "out there," independently of us; they map sensory/motor processing, and are *about* the anticipation of experience.

(b) Reasoning. Accordingly, change-in-view proceeds as a reassembling of representational schemas—a reassembling that settles into a stable state when emotional resonance stabilizes reentrant activation patterns. What we call "reasoning" is *not* the discovery of what beliefs justify or *make true* a given assertion, but an ongoing symbolic portrait of the current body-state engaged in measuring the outcome of a hypothetical action.

(c) Metaphysics. The implications for metaphysics are extensive. As Llinás persuasively contends: "Although the brain may use the senses to take in the richness of the world, it is not limited to those senses; it is capable of doing what it does without any sensory input whatsoever. The nature of the brain and what it does makes the nervous system a very different type of entity from the rest of the universe. It is ... a reality emulator. Suggesting that the system is closed, and so very different, means that it must be another way of expressing 'everything.' In other words, *brain activity is a metaphor for everything else*."[10]

(d) Historical Note. I should like to close with an historical note. Hobbes and Spinoza predicted all this in their own early efforts to purge philosophical psychology of its traditional teleological baggage, noting that the train of thought isn't pulled by its objects but pushed by a succession of body-states, and that the behavioral potential of the intellect is as constrained by natural law as any other star-stuff. Spinoza had the matter exactly right: "*The object of the idea constituting the human mind is the body*," (E2, Prop.13) or, conversely, as Llinás says, "brain activity is a metaphor for everything else."

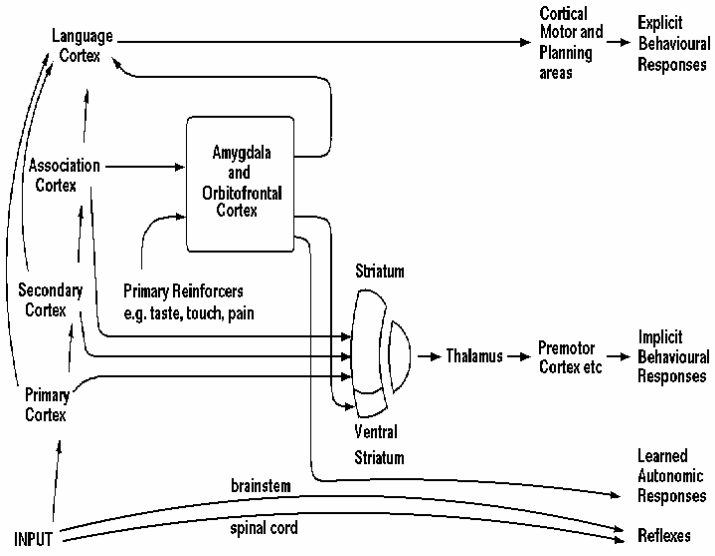

Figure 4. Summary of the organization of some of the brain mechanisms underlying emotion, showing dual routes to the initiation of action in response to rewarding and punishing, that is emotion-producing, stimuli. The inputs from different sensory systems to brain structures such as the orbitofrontal cortex and amygdala allow these brain structures to evaluate the reward- or punishment-related value of incoming stimuli, or of remembered stimuli. From Edmund T. Rolls, *The Brain and Emotion*, Oxford University Press, 1998. Fig. 9.4.

Notes

[1] Hacking, Ian. *The Social Construction of What?* Cambridge, MA: Harvard University Press, 1999, p. 21.

[2] *Metaphysics* 1011b, 25-28; cp. Plato, *Sophist* 263d.

[3] Moore, G.E. *Some Main Problems of Philosophy.* London: George Allen & Unwin, 1953, p. 247.

[4] Quine, Willard Van Orman. *From a Logical Point of View.* 2nd Edition. New York: Harper & Row, 1963, p. 43.

[5] Dewey, John. *Reconstruction in Philosophy.* New York, 1920, Chapter 6.

[6] Dennett, Daniel. *Kinds of Minds.* New York: Persus Press, 1996, p. 88.

[7] Calvin, William. *The Cerebral Code: Thinking a Thought in the Mosaics of the Mind.* New York: 1999, p. 92.

[8] Llinás, Rodolfo R. *I of the Vortex: From Neurons to Self.* Cambridge, MA: MIT Press, 2002, p. 8.

[9] Damasio, Antonio. *Descartes' Error.* New York: Harper & Row, 1999.

[10] Llinás, Rodolfo R. *I of the Vortex: From Neurons to Self.* Cambridge, MA: MIT Press, 2002, p. 94.