PRINCIPIA **30**(1): 7–15 (2026) doi: 10.5007/1808-1711.2026.e110910
Published by NEL — Epistemology and Logic Research Group, Federal University of Santa Catarina (UFSC), Brazil.

INTRODUCTION TO ISSUE OF *PRINCIPIA* ON NEUROPHILOSOPHY

PATRICIA SMITH CHURCHLAND
University of California San Diego, USA
`pschurchland@ucsd.edu`
https://orcid.org/0000-0002-7330-7287

Abstract. This is an introduction to the special issue dedicated to the 40th anniversary of *Neurophilosophy: Toward a unified science of the mind/brain*.

Keywords: neurophilosophy • mental processes • brain processes • cognitive sciences • neuroethics

---

RECEIVED: 14/01/2026 ACCEPTED: 30/01/2026

Neurophilosophy was published in 1986, almost forty years ago. During those forty years, neuroscience has developed in revolutionary ways. Although important progress was made in the 1950's and beyond regarding how an *individual* neuron in V1, for example, responds under various conditions, the holy grail was this: how do neuronal *networks* work? In more detail, how does the interaction among neurons in networks yield a complex output, such as recognizing the face of Obama or deciding to reach for a biscuit? To address network mysteries, a great deal of inspired engineering as well as clever miniaturization in electronic devices had first to be achieved. Only within the last fifteen years has it become possible to record from thousands of neurons with multi-electrode arrays or neuropixel probes, and to use calcium imaging with optogenetics to manipulate neuronal activity using light. (Carter et al 2022.)

These technological developments encouraged research into the multi neuronal activity during feeling fear, learning, decision-making, and motor control. The results in many respects were unpredicted. In many cases, the results provoked re-evaluation of long-standing hypotheses. One surprising result was that motor-decision signals are found very broadly all across cortex, not just in the hitherto identified "motor-decision" areas (Musall et al 2019; Findling, C. et al 2025). This raises the questions concerning whether in general cortical areas are due for a reconsideration of functions. It is questionable whether the functions of certain areas are best named by labels we are currently familiar with in cognitive science, such as "logic module", or "decision module". The customary labels may be more inappropriate than even close to being accurate to the brain.

Technical developments have also enabled the emergence of quite new methods for treating patients who lost the capacity for speech as a result of ALS (amyotrophic lateral sclerosis) or stroke. By strategic placement of semi-implanted microelectrode arrays, a mute subject can control a computer with their silent speech (thoughts) to signal vocal speech (Wairagkar, Card et al 2025). At first, not all speech nuances could be captured using these methods, but recent technical improvements have come closer to fixing that. Die-hard dualists must find it increasingly challenging to specify some role for the nonmaterial soul given these results.

Neurophysiological breakthroughs have been matched with neuroanatomical progress using new techniques with the hope of revealing computationally relevant structure. For example, Martin Sereno and colleagues found that even so-called "higher cortical areas" such as prefrontal cortex as well as "higher visual areas" were topographically organized. This result ran against the long unchallenged assumption that apart from early sensory areas such as V1 and S1, higher areas were unmapped, since they were engaged in map-independent "association" functions. That turned out to be just wrong, and no one knew what "association" was anyhow. Maps now appear to be a fundamental organizing principle, not only in cortex but also in the cerebellum, and even in noncortical areas such as the caudate and the putamen which have mapped projections to the substantia nigra. What such mapping, and other "odd" mappings in basal ganglia imply regarding function is not understood.

Another important anatomical discovery made by Sereno and colleagues is this: in cortex, local patch structures are found all over cortex, and most of any given neuron's connectivity lies within about one millimeter radius of the patch center. (Sereno et al 2022) This gives us a lot to ponder as we try to understand what exactly networks and their subnetworks are doing. A related unanswered question is this: how does the brain integrate so much information in so brief a time and how are motor outputs organized in milliseconds? In artificial neural networks signals travel vastly faster than neural signals — on the order of nanoseconds — whereas brain circuitry works on millisecond time scales.

Although the technological developments in neuroscience are wonderfully exciting, we are apt to ignore the sobering truth that we still understand very little about how those data help us explain how nervous systems actually work. Yes, the data are magnificent and then some. But what new *explanations* are based on that data? What have we learned about the causal mechanisms whereby a monkey can swing through the trees? Or mechanisms whereby an octopus solves a spatial problem to get food? As it is sometimes ruefully admitted, neuroscience is data rich and theory poor (Sejnowski, Churchland, & Movshon 2014). Neuroscientists are not unaware of the lopsided state of the data/theory duet (Rust 2025). It is not that there is no theorizing, but only that currently, at least some of the big mysteries still need to be cracked, easier said than done to put it mildly. Increasingly, it is evident that this requires collaboration across many labs and across many disciplines. Such collaboration may not be as straightforward as it sounds. With in lab, a culture of easy exchange, common vocabularly and shared assumptions quietly build up. In cross-lab exchanges, there needs to be culture sharing, and that takes time and good-natured camaraderie. Fortunately, it is happening in many labs (see International Brain Laboratory et al 2025)

One momentous and semi-philosophical observation is if that we want to understand the brain, conceptual innovation is essential to bridge the gap between the activity of individual neurons and the jobs performed by the networks in which they participate. For example, no simple causal line goes from a single neuron in motor cortex to reaching for an apple, a relatively complex action. By contrast, in spinal reflexes such as a stretch (knee jerk) reflex, motor cortex plays no role, and a relatively simple causal line well explains the behavior. As we all learned in basic science, a smart tap below the kneecap stretches the tendon, which is detected by a sensory neuron that synapses onto a short interneuron that connects to a spinal motor neuron that makes the quadricep muscle contract. (see familiar diagram below) Causality in the knee jerk looks pretty straightforward, though you can add a bit of additional complexity if you want to get meticulous.

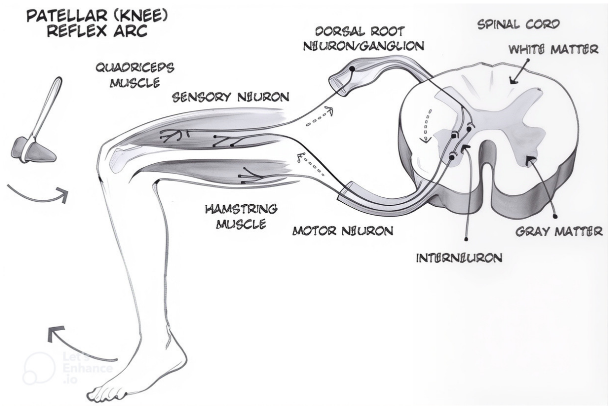

Figure 1: Knee Jerk Reflex. This simple wiring diagram shows the circuitry whereby tapping just below the knee activates the tendon, and a sensory neuron picks up the signal which then travels to the spinal cord. In the cord area, the sensory neuron synapses on an interneuron which generates a signal in response. In turn, that signal is then passed onto the motor neuron dendrite synapse which is also in the cord. The axon of the motor neuron then exits the spinal cord, and connects via a synapse on a region of muscle, causing the muscle to briefly contract.

By contrast, for actions such as reaching or throwing, simple wiring diagrams have long been known to be explanatorily worse than useless. The capacity for behavior of mammals, for example, is very rich, appropriate to changing conditions and even quite novel conditions. Often motor behavior is highly skilled (I think of ice hockey, or surfing). And success may depend on others in the group, as in wolves and humans. Even the ostensibly humdrum action of moving a hand exactly right to grasp the apple from a tree branch involves complex mechanisms intervening between visual recognition of an apple and reaching for it, all the while maintaining whole body balance. We can move the arm at different speeds, we can reach overhead or to one side or move the whole body forward along with the arm, and so on. Sometimes a person will catch a football while running and reaching. Our ease and nonchalance in moving as we do mask the very difficult problems evolution had to solve in order to move and survive. How do we manage to move in so many ways? What are neurons in motor cortex doing? There are many brain areas with a role in movement, but let's start with neurons in cortex that project to spinal cord to allow us to catch that ball.

Motor cortex has long been studied. Despite thoughtful experimental design, uncertainty about what the neurons are representing or whether they are representing anything at all, continued to linger. Not unreasonably, neurons in motor cortex were generally expected to be tuned to something external, such as a "preferred movement", just as sensory neurons had a fairly stable stimulus that would drive them – upgoing light, for example. The fly in the ointment is this: recording from motor cortex neuron activity does not reveal any such movement preferences, such as raising an arm or scratching your nose, ingenious though experimenters were. The unavoidable but unwelcome conclusion is that neurons in motor cortex are just not tuned to external parameters. They are not even tuned to specific profiles of muscle activity. So what *are* motor neurons in cortex doing? And how can we find out?

One obvious but easily ignored feature of the motor system is that it is for doing things — it is dynamical. Crudely, things necessarily change over time as you move and because you move. So you have to change appropriately what you do next which then changes what happens and that changes what you will do after that. More succinctly, in a dynamical system, "the state determines the derivative". We casually and comfortably live with that in our day-to-day lives. But nature had to be very inventive to make it so.

At the *population* level, the motor cortex can be seen to change its activity profile during three phases: "*preparation*", "*execution*", (the main focus here), and then "*during the movement*". If we expected the responses of each and every neuron in a network to be seen to do much the same things during, say, preparation, we would be flatly disappointed. What is actually seen is that the activity profiles of neurons one-by-one vary greatly within *every* phase, as though each is doing its own thing. Even more puzzling, single neurons do not shift from preparation phase to execution phase in an obvious pattern-preserving way. Once a trigger signal is received the cortical motor neurons activity looks very different from was it was doing during preparation. So what was the point of the preparation?

In the early 2000's, Mark Churchland the Krishna Shenoy figured that maybe they could find the answer at the population/network level, and then work back down to the single neuronal level, instead of the other — traditional — way around. Is there a population level organization, though invisible at the neuronal level, that plays a causal role in a well-organized reach of the arm? Will it inform us about what the individual neurons in motor cortex are up to?

Churchland and Shenoy realized that the concept of a *factor*, familiar in statistics, could be deployed here as an abstraction from individual neuronal events in motor cortex and give them an entry to the network level. Think of the strategy this way: we were familiar with pressure as a phenomenon before we knew that pressure is a statistical abstraction, not a property of individual particles. They proposed that in neuroscience a factor is a population level feature, extracted from the otherwise puzzling responses of individual motor neurons during the preparation and execution phases involved in a reaching movement. Here is how Mark Churchland (2025) explains what, in the context of neuroscience, factors are:

> "Think of it this way: When you record from 100 neurons simultaneously, each firing at different rates over time, you have 100-dimensional activity. A "factor" is a pattern or dimension that captures systematic co-variation across those neurons. It's extracted through dimensionality reduction techniques (like PCA, factor analysis, or more sophisticated methods)."

Thus conceived, a *factor* is a lower dimensional-description of a network rather than a neuron-by-neuron description. The data also showed that the preparation factors were rather different from the execution factors. Consequently, the relation between them was worth discovering to see whether explanations might be forthcoming. One deeply puzzling result was that preparation level factors are, as they put it, "output-null", in the sense that they did not appear to have the predicted impact on execution of the action. Specifically, output-potent factors seen during the execution phase were very dissimilar to factors in the preparation phase. At this point, one might conclude, well then you have the wrong "higher order" abstraction. Or you could say that preparation factors are just the brain's version of the cactus spine, with no real function. Yet Nature rarely puts that much effort and energy into just horsing around. So perhaps output-null factors are doing an important job if you look more closely. At this point, I find myself recalling the lesson of so-called "junk DNA". As geneticists looked more broadly and carefully, "junk DNA" turns out not to be useless at all but generally has complex functions, albeit *noncoding* functions.

From my perspective, what is exciting about output null factors in motor cortex is that they do affect execution factors that affect muscles, so they are not causally irrelevant. To see the causal links, however, you need the right kind of conceptual framework that uses the state spaces, vectors representing movements, and the relations between state spaces. Crudely, this mechanism just falls out of evolution's way of arranging neurons to change in time to handle complex space-time movements such as we see in monkeys and humans, though not only monkeys and humans.[1]

The bottom line is that preparation factors appear to be essential to the profile of execution factors and hence indirectly affect neurons in the spinal cord. For starters, output-null factors affect the choice of speed of the movement, such as the apple reach. They seem also to have a role in triggering the movement — in shifting to execution factors. Owing to output null factors, we can make several movements at the same time without getting tangled up, such as pushing away a branch with the upper arm while reaching for the apple with the lower arm, while not falling off the ladder, all the while singing "Hello Dolly". Sometimes, of course, we do get tangled up, especially when we try something very new such as riding a bicycle or riding a horse with a bridle but no saddle. Perhaps learning through trial and error tunes up our output-null factors. In short, output-null factors seem to be one of evolution's solutions to the difficult problem of allowing a body to make a vast range of movements despite using the same limbs and muscles. Unlike plants, multicellular animals have a nervous system rigged to handle movement, and movement means the system has to handle constant change. Even as I sit here writing this, my head is moving, my hands are moving, my whole body shifts a bit now and again. Maybe output-null factors are really output-subtle factors.

As Churchland and Shenoy show, the sort of conceptual strategy they deployed in using network-level factors helps explain properties of motor neuron activity that hitherto were deeply puzzling. This strategy does mean we have to think about the motor system rather differently than hitherto, and maybe that means we shall end up thinking about many properties of a nervous systems -- even cognition -- a bit differently.

Within the domain of neuroscience, this approach to winnowing out population-level concepts that can carry a significant explanatory load seems an example of the conceptual innovation I earlier mentioned as needed in neuroscience.

Science is replete with conceptual innovations that strike thoughtful people as nutty, but that over time became so entrenched that with little to no fanfare, they graduate from "nutty" to "intuitively obvious", often becoming ever more complex as they help us uncover yet more facts. In biology, the nonintuitive conceptual innovation list would include "gene", first introduced in 1909 by Danish botanist, Wilhelm Johannson, who explicitly saw the word as an abstraction. There was significant "be sensible" pushback, and even by 1926, Johannson was not sure whether the concept referred to something real or whether it was just a tool. Hedging his bets, he said that genes were still "entities of calculation, expressions of realities of unknown nature, but with familiar effects". By the 1940's, however, many biologists acknowledged that genes are indeed real and are located in specific places on the chromosome.

Later, it was found that that genes were made of DNA though that too was rejected for about eight years on grounds that proteins were obviously the hereditary material. Finally in 1952, Hershey and Chase published their infamous "Waring blender experiment". They definitively showed that yes, genes *are* made of DNA. Even after Watson and Crick reported on how DNA codes information, complexities continued to be unearthed. For example, some segments of DNA produce (via RNA, needless to say) a specific protein depending on whether the segment is processed in fetal development or in maturity, or in other cases, depending on tissue type. Then there are segments that appear to do regulatory jobs rather than code for proteins — are they really genes? This is a familiar example where a conceptual innovation initially seemed counterintuitive but now has settled comfortably into the realm of real things and has provided the mechanistic basis for new discoveries. Additionally, especially in biology it may be wise not to insist that concepts either be precisely defined or not used at all. Core scientific concepts may well have fuzzy boundaries, just as many everyday concepts such as "house" or "mature" do. (E. Chen 2022)

Aside: Leslie Orgel, a British biochemist at the Salk Institute, regularly cautioned against expecting matters in biology to be simple. *Orgel's Second Law* says this: evolution is cleverer than you are, even when you take Orgel's Second Law into account. For completeness, here is *Orgel's First law*: "Whenever a *spontaneous process* (i.e. one not externally caused) is too slow or too inefficient, a protein will evolve to speed it up or make it more efficient." His observation helps us understand why there are so many proteins controlling biological processes.

Finally, it is perhaps worth recalling that in physics the conception of *force field* was initially rejected as nonsense many physicists. Indeed, the revered Isaac Newton famously called the idea of causal interaction without contact "absurd". He wrote in a letter to Richard Bentley in 1692, "That gravity should be innate inherent & essential to matter so that one body may act upon another at a distance through a vacuum without the mediation of any thing else by & through which their action or force may be conveyed from one to another is to me so great an absurdity that I believe no man who has in philosophical matters any competent faculty of thinking can ever fall into it." [2] If you put yourself in Newton's 17th century shoes, his resistance to the idea of "force field" seems not unreasonable. Surprisingly perhaps, to his everlasting credit Newton not only changed his mind, but figured out what now we all learn as Newton's Universal Law of Gravitation.

It is important that a significant number of philosophers have begun to see the value of empirical discoveries, whether in neuroscience or genetics, or in anthropology or ethological studies of nonhuman organisms such as the octopus. A splendid contribution to such research is found in the twelve papers of this issue of Principia: an international journal of epistemology. I am immensely pleased to see the intense and careful work each author has put into his or her selected topic, and the thoughtful reservations regarding various hypotheses show the payoff of multidisciplinary research.

References

Carter, M.; Essner, R.; Goldstein N.; Iyer, M. 2022. *Guide to Research Techniques in Neuroscience*. Third Edition. New York: Academic Press.

Chen, E. K. 2022. Fundamental nomic vagueness. *Philosophical Review* 131(1): 1–49.

Churchland, M. M. 2025. *Understanding Large-Scale Neural Computations*. Video. Kempner Institute at Harvard University. https://www.youtube.com/watch?v=aFkwFTI8RnQ. Access 14.01.2026

Churchland, M. M.; Cunningham, J. P.; Kaufman, M. T.; et al.. 2012. Neural population dynamics during reaching. *Nature* 487(7405): 51–6.

Churchland, M. M.; Shenoy, K. V. 2024. Preparatory Activity and the Expansive Null-Space. *Nature Reviews Neuroscience* 25(4): 213–236.

Findling, C. et al. 2025. Brain-wide representations of prior information in mouse decision-making. *Nature* 645: 192–200.

International Brain Laboratory; Angelaki, D. et al. 2025. A brain-wide map of neural activity during complex behavior. *Nature* 645(8079): 177–192.

Musall, S.; Kaufman, M. T.; Juavinett, A. L.; Gluf, S.; Churchland, A. K. 2019. Single-trial neural dynamics are dominated by richly varied movements. *Nature Neuroscience* 22: 1677–1686.

Perkins, S. M.; Amematsro, E. A.; Cunningham J.; Wang, Q.; Churchland, M. M. 2025. An emerging view of neural geometry in motor cortex supports high-performance decoding. *eLife*. https://doi.org/10.7554/eLife.89421.3

Rust, N. C. 2025. The Unexpected value of communicating science to the public. *Nature Reviews Neuroscience* 26(10): 572-573.

Sejnowski, T. J.; Churchland, P.S.; Movshon, J. A. 2014. Putting big data to good use. *Nature Neuroscience* 17(11): 1440–11441.

Sereno, M.I.; Sood, M. R.; Huang, R. -S. 2022. Topological maps and brain computations from low to high. *Frontiers in System Neuroscience* 16: 787737

Sussillo, D.; Churchland, M. M.; Kaufman, M. T.; et al. 2015. A neural network that finds a naturalistic solution for the production of muscle activity. *Nature Neuroscience*. 18(7): 1025–33.

Wairagkar, M.; Card, N. S.; Singer-Clark, T.; Hou, X.; Iacobacci, C.; Miller, L. M.; Hochberg, L. R.; Brandman, D. M.; Stavisky, S. D. 2025. An instantaneous voice-synthesis neuroprosthesis. *Nature* 644(8075): 145–152.

Notes

[1] I have not here provided their explanation which would require adding a lot of background material. But see the recent review article and helpful figures in Churchland and Shenoy 2024.

[2] The quote comes from Newton's fourth letter to Richard Bentley, dated February 25, 1692/93. The original letter from Isaac Newton to Richard Bentley was published online in 2007 via *The Newton Project*. The online address where the whole letter can be read is: https://www.newtonproject.ox.ac.uk/view/texts/normalized/THEM00258.