The History of Neuroscience in Autobiography Volume 13

Edited by Thomas D. Albright and Larry R. Squire
Published by Society for Neuroscience
ISBN: 978-0-916110-12-3

Patricia Smith Churchland

pp. 194–219

https://doi.org/10.1523/hon.013005

Patricia Smith Churchland

Born:

Oliver, British Columbia, Canada
July 16, 1943

Education:

University of British Columbia, BA (1965)
University of Pittsburgh, MA (1966)
Oxford University, BPhil (1969)

Appointments:

Assistant Professor, Philosophy, University of Manitoba (1969–1972)
Associate Professor, Philosophy, University of Manitoba (1972–1978)
Professor, Philosophy, University of Manitoba (1978–1984)
Professor, Philosophy, University of California San Diego (1984–2011)
Adjunct Professor, Salk Institute, (1994–present)
University of California President's Professor, Philosophy (1999–present)
Chair, Philosophy Department, University of California San Diego, (2000–2007)

Honors and Awards (Selected):

MacArthur Foundation Research Fellow (1991–1996)
Elected, Academy of Humanism (1993)
Elected, American Academy of Arts and Sciences (2003)
Honorary Doctor of Letters, University of Victoria (1996)
Honorary Doctor of Laws, University of Alberta (2007)
Rossi Prize in Neuroscience (2008)
Prose Prize for Science Writing (2013)
Revelle Medal for distinguished service, University of California San Diego (2022)

Pat Churchland recognized that speculative philosophical approaches to questions about knowledge, consciousness, and decision-making could benefit from emerging data in the neurosciences. Her first book Neurophilosophy: Towards a Unified Science of the Mind-Brain *(MIT Press, 1986) made the case for incorporating data from neuroscience, evolutionary biology, and genetics in developing theories about our mental life. Extending those ideas, she collaborated with physicist/neuroscientist Terry Sejnowski to publish* The Computational Brain *(MIT Press, 1992). The book's central idea is that brains have many levels of organization from neurons to circuits to systems, and explanatory hypotheses at many levels can draw on computational principles. As one of the first comprehensive treatments of the subject, this book was widely read and republished in 2017. Her later work focused on social neuroscience especially in mammals and birds, drawing on data showing that social attachment is mediated by neuropeptides oxytocin and vasopressin. Her hypothesis was that social attachment is the neural platform for moral norms that are learned after birth. Her other publications include* BrainTrust: What Neuroscience Tells Us about Morality *(Princeton University Press, 2011),* Touching a Nerve: The Self as Brain *(Norton, 2013), and* Conscience: The Origin of Moral Intuition *(Norton, 2019).*

Patricia Smith Churchland

Early Life

My dad left the farm when he was 12 to begin working as a "printer's devil" for a small-town newspaper, *The Brooks Bulletin*, in Alberta. Further schooling was not available. As a toddler, he had suffered a bout of polio that weakened one leg so that he had a limp and could not run. Consequently, he was ill-suited for the otherwise likely jobs of wrangling cattle or logging or joining the army. Invoking a favored cliché, he would say that *printer's ink was in his veins*. Indeed, he did know the printing trade well, from how to fix the temperamental linotype machine that melted the lead and dimpled out the lead type, to how to write engagingly about erecting a telephone line over the Monashee mountains. I saw his expertise because he opened a village newspaper in British Columbia's Okanagan Valley, well before I was born. After school, I would often stop by the Oliver News office to sniff the heady smell of printer's ink or eavesdrop as the local Mountie listed the lockups after brawls in the beer parlor.

As children do, I began to imitate my father. I thought writing was fun. While some school tasks seemed a bit tedious, it was generally fun to write. Once I reached high school age, I was permitted to use my dad's home typewriter, a 1928 Underwood. It was a hallowed thing. As I now use a computer, his Underwood sits on a hall table, a reminder of nights of key-pounding while the Northern Lights curtained outside.

Although my mother, Katie James, had attended school until eighth grade after which no schooling was available, she trained as a registered nurse in a small far north hospital that had only two doctors. In the 1920s in Canada, a high school diploma was not required to earn the title of registered nurse (RN). Working in a remote community, the medical staff were dedicated, practical, and kindly. They treated for free many of the local Salish people, quite a few of whom had tuberculosis or smallpox. Katie's mother was an orphan who had been shipped at age six from Scotland to work in Ontario as an indentured servant until she reached seventeen. My grandmother could barely read or write when I knew her, although she did know how to keep bees and gather honey. My mother was the first of her family to have the privilege of any schooling at all, meager though it was. Like my father, she was highly practical and taught us that you manage, no matter what turns up. Until a hospital was built in Oliver, she worked as a visiting nurse, helping people live and, sometimes, helping them as they died. Like my father, she read every chance she could and learned voraciously.

The formal education—or lack of same—of my parents is worth mentioning because one thing they excelled at, above all, is logically working things out, checking evidence when available, and discussing all manner of problems and issues from every angle. The dinner table and the cow barn were sites of extended discussions of one matter or another. Lacking "common sense" seemed to be the worst fault one could have.

When I was at home on the farm, I had chores—collecting eggs, cleaning out the hen house, weeding gardens, canning, doing laundry (wringer style), and so forth. Given the workload, for me, going to school was a pleasure—and this was true for most kids in the Okanagan Valley at that time. Being at school was vastly easier than working the farm, and besides, we had the joy of friends. As it happened, our school was exceptionally strong, because the pioneers in the Okanagan Valley keenly understood the value of education despite having little money, and they put their scanty resources into the school. After the end of World War II, the school had a windfall. Highly educated Scots and English men who were restless after their war years or who could not get jobs after the war, emigrated to Canada, and our school board wisely arranged to bring a batch to our school. They picked up teaching skills quickly, and their authoritative and clear but amusing manner of teaching elicited from us a strong desire for their approval. When I went to college, expecting to find that I was poorly trained relative to well-heeled city kids, I discovered quite the opposite. I was way ahead in every subject, including physics and math. Unlike most of them, I had been taught how to work—to work hard, efficiently, and systematically until the job was done right. Goofing off, as I saw some prosperous students do, seemed as stupid as killing a good dog just for fun.

I fell in love with chemistry in 11th grade. Its beauty and explanatory power amazed me. The mathematization of interactions was dazzling. Noticing my eagerness, my teachers warned that females did not become chemists, so I had best think of something else to do. This warning was certainly not mean-spirited but merely reflected the assumptions of their times. I did consider that a career in law might work out instead, although a little sleuthing revealed two obstacles: (1) women were rarely if ever admitted to law school, and (2) I could not possibly begin to pay for law school. I expected I could talk my way around the "woman problem," but the second, well, the money angle really stumped me. The blokes could earn money for tuition during the summer by logging, which was not an option for women. The money issue again loomed as the major problem when I considered medical school. I ended up going to graduate school in philosophy because, as it happened, I was awarded scholarships to do that. I would like to say I followed my greatest passion, but when you are "as poor Job's turkeys" as they intoned in Oliver, your greatest passion has to take the hindmost.

After High School

I studied philosophy as an undergraduate because I was introduced to the ancient Greeks in my freshman year. Socrates and Aristotle stood out as especially fascinating; Plato, in contrast, seemed lacking in common sense, and downright goofy at times. I was under the misapprehension that current philosophers were trying to understand how the mind works, as I took Aristotle to be. It took me a while to fathom that they actually were interested primarily in words, not things. They staged little contests in which flamboyant cleverness in wordplay counted as progress. One faculty member kindly nominated me for a Woodrow Wilson Fellowship to do graduate work in philosophy, and surprisingly, I won. With no money to do anything else, I followed that option with mixed feelings. I started a graduate program at the University of Pittsburgh, known to be a philosophy of science bastion. I did learn a lot there, and mercifully, the fashionable focus on words instead of things was regularly satirized, especially by the brilliantly witty Richie Schuldenfrei from the Bronx. Given how splendid the program was, it may seem odd that I stayed at Pittsburgh for only one year. Embarrassing to say, in that one year, I had rather foolishly stumbled into a complicated personal life. It seemed least ridiculous to just hightail it to another country, so I applied for a British Council Fellowship to go to Oxford. When the news came that the British Council would supply funds to send me to Oxford, I left Pittsburgh and the social complications behind. Needless to say, I had never traveled to Europe, and Oxford philosophy was then considered to be the fountainhead of philosophy. Highly enticing was the prospect of living in England for a while. Once settled in, I soon realized that Oxford philosophy was actually the fountainhead of wordplay and "analyzing meanings of words," whatever that really was. The Schuldenfrei satires rang in my head, and I could not fail to realize that current Oxford philosophy should be taken with more salt than a mere pinch. Oxford being Oxford, they did not throw me out. Instead, Philippa Foot found me a supervisor who was happy to know nothing about my project (the importance of unconscious processing in decision-making) and would be obliged only to have tea with me twice. Justine Gosling, a Plato scholar at Jesus College, was perfect for me, and humored my fascination with science and the brain. I passed.

The one good thing to emerge from my early years in philosophy in British Columbia was that I met Paul Churchland, who not only was tall (nice because I was very tall for a female) but also studied philosophical questions from a knowledge of the history of science and of physics in particular. Outrageous for the time, Paul suggested that the typical common explanations we give for human behavior in terms of intentions, beliefs, and so forth might eventually be revised, perhaps quite radically, just as "element" no longer meant "earth, air, fire, and water." His hunch was that just as medieval assumptions such as that the earth is the center of the universe turned out to be quite wrong, so too might current assumptions about the mind, such as that beliefs required language skills, be wrong. Furthermore, concepts such as" self-control" or even "reasoning" might embody false assumptions comparable the concepts of "vital spirit" or "phlogiston." Encouraged by his Pittsburgh supervisor, Wilfrid Sellars, Paul envisaged a deeper understanding of the brain that might lead us to discover functions hitherto unknown, and that "folk psychology" might be radically modified to reflect neuroscientific discoveries (P. M. Churchland 1989). This was a commonality between us right from the beginning (see Hirstein 2004).

After he was awarded his doctorate from Pittsburgh, Paul got a job at the University of Toronto, rather a feather in his cap at the time. When I graduated from Oxford, the University of Toronto explained to Paul that they could not possibly hire me—I was a woman, and women could not really make a success of philosophy, Oxford degree notwithstanding. Yes, they already had one female philosopher on the faculty, but that was more than enough. As luck would have it, Paul and I then landed two jobs in philosophy at the University of Manitoba through the kindness of an old friend, Jack Bailey, from Pittsburgh.

From Philosophy to Neuroscience

How did the transition to neuroscience come about? First, it was sheer luck that my first job was at a university (Manitoba) that took itself seriously enough for there to be some terrific faculty, but not so seriously as to be arrogant and self-important. The story of my integration with neuroscience in Manitoba began with my exasperation in trying to learn neuroanatomy from books.

Throwing my anatomy textbook at the wall, I was forced to admit that my goal was doomed. What goal? To teach myself the organization of the thalamus, a brain structure deep under the cortex. Why the thalamus? Because in mammals, all sensory inputs (except smells) go through the thalamus and from there to cortex. It is kind of the heart and soul of all mammalian brains. In my expensive textbook, each page showed a slice of the thalamus, advancing from back to front, each two millimeters thick. The trouble was that from one page to the next the thalamus looked completely different to me. From slice to slice, I could not get any sense of continuity, of what this walnut-size thing really looked like. Visualizing a three-dimensional (3D) thalamus from two-dimensional (2D) illustrations was a dead end. Maybe going whole-hog on the real 3D object was the better way. The University of Manitoba had a medical school, and the medical school had a department of brain anatomy. Surely some neuroanatomist could show me what the thalamus looks like. Thus, my state of mind (P. S. Churchland 2021).

Cancelling my office hours with students (okay, my bad), I drove my rusty VW downtown and found the neuroanatomy department. The sign on the door read: John Baskerville-Hyde, Head. Warmly welcoming my interruption, Baskerville-Hyde listened to my tale of woe with amusement and encouragement. “Yes,” he agreed, “philosophers need to know about the brain if they want to understand the mind. Here is what we can do. Can you attend regular neuroscience lectures with the first-year medical students? For the lab component, the students will get a human brain to dissect. I will arrange for you to have your own human brain to dissect. Then you’ll begin to learn brain anatomy. In lectures, you will also learn physiology.”

Just like that, after a mere 40 minutes’ conversation, my intellectual life changed forever.

A human brain was indeed delivered to me in the anatomy lab, and holding it my hands, I felt an almost reverential humility toward this tissue that had embodied someone’s love and knowledge and skills. It looked so small relative to what a human brain can do.

The world of neuroscience was opening up to me. At the clinicians’ weekly meeting—neurology rounds—a patient with unusual or puzzling symptoms would be presented and later discussed. To my everlasting gratitude, the clinicians invited me to join their rounds. One stroke patient was a dairy farmer whose only symptoms were that he could no longer recognize faces—not those of his wife or children, or even his own face in a mirror. Particularly disappointing to him was his inability to recognize the faces of his beloved cows. Encouraged to wander the neurology ward of the hospital, I visited a stroke patient with hemineglect. She was unaware of any stimulus or event in the left-side of her space and paid no attention to the left side of her body. When I raised her left hand and brought it into her right hemispace, she said did not know who it belonged to. Another stroke patient suffered from pseudobulbar affect (brainstem lesion is the cause), which meant that when stressed, he would suddenly sob piteously for about 30 seconds or so. When the sobbing ceased, he explained that he had felt no sadness whatever, but the pseudobulbar affect was certainly a nuisance. Because Manitoba is a province spanning a vast geographic area, and because only Winnipeg had advanced medical facilities, any patient with challenging symptoms eventually came to the medical school hospital. I was fortunate to see cases that many neurologists see only rarely, if ever.

Although I prized the connection to the clinicians, it was in Larry Jordan’s local neurobiology lab that I really began to understand the challenges and the thrill of getting data from nervous systems. Larry’s lab was focused on circuitry for rhythmic behaviors—more exactly, on walking in the spinalized cat. At that time, rhythmic behaviors were usually studied in simple nonmamalian animals such as the lobster (its stomatogastric ganglion rhythmically ground up food). Rhythmic behaviors were targeted because they were an entry to the circuit level. Although no individual neuron might be producing rhythmic output, it was intriguing that the *circuit* generates rhythm and modifications to a basic rhythm. What drew researchers in was the prospect of understanding a circuit output in terms of the input signals, along with the physiology and the connectivity of all the neurons comprising the circuit.

Larry Jordan's lab examined walking as controlled by the spinal cord. The experimental animal was the cat. In that lab, I came to see the wisdom of studying the rhythmic behavior as a circuit level phenomenon, not the selective business of a single rhythmic neuron or single pulsating neuron type. To isolate the circuitry in the spinal cord, the cerebrum of the animal was surgically removed under anesthesia, and a sling was used to keep the cat upright. Walking in the decerebrate cat could be sped up or slowed down by changing the speed of the treadmill its paws touched or by altering chemical balances in the spinal cord. (Jordan and Sławińska 2011) What exactly was the circuitry and how did it work, were the lab's defining questions. Paul got very involved in the lab's work as well, and we egged each other on as we basked in the joys of this hands-on direction. We jointly wrote a paper critical of conventional "truths" about qualia and visual experience (Churchland and Churchland 1981), sometimes drawing on data, sometimes just exposing cock-eyed logic relied upon by philosophers.

Untenured philosophy professors, as I was at the time, are typically advised to avoid risk and stay well-focused on the job they were hired to do. Providentially, my dean, Fred Stambrook, an historian with a dash of intellectual derring-do, skillfully encouraged my hands-on initiative, even while members of the Philosophy Department were faintly bewildered. In time, papers in philosophy journals were published, tenure was granted. Quietly, I thanked my lucky stars that I was not at Toronto where my unconventional shenanigans would have been seriously frowned upon.

From me, the neurologists wanted to know whether philosophers thought there was a nonphysical soul, and if so, why. For example, we discussed the split-brain results in detail. They were puzzled by those who applied the computer metaphor to the brain function, which claimed that the brain is merely the hardware on which to run the software of cognition. According to a popular philosophical assumption favored at MIT, marketed vigorously by Noam Chomsky and Jerry Fodor, was that to explain the human mind, only the software needs to be understood. Leave the hardware—the brain—out of it. Brain-phobic philosophers found that assumption reassuring and argued rather creatively in defense of it. The neurologists, however, were unmoved, compelling MIT rhetoric notwithstanding. Much later, undermined by accumulating neurobiological evidence, the misbegotten software-hardware metaphor was quietly shelved, or mostly anyhow.

To great philosophical acclaim, Jerry Fodor wrote a book, *The Language of Thought* (1975), supposedly showing that a language could be learned *only if* the child already had an innate language of thought; language learning is, according to Fodor, nothing but translation. Motivated by evolutionary biology and some knowledge of the brain, I published a paper (P. S. Churchland 1978) in the philosophy journal, *Synthese*, emphasizing the illogical underpinning of Fodor's view, bemoaning the lack of data, and more generally, poking fun at the Fodor method that (only in conversation) I had rather ungraciously called "pulling science out of your ass."

Inevitably I was drawn in to the topic of consciousness, especially because I found myself in opposition to the conviction, common among philosophers and some scientists, that only humans are conscious. More exactly, the idea was that *only if you have a language* can you be conscious, and of course, only we humans have language. This conviction goes back at least to Descartes and his 17th-century introspection, but perhaps even further. Still, contemporary philosophers of the "analyze-word-meanings" approach to understanding consciousness made much of an alleged "conceptual connection" between using language and being aware. "Consciousness" means, so they claimed, the ability to *say* what you are conscious of. Daniel Dennett, argued extensively for this view. For example, Dennett (1979) says that what we typically *mean* by "consciousness" requires that a conscious organism must have verbal skills, but a "what a dog has, or the right hemisphere has, is a radically different phenomenon." I saw this as mistaking a test for consciousness (reporting verbally) with the phenomenon itself.

In my paper, "Consciousness: The Transmutation of a Concept" (1983), I marshaled evidence from many scientific sources (e.g., psychophysics, neurology, ethology, evolutionary biology) to demonstrate how feeble were Dennett's claims. For example, Gordon Gallup had shown that chimpanzees, if given a mirror, will study themselves in the mirror, picking their teeth, and checking their behinds. If, under anesthesia, a red mark was put above their eyebrows, later they would study that mark in the mirror, gently check it with their finger and then check the finger. Lacking self-consciousness because they have no language? Surely, you jest. Undeterred by negative evidence, Dennett adheres to his speculation that acquiring language skills is necessary for consciousness. His idea is that acquiring language skills, as opposed to hiding or hunting skills, *uniquely* reorganize the brain. The result of such special reorganization is consciousness (see especially Ch. 14 in Dennett, 2017). In the absence of positive evidence for Dennett's ideas here, I have found it more fruitful to look to the neurobiology of attention (Graziano, Webb, 2017) and cortical self-control of movement (Schroeder et al 2022), for example.

By the time I was dissecting human brains, I jokingly thought of myself as a *neurophilosopher*. The name stuck, but the risibility retreated as I began to think hard about what neuroscience teaches us about "big-question" philosophy—questions about knowledge and morality and sleep. It seemed ever more evident that I am as I am because my *brain*—with its genes and neurons and experiences—is as *it* is. *Neurophilosophy* as a new paradigm took hold in me because it became gloriously obvious that progress on various time-hallowed philosophical questions, such as how we learn and see and think, was beginning to be made in neuroscience, and more progress was surely in store. The way forward was unmistakable: follow the science, wallow in the science, think without fear, and ditch bad ideas without shame when you see they are duds. Still, predicting how neuroscience will proceed in the distant future is risky, to put it mildly. As our son Mark asked with six-year-old naivete, "What if the human brain is more complicated than it is smart?" Well, we may never know for sure. We can just keep working to make new tools and new discoveries until we hit the wall—but how would we know it was the wall or just a local minimum?

I should mention here that I immensely enjoyed having children, and I was particularly fortunate that giving birth and lactating were very easy for my body. Paul, too, hugely enjoyed the children, and we both found ourselves gleefully reverting to our childhood mindset in playing with our own two children, Mark and Anne. That they both went on as adults to study neuroscience was a delightful surprise, a choice shaped mostly, I think, by the fact neuroscience is now the locus of so much scientific excitement, and there is so much to be discovered, and not because we had especially prodded them in any particular direction.

We spent a year (1978–1979) at the Institute for Advanced Study in Princeton, which gave me time not only to read and think but also to visit with evolutionary biologists such as Donald Griffin, whose data on animal hunting and foraging behavior strongly indicates that they think and reason and solve problems. I also visited Charlie Gross, whose lab had discovered neurons that respond only to faces and went on to explore cognitive representations more generally. Michael Gazzaniga was at New York University, so I hopped on a train to pay a visit to him and Joe LeDoux, mostly to discuss consciousness, and especially Don Griffin's work, showing the high probability that nonlinguistic animals can think and are conscious.

By the early 1980s, I toyed with the possibility of writing a book to illustrate the interconnections between neuroscience, psychology, and philosophy. Before long, the idea did assemble itself into an actual book, *Neurophilosophy: Towards a Unified Science of the Mind-Brain*, with lots of financial support from the Canada Council. The writing was not completed in Manitoba, but rather in San Diego where we had been offered jobs. Landing jobs at the University of California–San Diego (UCSD) in 1984 was an incomparable piece of luck for me and for Paul. I quickly came to appreciate this when I discovered I could walk from my office to the labs of Larry Squire and Stuart Zola, as well as to Ted Bullock's lab at Scripps down the hill by the beach. There were, however, some uneven moments. Knocking on my office door as I was unpacking books, a rather prominent member of the UCSD Philosophy Department felt the need to explain in painful detail that he had been totally opposed to hiring women, whatever their status. Moreover, he especially opposed hiring me because he though me "mouthy." I found myself laughing resoundingly, inadvertently confirming his "mouthy" description. I offered him a glass of sherry to share the mirth, a glass he politely refused.

In 1986, *Neurophilosophy* was published by MIT Press. By and large, established philosophers loathed or ignored the book. Graduate students, in contrast, snuck off to read it on their own and figured out how to collaborate with neuroscientists to do philosophy in a new way. Beth Buffalo—then a graduate student in philosophy at UCSD and now a highly distinguished neuroscientist, was an early case in point. Her work in Larry Squire's lab on the hippocampus not only showed philosophers that neuroscience was relevant to understanding memory, but also showed neuroscientists that philosophers could be creative and skilled scientists.

Mercifully, lots of scientists, including Francis Crick and Jonas Salk, thought *Neurophilosophy* was a book whose time had come. Their support encouraged me to take the philosophers' ridicule in my stride. They also arranged for me to become an adjunct professor at the Salk Institute, and I have always felt completely at home there. In the period following the publication of *Neurophilosophy*, I was gratified to find that science journals were ready to publish my ideas, including "From Descartes to Neural Networks" in *Scientific American* (1989) and "A Neurobiological Slant on Consciousness Research" in *Progress in Brain Research* (2005). I was also invited to give talks at many neuroscience departments who were keen to hear about why most philosophers thought that studying the brain would tell us nothing about our mental life. These visits also gave me wonderful opportunities to learn about what experiments neuroscientists were doing and why.

One of the hypotheses that Paul and I jointly developed, and which I discussed extensively in *Neurophilosophy*, was that brains grow high-dimensional parameter spaces (maps) and that representations, such as the layout of a house, or of things in the house, could be understood as vectors in a parameter space. On this idea, internal distances in the relationships in the maps might correspond to similarity relationships in the categories of things, such as the similarities we see between kinds of bears or between kinds of evergreen trees or kinds of footwear, and so forth. We expected that there could be integration across parameter spaces, as neurons made connections with other neurons, although how integration was achieved we had no idea. This hypothesis was sketchy, but what we wondered was whether something roughly like this, consisting of neural networks, might be a better approach to cognition than the analogy to digital computers and programs favored by Chomsky, Fodor, and Dennett (P. S. Churchland and P. M. Churchland 2002). In 2007, we were interviewed by a *New Yorker* journalist (Larissa MacFarquhar), who published an article about Paul and me titled "Two Heads: A Marriage Devoted to the Mind-Body Problem."

A physicist/neuroscientist I had met at a meeting at John Hopkins, Terry Sejnowski, was persuaded to take a job at the Salk Institute in 1989, and we shared the hunch that understanding how artificial neural networks learn and represent might yield clues useful in understanding computational principles of real nervous systems. I became an informal but regular member of Terry's computational neuroscience lab. Francis Crick, also at the Salk, adapted the Cambridge institution of afternoon tea every day at 4 p.m. in Terry's lab. We labbies (including in those first days Read Montague, Peter Dayan, Steve Quartz, and Alex Pouget) took turns making tea and rustling up biscuits. Scientists from various other labs, such as Chuck Stevens, Leslie Orgel, and Tom Albright, often dropped in for tea, too. Teatime quickly became the daily occasion for close discussion of ideas and data, flying untried balloons, and giving the broad questions a hearing. It was a time for emerging from the comfortable burrows of safe detail into the wide-open prairie of no-holds-barred. Virtually everyone who visited the lab was coaxed or gently bullied into dilating on the philosophical (grand-scale, background, or fuzzy) questions facing computational neuroscience.

Terry and I decided to write a book, mainly for neuroscientists but also for computer scientists, to explain artificial neural net computing and learning as well as to explore ways in which brain achievements can be seen as computational. In the ensuing publication, *The Computational Brain* (1992), we noted that the style and principles typified by prevailing computer programs—coding, rules, exceptions to rules, and exceptions to the exceptions—were unlikely to reflect computation in nervous systems, given what is known about the basic properties of neurons and the modification of neurons through learning. Although Paul and I had thought a lot about vector-to-vector transformation in high-dimensional spaces, we had not actually built a neural network. But Terry had.

An early Sejnowski project, collaborating with graduate student Charlie Rosenberg, involved constructing an electronic neural network to translate a short piece of written text into speech. (Sejnowski and Rosenberg, 1987). The network was recurrent (loopy), so error signals would be looped back to units in the middle level, and the network would adjust the units' weights accordingly. This error-correcting feedback to the middle layers of the network was mainly a Geoff Hinton invention, known as *back propagation of error* (aka *backprop*). The result of an error correction step was a new output sound to the same input word, generally getting a little closer to the right sound for that text, a process known as gradient descent. Left overnight to run and error-correct according to the algorithm, the artificial neural network (called NetTalk) was surprisingly accurate when tested aloud in the morning. One could easily understand its vocalizations as meaningful text. Probably, it seemed to Terry—and to me—then, that although this small network could not do very much—just map text onto speech—a greatly cranked-up colossal artificial neural network might be able to do vastly more. And that hunch turned out to be correct. In any case, error correction through negative feedback is possibly one prominent way organic brains learn, when they do learn. Error correction via backprop is how NetTalk learned.

A much harder question was whether there were going to be instructive parallels between organic brains and artificial networks concerning the mechanisms of error-correction learning. In short, does my brain use backprop? At that time (1980s) little was known about the highly complex details of brain circuits, although at the behavioral level, the importance of positive and negative feedback was well studied. The ventral tegmental area (VTA) was known to regulate reward learning, and the release of dopamine was a crucial element in how animals learned to approach or avoid a stimulus. As intense research on learning continued, it became evident that the VTA, although important, was only one part of a much larger neural system. Other important components were the hypothalamus, regions of the prefrontal cortex and the inferior parietal cortex, the amygdala, and the basal ganglia (see Averbeck and Murray 2020).

From my perspective, the algorithms by which artificial intelligence (AI) networks learn, cool though they are, appear to be much simpler than the computations used by real nervous systems as they learn. For one thing, AI machines do not have motivations or emotions: they do not run or hide or move at all; they lack diurnal cycles and do not need to forage for food or mates; they do not tend offspring. Whether these factors mean that the intelligence in AI machines is different in kind or just in degree from the intelligence of a rat or monkey is a favorite topic of discussion now in Terry's lab. Owing to lessons learned in Larry Jordan's lab as well as from Rodolfo Llinás, I tend to see evolution of nervous systems as ultimately serving the four Fs of motor control—feeding, fleeing, fighting, and reproducing. Fancy sensory systems and intelligence earn their keep only by serving motor control, either in the immediate future or in the longer term. Nevertheless, deep mysteries remain concerning the mechanisms for sensorimotor control.

Consequently, I wonder how much we really understand about intelligence in ourselves or in other mammals and in birds. Yes, pattern recognition as seen in AI machine output is remarkable, but living organisms have to *generate* adaptive movement, not just recognize patterns. In mammals, not only is the spinal cord involved, but so are the cerebellum, the basal ganglia, the motor cortex, and the contribution of sensory signals. Because sophisticated motor control is a circuit-level function and is adaptable as external conditions change, computing using only the length of an animal's limbs, for example, is not nearly enough to approximate a cat's movements. Thus, an AI system that learns to move from mimicking limb movements may itself move like a stick figure. Sensorimotor learning in a dog's brain as it acquires a ball-catching skill may seem relatively easy to the casual observer, but the computations governing the feat are profoundly difficult to understand. Still, as Mark Churchland has explained to me, "many movements are executed with the same neurons without mixing them up, and one way to accomplish this is to use a different low-dimensional subspace for each task, which is learned. The arm motion you use to serve a tennis ball is different from the one you use to throw a baseball, even though the same muscles are activated. These are still early days, and not all cortex areas have been analyzed. The same dynamics may occur in the prefrontal cortex, where thoughts are dynamical trajectories in low-dimensional concept subspaces." (See also Churchland, M.M., Shenoy, K. 2024.)

The revolution of my lifetime is that both the neurophysiology of real neural circuits and the enlargement of electronic networks have advanced in dramatic fashion, especially beginning about 2000. Experimental use of hundreds of electrodes is now common in studying neuronal circuits, while trying to dope out circuit function using a single electrode now seems quaint if heroic. Thanks to miniaturization in electronics, large language models can be not merely large but gigantic, flaunting upward of a hundred billion parameters, requiring 200 gigabytes to load. (NetTalk, the pioneer, had a measly 300 or so units.) We shall see how much we can learn about brain functions from these devices (Muller, Churchland, and Sejnowski 2024).

Although some philosophers had come to appreciate that empirical data can constrain hypotheses about the nature of the mind, the weight of opinion was that talking to and arguing with other philosophers, and perhaps dwelling on the writings of historical figures such as Immanuel Kant, is the way philosophy works. I once ventured to make a slide quoting a famous comment by economist and engineer, Edwards Deming, "Without data you are just another person with an opinion." Although I regularly showed the slide when I give a talk, the philosophers typically dismiss it as displaying my ignorance of the way philosophy is done and *should* be done.

In particular, positing a nonphysical soul continued to find favor in the philosophical community. Famously, the Australian philosopher, David Chalmers, finds this idea compelling (Chalmers, 1996). His central argument consists of what he dignifies as a *thought-experiment*, which roughly goes as follows: I can imagine a person, like me in every way (attention, short-term memory, use of language, laughs at jokes), but completely lacking in qualia—qualitative experiences, such as feeling short of breath or seeing the colors of a rainbow fade. My brain and the Zombie's brain are, in Chalmers' story, *exactly* the same. In sum, this individual would be exactly like me, save that he would be a Zombie.

"So what if you can imagine such a thing?" you might ask. Here is the conclusion Chalmers draws: because the scenario is imaginable, it is *possible*; because it is *possible*, then whatever consciousness is, it is independent of the brain.

Does Chalmer's conclusion follow? No, not even a little bit. Not even if you are charitable. The glaring flaw lies in relying solely on what *seems* *possible* or imaginable to establish an hypothesis about what *is actual*. After all, what is and is not conceivable is merely *a psychological fact about us*—about what we can and cannot imagine, given our capacity for imagination. It does not constitute factual evidence about the nature of things. I can imagine running faster than the speed of light, but in reality, I cannot. I can conceive of waking up some morning to find that I am a newly hatched chicken. Nothing follows about me, or chickens, except that I have a vivid imagination. (This version of my criticism was drawn from my 2023 article, “Brains and Minds.”)

Additional problems loom: if the imagined Zombie is, as the thought-experiment requires, *exactly* like me, then can it too imagine a world in which there are zombies without consciousness? It is not clear how to make sense of this. Incidentally, notice too that if Chalmers acknowledges that the Zombie has *attentional* capacities but no conscious awareness, he also runs up against the neuroscientific data showing that attention, a neurobiological trait, is typically an aspect of conscious states. So perhaps the fanciful Zombie is not exactly like me, after all. But wait: why not say the Zombie *is* like me and hence has conscious experiences *because* its brain is exactly like mine? In summary, too little thought, and not enough experiment.

I am sometimes cheered up by a casual comment by the Princeton physicist John Wheeler. He quipped that philosophers are like tin cans tied to the back of a car—they make a lot of noise, they do not move the project forward, and they are always behind.

The history of science has a rough parallel to dualism—namely, *vitalism*. Typical of vitalists generally, my high school biology teacher argued thus: no one can explain how living things can emerge from dead molecules. Out of bits of dead proteins, fats, and sugars, how could life itself emerge? He thought it was obvious from the sheer mysteriousness of life that the nature of life could not possibly have an explanation in biology or chemistry. His unwavering intuition about mysteriousness assured him he could just tell that life would require a *nonbiological* solution—that is, *vital spirit*. By 1953, with the discovery of the molecular structure of DNA and how its organization embodied a code for making proteins, the vitalist game folded. Done for.

Both Chalmers’s argument and the vitalist arguments are examples of the well-known *argument from ignorance*. Here is the general form of the fallacy: I do *not* know something (e.g., how the brain produces consciousness), so I *do* know something (e.g., that the brain does not produce consciousness). The fallacy is well named. Another common fallacy (false analogy) is to drum up this syllogism: Consciousness is a mystery, quantum phenomena are a mystery, and so they probably are the same mystery. Roger Penrose, a mathematician at Oxford, claimed that quantum gravity inside a neuron’s microtubules was the cause of consciousness—the same mystery. Although bedecked with various bits of mathematics to provide some respectability, the heart of the argument is obviously a false analogy. (The philosopher Rick Grush and I wrote a paper criticizing Penrose in detail thatappeared in *Pacific Philosophical Quarterly* in 1983.)

To be sure, there is much about the brain that is not understood, including the precise nature of the many mechanisms involved in timing, learning, speech, and stereo vision, or the brain configurations that cause schizophrenia or bipolar disorder, to mention only a few mysteries. Although Chalmers identifies consciousness as a uniquely *hard problem*, from the vantage point of ignorance, it is difficult to know which of the vast number of brain unknowns will more easily submit to progress. This tends to be the way of things in science more generally. For example, the problem of how to predict the folding pattern of a large protein from its amino acid sequence was long thought to be "a grand challenge" in biophysics, never to be automated. From the perspective of 1990, this was a genuinely *hard problem*. Yet by 2021, AI techniques (*Alphafold2*) made that very problem more tractable. When will the problem of consciousness succumb? Next decade? Next century?

How I Came to See Neuroscience as Having an Impact on the Nature of Morality

A longstanding issue discussed by philosophers concerns the origin and nature of moral convictions, and although I did not expect to have anything to say on that matter, everything changed in the 2000s.

Western philosophers have commonly seen morality as a set of rules derived either from a supernatural being, as in the Ten Commandments, or from reason. Or by combining both God and reason, as St. Thomas Aquinas proposed. In Aquinas's view, our reason can discover—and espouse—the correct moral rules because God kindly arranges things thus. Our job is to be in a pious state of mind and deploy our God-given reason. On this approach, the aim of a moral philosopher is to unearth those rules, thus pleasing God and then to advertise the rules broadly so others will follow your lead.

As a child, I had long suspected that the supernatural explanation for the origin of moral norms lacked coherence. The Ten Commandments were ostentatiously implausible as exceptionless rules. I had known of one father of a friend whose conduct implied he deserved no honor whatever. As for keeping the sabbath, that was a luxury farmers did not have.

My callow skepticism got a refinement when, as a freshman in college, I was introduced to the Platonic dialogue, *Euthyphro*. Conversing outside a court of justice are the know-it-all Euthyphro and the faux-featherbrain, Socrates. Euthyphro is at the court to bring a charge of murder against his own father, who killed a mad man, albeit inadvertently. For his part, Socrates attends court to face charges of impiety—really, for asking embarrassing, if polite, questions of the high-handed theologians. As their conversation about gods and morality drifts along, Socrates, feigning innocence, asks the pious Euthyphro: is something right *because* the gods say it is right, or do the gods say it is right because it *is* right? The dilemma for a theistic approach is unavoidable. Either what is right is *arbitrary* (the first horn of the dilemma—merely the gods' say-so) or the gods are irrelevant to the explanation of what makes an action right (the second horn of the dilemma—the gods merely recognize what is right). Either way, appealing to the gods does not illuminate anything about morality. What remains to be explained is what makes an action right and how we learn that. Secondarily, there is the further skeptical challenge: can we be wrong in an ethical judgment even when, like Euthyphro, we are entirely *certain* of our judgment?

As far as I could tell, Socrates's dilemma had hobbled the theological approach, and it also cast doubt on the idea that human reason could find absolute rules that have no exceptions and are immune to counterexamples. Nevertheless, I was disappointed to realize that like Aquinas, many moral philosophers still assumed that reason (perhaps "God-given reason") would reveal the deepest moral rules. Immanuel Kant, for example, held that the moral rules must be absolute and universal, and as an example, cited "always tell the truth." Always? Always. Undergraduates, to the dismay of steadfast Kantians, found plenty of obvious counterexamples—armed Nazis questioning the doorman at Jewish hideouts and so forth—in which telling the truth would be exactly the wrong thing to do.

Jeremy Bentham and James Mill, though atheists, shared with Aquinas the assumption that reason could be relied upon to find the absolute basic rule. Additionally, according to them, they had indeed found the rule by logic and reason alone. Maximizing aggregate utility (acting so as to yield the greatest happiness for the greatest number) is the absolutely fundamental moral rule. Although the counterexample shops turned out oodles of troubling cases, the Bentham-Mill proposal, also known as *utilitarianism*, remains even now appealing to many contemporary moral philosophers (e.g., Peter Singer, Nick Bostrom and so-called "Effective Altruism;" for recent criticism, see L.Wenar, 2024).

A not uncommon philosophical career consists in finding ever-more subtle ways to fend off the inevitable counterexamples to utilitarianism. In graduate school, I ruefully concluded that Utilitarians and Kantians were contentedly engaged in a sterile back and forth, showing again and again how foolish the other side was, but how *their* side alone conforms to reason. I never expected to have anything useful to say on matters concerning morality and vowed to steer clear of morality as a topic. Neurobiology, however, changed all that.

One day in 2007, I went to hear a talk at the Salk Institute on social behavior in voles. The advertised topic was not especially compelling, but I had a couple of hours before teaching my logic class, so I found a seat in the auditorium and settled down to listen. The speaker was Larry Young from Emory University, and he began to describe the behavior of montane voles. Male and female meet, they mate, and they go their separate ways. Ho hum. Then he described the behavior of prairie voles: they meet, they mate, and then they are bonded forever. They form lifelong attachments. The male helps guard and nest and takes care of the babies. Montane voles are loners, whereas prairie voles live in large communities of voles. What, he asked, might be the differences in the brains of montane voles and prairie voles that accounts for these social differences?

Now I am on the edge of my chair. Young's slides reveal his lab's anatomical results. The main neurobiological contrast they found is that prairie voles have a much higher density of receptors for the neuropeptides arginine vasopressin (AVP) and oxytocin (OT) in the ventral pallidum and the nucleus accumbens, respectively, than do montane voles (Lim, Murphy, and Young 2004; Lim and Young 2006). Although all mammals have both OT and AVP centrally, it is the receptor density in specific and highly interconnected regions that marks the crucial difference in social behavior. He went on to describe the supporting evidence for the role of OT and AVP, including for example, the behavioral changes that happen after the receptors are blocked. Inspired to know more, I read everything about OT that I could find on PubMed.

Vole mating behavior is all very interesting, but whence the link to morality? The background that inspired me to think broadly about social behavior given the vole results, came from many years of attending symposia arranged by CARTA (the Center for Academic Research and Training in Anthropogeny). This institution was the brainchild of a diverse group of scientists at UCSD and Salk who in the 1990s formed a casual group called the La Jolla Group for Explaining the Origin of Humans. From different subfields, one and all they were fascinated by questions about human evolution and taught each other what they knew. They understood well that to address the differences between humans and other mammals, a very broad sweep of inquiry was needed. They wanted information from all levels of organization—from the glycobiological and genetic comparisons among mammalian species to comparisons in sensory systems and in behavior, including social behavior of many species studied by field anthropologists. This was cross-disciplinary inquiry at its fermenting best (see the CARTA website at https://carta.anthropogeny.org).

Receiving funding from like-minded donors, they formalized the group into CARTA. I was very happy to be included in CARTA from its beginning in 1998. ("Anthropogeny" is an old term, seldom used now, meaning the attempt to explore and explain the origin of humans.) In this thriving but easy-going milieu, the role of OT and AVP in prairie vole sociality raised questions about the evolution of such behavior, and what those peptides did in nonmammalian species. Drawing on the research of my CARTA colleagues, I saw a way to make progress on long-standing philosophical debates about the nature of morality. The key involved mulling the matter of sociality from an evolutionary perspective. What could be the benefits and costs of sharing food, cooperating, or showing compassion?

Despite the strong claim by Richard Dawkins (1976) that we are all born utterly selfish and must learn altruism by pressure and coercion, field studies of social behavior in many species show that the opposite holds surprisingly often. For example, at a CARTA meeting, primatologist Christophe Boesch presented data showing five distinct cases of a male chimpanzee in the wild adopting an orphan baby and raising it to maturity over an extended period of time—not an easy job. Genetic data (from carefully collected urine samples) revealed that in no case was the male chimpanzee the biological father of the adoptee. Infant adoption has also been documented in bonobos, brown howler monkeys, and Tibetan macaques. On a different note, wolves and elephants will share food and show compassion to conspecifics in the group and sometimes even to allospecifics. Black bears, though long assumed to be cranky loners, actually have rather complicated social relationships with kin across large territories, involving sharing of food and playing. (Getting this data on bears, however, took Ben Kilham some 20 years of careful observation.) The nonhuman data on altruism shows that neither a supernatural being nor language-dependent reason seems to play a crucial role in altruism (De Waal 2009, 2013).

Distinct from the Dawkins model of coerced altruism, a significant philosophical tradition, starring Aristotle, but also the Scottish geniuses Adam Smith and David Hume, sees other-regarding impulses as a deep feature of our nature. On this view, heritable social impulses form the basis for social behavior, and such impulses can be shaped and modified by the environment. Norms emerge from customs, typically making pragmatic sense and reflecting ecological conditions. Dawkins, wrong though he was about the necessity of coercion, provoked us to ask about the evolution of mammals, and in particular, to ask about the evolution of social bonding with offspring and also with others such as mates. Why did evolution select for mammalian and avian styles of sociality (De Waal 2013)?

The succinct version of the story that seems to me probable runs as follows—which I published in book form in *Braintrust* (2011), and in further development in *Touching a Nerve* (2013). The event that triggered the huge cascade of changes in the brain that eventually wired our style of sociality was *homeothermy*. Surprising, but probable, nonetheless. About 200 million years ago, warm-blooded animals appeared for the first time. Even in its rudimentary forms, homeothermy was highly advantageous because warm-blooded animals could forage at night, when their cold-blooded competitors had to slow down and sleep. Additionally, it allowed animals to venture into colder climates. Although homeothermy is beneficial in many respects, it has a cost. Gram for gram, homeotherms need about ten times the calories as poikilotherms. That is a huge constraint, but given the benefits of homeothermy, it favored the evolution of a new neural structures in homeotherms—structures that could learn a lot, learn very fast, and be flexible in guiding behavior, such as defense and finding food. Enter the cortex: immature at birth, it grows very quickly in response to experience, has flexibility, and enables figuring things out. Advance warning: how exactly cortical structures evolved and integrated with "subcortical" structures is not understood. Early homeotherms and premammals are no longer with us (but see Van Essen, Donahue, and Glasser 2018).

Laminar cortex is unique to mammals, though birds have been discovered to have structures anatomically comparable, with an organization more clumpy than laminar (Karten, 2012). The cortical capacity for large-scale learning makes it especially valuable, even though its additional neuronal numbers add to the energy cost that must be made up with yet more food. Benefits notwithstanding, selecting for creatures with cortex introduced yet another problem. Learning entails neuronal growth; growth is achieved by adding branches, synapses, and connections in an organized way to existing neurons. Consequently, at birth, the cortex needs plenty of room for neuronal growth to embody what is learned. How to ensure cortical space for learning and hence for neuronal growth? Immaturity of the neurons (especially those in cortex), which demands immaturity of offspring at birth.

A female turtle will lay her eggs in the sand, and then waddle off to do other things. The baby turtles upon hatching must manage on their own. But for homeotherms, if the offspring are born immature enough to make big learning advantageous enough for survival, the next problem for Mother Nature is how to rig things so that the immature offspring grow to maturity so that they too can reproduce. The solution? Well, mothers are nearby the neonates, because mothers give birth. Modify the brain so that mothers care for infants. In other words, those early mammalian mothers who tended their immature babies were more likely to have surviving offspring than those who wandered off, like mother turtles.

In mammals, vasotocin, an ancient precursor to AVP and OT, was modified and put to a new use in the mammalian brain. Jointly, OT and AVP cause strong attachment of caregivers to offspring and of offspring to caregivers. AVP on its own increases stress levels and aggression. OT on its own increases general sociality and lowers stress levels. Other neuromodulators play a role too, such as the endocannabinoids and endogenous opioids, which make us feel good when released. (Nummenmaa et al, 2016) Along with the new neurohormones, suitable receptors evolved from vasotocin receptors. OT receptors are distributed quite widely in the mammalian brain, including in the amygdala, hypothalamus, and prefrontal cortex. Some 50 million years after the emergence of mammals, the same basic neurohormonal trick was used to ensure attachment by avian mothers and fathers to their hatchlings. Wiring homeotherms for sociality depends on a demanding style of caring, although to judge from existing species, caring seems to be a powerful force.

Apart from highly dependable mother-care, styles of sociality vary considerably among mammalian species. Monogamy is seen in only about 8% of mammalian species, including beavers, marmosets, titi monkeys, gibbons, wolves, and of course prairie voles. Attachment to various others, such as kin and friends, is seen rather more frequently. Highly social species, such as elephants, macaques, and baboons, live in groups and often show care for their kin and friends but are not monogamous. Even mice and rats seek the company of conspecifics and will become depressed when isolated. The social behavior typical of a species is related to its ecology and way of making a living. As Joan Silk (2007, p. 539) succinctly put it, "According to behavioral ecology theory, sociality evolves when the net benefits of close association with conspecifics exceed the costs. The nature and relative magnitude of the benefits and costs of sociality are expected to vary across species and habitats." And vary, indeed they do.

Extending care and attachment beyond offspring to others, such as mates or kin or friends, probably involves relatively minor genetic changes. Small tweaks to OT and AVP receptor distribution may be all it takes to yield alterations to the social style of a particular species. An example used by Frans de Waal (2013) involves comparisons between the behavior of chimpanzees and bonobos, who now live on opposite sides of the Congo River. Initially just one species, they began to differentiate into two species once the animals were separated by an enlarged and swifter Congo River, about two million years ago. Observations show that small changes in dominance hierarchies, for example, can occur quite quickly. Bear in mind that resources are more plentiful and competitors fewer in the territory where the bonobos range, on the south side of the Congo River. Bonobos have matriarchies, whereas chimpanzees are patriarchal. Recently, bonobos have been observed cooperating not only with in-group members but also with unrelated bonobos in other groups. This too is very different from chimpanzees who typically do not tolerate chimps from other groups. Other changes such as hair density, diet, and tail tufts are also observed. Just as there is variability in cooperativeness among humans, there is variability in this trait among bonobos (Samuni and Surbeck 2023). Some variability in sociality may be related to polymorphisms in the gene for OT receptors, and some may be related to wiring differences resulting from differences in early nurturing (Carter 2017).

Cortex learns many things as the infant grows up, and among the things it learns is how to get along and manage in its social world. Basic social impulses are in play from the beginning, but learning social practices and acquiring social skills are critical to all mammalian and avian species. Their lives depend on it. The adult wiring, therefore, will reflect not only genetic endowment but also epigenetic changes resulting from interactions with the environment (Wilczyński, Siwiec, and Janas-Kozik 2019). Depending on such matters as infant cuddling or lack thereof, there may also be methylation of genes for OT and for OT receptors (Wismer Fries et al. 2005; Carter 2017).

Human societies can have rather complex sets of norms that guide the behavior of individuals, and much of moral philosophy begins by assuming that the complexities governing social behavior in, say, European countries in the 19th century, must be typical of humans generally. An evolutionary perspective, however, indicates that norms and traditions prevailing in small groups of hunter-gatherers can be rather simple and flexible, although exquisitely well suited to their specific way of life. There is no benefit to having normative complexity merely for the sake of complexity. (My 2019 book *Conscience* contains an extended discussion of within-species variability.)

Our hominin ancestors lived in small groups of about 25 to 100 individuals—hunting, fishing, and gathering. Some hints of their sociality can be gleaned from anthropological studies of such groups of humans in the 19th and 20th centuries, for example, the *Chinook People of the Pacific Northwest* by James Gilchrist Swan (*The Northwest Coast* 1857), and later, studies of the Inuit of the far north by Franz Boas (*The Central Eskimo* 1888). These studies reveal that in comparison to highly developed cultures in Europe and North America, social norms were simpler, more flexible in application, and interleaved with cherished stories of past disasters and successes. Among hunter-gatherer groups, moral norms varied somewhat and generally did not allocate a role for the supernatural in moral norms. Even when some kinds of supernatural beings were thought to exist, they might be limited to governing bear behavior or to the weather. Differing ecological conditions likely had a significant effect on customs and norms and on their modification as conditions changed. For example, customs for resolving within-group conflicts vary—for example, between the Inuit of the Arctic and the Haida of the Pacific Northwest—probably reflecting tremendous differences in ecology that affect population size and availability of food. To an unknown extent, the preferences and temperaments of persuasive individuals also may have shaped certain customs and norms (Chacon and Mendoza 2007).

The advent of agriculture about 12,000 years ago ushered in tremendous social changes wherever farming replaced hunter-gatherer ways of making a living. Among the changes are included the emergence of large-scale religions and their role as custodians and enforcers of moral norms. Known as the Neolithic Revolution, farming led to a comparative stability in food supply, settled villages, and then larger towns. It resulted in the domestication of plants and animals, the division of labor and invention of new tools and techniques, and a major expansion of the human population. Violence between large groups as they competed for resources became commonplace, with king-priest hierarchies taking the place of loose social structures organized around skills—good hunters for leading the hunt, and good storytellers or good fishers for leadership in those jobs (see North, Wallis, and Weingast 2009).

What about norms and laws as they affect us here and now? When is abortion ethically appropriate? When is war ethically appropriate? How can we settle these issues when we disagree? My provisional response: it is messy—really messy. For one thing, to come to a judgment, one must learn a lot about background matters, although admittedly, time may often be short, for example, when serving on a jury. For another, moral norms often conflict, so it is not obvious which one to prioritize. Finally, as Socrates constantly reminded us, those who proclaim themselves to be ethical experts, to whom the rest of us should defer without question, are exactly the ones we need to size up skeptically. In my Socratic moments, I tend to think it is best to be especially skeptical about the moral certainties of one religion or another. Religions and religious leaders *qua* religious leaders have no particular right to the moral high ground. Some, such as the Dalai Lama may come to deserve respect for their moral views, whereas others, such as Warren Jeffs of the Church of Jesus Christ of Latter Day Saints (also known as the Mormon Church), have shown themselves to be unworthy of respect.

Finally, scientists *qua* scientists have no special normative expertise, although of course they may have factual evidence that is especially relevant. Like anyone else, scientists *qua* social human beings may have acquired morally wise views about how certain norms should or should not be modified in view of the facts. As in any other profession, a scientist might be a moral cretin. So who can we turn to as a moral authority? Socrates's counsel in my paraphrase: Mistrust any single moral authority. Seek advice widely, do your best to figure out what to do. You will not always get it right. Yes of course absolute rules we are certain we can trust would be so very handy and decision-making would be vastly easier. Alas. Life is hard.

The brain's networks *continuously* face constraint satisfaction problems, both social and otherwise. In dilemmas, some considerations are not mutually satisfiable (e.g., saving one child vs. saving another). Typically, constraints are not measurable against each other—for example, how do we measure the value of training soldiers to kill against the cost to them of becoming killers? To a first approximation, the constraints will include immediate desires, but also the force of habits, reputations, the expectations of others, and the evaluation of relevant options. As the relevant constraints weigh in, the networks settle into a solution—the brain's decision. The exact nature of the process whereby networks settle is a largely unsolved problem in computational neuroscience. But the representation of rules and their applicability to the situation at hand seems to be only one constraint among others. As I imagine a contemporary Socrates saying, "it seems that practical reasoning—what to do—mainly consists in finding a good solution to a constraint satisfaction problem. Do your best." That response is not dramatic or flashy or rhetorically resonant. What it does have is more wisdom than the counsel of many a preacher or priest (P. S. Churchland 2008, 2019).

As I am now retired from teaching at UCSD, I have more time now to keep abreast of discoveries in neuroscience and sociality as well as developments in the legal world. Both Tom Albright (Salk Institute) and Francis Shen (Harvard Law) are working on very important problems at the interface of neuroscience and the law, concerning, for example, the courtroom rules regarding which scientific evidence is trustworthy, and the likelihood that a jury will be composed largely of nonscientists and must come to a decision (Albright 2023). This work has important practical implications, and it fascinates me.

Selected Bibliography

Albright, T. D. (2023) A scientist's take on scientific evidence in the courtroom. *Proceedings of the National Academy of Science USA* 120(41): e2301839120.

Averbeck, Murray. (2020) Hypothalamic interactions with large-scale neural circuits underlying reinforcement learning and motivated behavior. *Trends in Neuroscience* 43(9): 649–738.

Boaz, F. (1888) *The Central Eskimo*. Sixth Annual Report of the Bureau of Ethnology, Smithsonian Institute, Washington, D.C. Reprint University of Nebraska Press, 1964.

Carter, C. S. (2017) The oxytocin-vasopressin pathway in the context of love and fear. *Frontiers in Endocrinology* 8: 356.

Chacon, R. J., Mendoza, R. G. (2007) *North American Indigenous Warfare and Ritual Violence*. Tucson: University of Arizona Press.

Chalmers, D. (1996) *The Conscious Mind*. Cambridge, MA: MIT Press.

Churchland, M. M., Shenoy, K.V. (2024) Preparatory activity and the expansive null-space. *Nature Reviews Neuroscience* 25(4): 213–236.

Churchland, P. M. (1989) *A Neurocomputational Perspective*. Cambridge, MA: MIT Press.

Churchland, P. M., Churchland, P. S. (1981) Functionalism, qualia and intentionality. *Philosophical Topics* 12(1): 121–145.

Churchland, P. S. (1978) Fodor on language learning. *Synthese* 38: 149–159.

Churchland, P. S. (1983) Consciousness: The transmutation of a concept. *Pacific Philosophical Quarterly* 64: 80–95.

Churchland, P. S. (1986) *Neurophilosophy: Towards a Unified Science of the Mind-Brain*. Cambridge, MA: MIT Press.

Churchland, P. S. (1989) From Descartes to neural networks. *Scientific American* 261(1): 118.

Churchland, P. S. (2002) *Brain-Wise: Studies in Neurophilosophy*. Cambridge, MA: MIT Press.

Churchland, P. S. (2005) A neurobiological slant on consciousness research. *Progress in Brain Research* 149: 285–293.

Churchland, P. S. (2008) The impact of neuroscience on philosophy. *Neuron* 60(3): 409–411.

Churchland, P. S. (2011) *Braintrust: What Neuroscience Tells Us About Morality*. Princeton, NJ: Princeton University Press.

Churchland, P. S. (2013) *Touching a Nerve: The Self as Brain*. New York: Norton

Churchland, P. S. (2019) *Conscience: The Origins of Moral Intuition*. New York: Norton.

Churchland, P. S. (2021) Hands On. *The Edge* (online). https://www.edge.org/conversation/patricia_s_churchland-hands-on

Churchland, P. S. (2023). Brains and minds. In *Think*, ed. S. Law, 17–23. Cambridge: Cambridge University Press.

Churchland, P. S., Churchland, P. M. (2002) Neural worlds and real worlds. *Nature Reviews Neuroscience* 3: 903–907.

Churchland, P. S., Sejnowski, T. J. (1992) *The Computational Brain*. Cambridge, MA: MIT Press.

Danoff, J. S., et al. (2021) Genetic, epigenetic, and environmental factors controlling oxytocin receptor gene expression. *Clinical Epigenetics* 13(1): 23.

Dawkins, R. (1976) *The Selfish Gene*. Oxford: Oxford University Press.

Dennett, D. (1979) "Consciousness." In: *Oxford Companion To The Mind,* ed. R. Gregory, (Oxford: Oxford University Press); 209–212.

Dennett, D. (2017) *From Bacteria to Bach and Back*. New York:Norton.

De Waal, F. (2009) *The Age of Empathy*. New York: Harmony Books.

De Waal, F. (2013) *The Bonobo and the Atheist*. New York: Norton.

Fodor, J. (1975) *The Language of Thought*. Cambridge, MA: Harvard University Press.

Graziano M.S.A. Webb T.W. (2017) From sponge to human: the evolution of consciousness. In: Kaas, J (ed.), *Evolution of Nervous Systems* 547–554. Oxford: Elsevier.

Grush, R., Churchland, P. S. (1995) Journal of consciousness studies 2(1): 10–29. https://philpapers.org/rec/CHUGIP

Hirstein, W. (2004). *On the Churchlands*. Belmont, CA: Wadsworth.

Jordan, L. M., Sławińska, U. (2011) Modulation of rhythmic movement: Control of coordination. *Progress in Brain Research* 188: 181–195.

Lim, M. M., Murphy, A. Z., Young L. J. (2004) Ventral striatopallidal oxytocin and vasopressin V1a receptors in the monogamous prairie vole (Microtus ochrogaster). *Journal of Comparative Neurology* 468(4): 555–570.

Lim, M. M., Young, L. J. (2006) Neuropeptidergic regulation of affiliative behavior and social bonding in animals. *Hormones and Behavior* 50(4): 506–517.

MacFarquhar, L. (2007). Two heads: A marriage devoted to the mind-body problem. *New Yorker*.

Muller, L., Churchland P. S., Sejnowski, T. J. (2024) Transformers and cortical waves: Encoders for pulling in context across time. Biorxiv.org

Nummenmaa et al. (2016) Social touch modulates endogenous μ-opioid system activity in humans. *NeuroImage* 138: 242–247.

North, D. C., Wallis, J. J., Weingast, B. R. (2009) *Violence and Social Orders*. Cambridge: Cambridge University Press.

Samuni, L., Surbeck, M. (2023) Cooperation across social borders in bonobos. *Science* 382: 805–809.

Sejnowski, T. J., Rosenberg, C. R. (1987) Parallel networks that learn to pronounce English text. *Complex Systems* 1: 145–168.

Schroeder, KE et al Cortical Control of Virtual Self-Motion Using Task-Specific Subspaces, *Journal of Neuroscience* (2022) 42:220–239.

Silk, J. B. (2007) The adaptive value of sociality in mammalian groups. *Philosophical Transactions of the Royal Society B: Biological Sciences* 362(1480): 539–559.

Van Essen, D. C., Donahue, C. J., Glasser, M. F. (2018) Development and evolution of cerebral and cerebellar cortex. *Brain Behavior and Evolution* 91(3): 158–169.

Wenar, L. (2024) The deaths of effective altruism. *Wired*, https://www.wired.com/story/deaths-of-effective-altruism/

Wilczyśski, K. M., Siwiec, A., Janas-Kozik, M. (2019) Systematic review of literature on single-nucleotide polymorphisms within the oxytocin and vasopressin receptor genes in the development of social cognition dysfunctions in individuals suffering from autism spectrum disorder. *Frontiers in Psychiatry* 31(10): 380.

Wismer Fries, A. B., Ziegler, T. E., Kurian, J. R., Jacoris, S., Pollak, S. D. (2005) Early experience in humans is associated with changes in neuropeptides critical for regulating social behavior. *Proceedings of the National Academy of Science USA* 102: 17237–17240.