

The History of Neuroscience in Autobiography
Volume 14

Edited by Thomas D. Albright
and Larry R. Squire
Published by Society for Neuroscience
ISBN: 978-0-916110-13-0

Terrence J. Sejnowski
pp. 380–425

https://www.doi.org/10.1523/hon.014009

Terrence Sejnowski



Terrence J. Sejnowski

Born:

Cleveland, Ohio
August 13, 1947

Education:

Case Western Reserve University, Summa cum Laude, BS (Physics, 1968)
Princeton University, MA (Physics, 1970)
Princeton University, PhD (Physics, 1978)

Appointments:

Postdoctoral Fellow, Princeton University (1978–1979)
Postdoctoral Fellow, Harvard Medical School (1979–1981)
Assistant Professor of Biophysics, Johns Hopkins (1982–1985)
Associate Professor of Biophysics, Johns Hopkins (1985–1988)
Professor of Biology, UC San Diego (1989–present)
Professor, Salk Institute for Biological Studies (1989–present)

Honors and Awards (Selected):

Wiersma Visiting Professor, California Institute of Technology (1987)
Fairchild Distinguished Scholar, California Institute of Technology (1992)
Wright Prize for Interdisciplinary Research, Harvey Mudd College (1996)
Elected, National Academy of Medicine (2008)
Elected, National Academy of Sciences (2010)
Elected, National Academy of Engineering (2011)
Swartz Prize for Computational and Theoretical Neuroscience, Society for Neuroscience (2015)
Gruber Neuroscience Prize, Gruber Foundation (2022)
Brain Prize, Lundbeck Foundation (2024)
Honorary Doctorate, Princeton University (2024)
Member, American Philosophical Society (2025)

At the dawn of computational neuroscience, Sejnowski trained a generation of leaders in this nascent field. He has been a pioneer in developing computational models of synapses, neurons, and neural systems, including the following: (1) MCell, a particle-based Monte Carlo program, which simulates biochemical interactions underlying synaptic plasticity; (2) experimental evidence for millisecond-precision spike initiation in cortical pyramidal neurons; (3) biophysical models of thalamocortical circuits underlying sleep spindles; (4) sleep spindles in human electrocorticograms, which are global, circular traveling waves; (5) transients in dopamine neurons, which represent reward prediction error; and (6) independent component analysis (ICA) for blind source separation, which is routinely used to analyze EEG, MEG, and fMRI recordings.

Terrence J. Sejnowski

Early Life

I was born in Cleveland, Ohio, and spoke Polish before I learned English. My parents said I gave Polish lessons to the kids I played with on the street. I was also fast at finding pieces in jigsaw puzzles, to the amazement of my parents and their friends at their dinner parties. These two early talents later became passions: Teaching and solving scientific puzzles. My father was an engineer who designed jet engine fan blades. My mother was a born organizer who ran a pool of typists during World War II (WWII) and organized her five sons and one daughter. I was the oldest. I attribute my self-confidence to my parents' trust in my early exploration of the world and the freedom they gave me to pursue my interests while growing up. I did not fully appreciate this then, but I now know how much I owe them.

I was the science guy in grade school. I vividly remember building a working papier-mâché volcano fueled by a mixture of homemade gunpowder and aluminum powder I had concocted in my basement chemistry lab (figure 1). It spouted an impressive, sparkling flame, and black smoke when I tested it outdoors. When I demonstrated it to my grade-school class, I failed to realize it would fill the classroom with smoke, trigger a fire alarm, and force an evacuation. This event made my reputation as the science guy.

I was president of the Radio Club at Saint Joseph High School. I stayed after school every day to build electronic devices and talk with ham radio enthusiasts worldwide. Mike Stimac, the faculty adviser, had tracked the first Sputnik satellite in 1957 with "the boys," a story picked up by the national press. We obtained a commercial radio transmitter and assembled



Fig. 1. Terry in his basement chemistry lab around the time he demonstrated a working volcano in his classroom.

a giant Yagi antenna on the school's roof for Project Moonbounce (figure 2). Mike once asked me, "What is your mission in life?" This question has been at the back of my mind ever since then. Mike was also the faculty adviser of the Aviation Club, where I learned how to navigate and fly light aircraft. However, I was not old enough to qualify for a pilot's license.

In my senior year, three of my stories appeared in *The Carpenter*, the high school literary magazine. One was about gravity and how it works; the second was a science fiction story about someone who experienced an inexplicable blank period in his life, which, at the end, was caused by a failure of a planet-wide computer that was simulating him; the third was a meditation by someone who was selling the house where he grew up. The first essay reflected a burgeoning interest in physics; the second story about the world as a digital simulation predated *The Matrix*; and the third story was a premonition of my family's move from the inner city to the suburbs.

I matriculated at Case Institute of Technology. My diploma was from Case Western Reserve University, after Case merged with Western Reserve University next door. I fell in love with the beauty of physics and its lore. By my senior year, I was taking graduate courses. The Andrew R. Jennings Computing Center had a UNIVAC® 1107 mainframe where I spent countless hours. I learned how to program in Algol by reading discarded programs. In addition to my scientific projects, I worked part-time as a digital records manager in the registrar's office. When the director abruptly left, I was responsible for processing all the grade transcripts.

I was intrigued by astrophysics and participated in a summer program at the National Radio Astronomy Observatory in Charlottesville, Virginia. I visited the radio telescopes in Green Bank, West Virginia, a radio-quiet valley. It was magical being in the presence of giant radio telescopes, pointed up at the stars, listening in on radio signals from distant galaxies that originated billions of years ago. I found a complete set of the Palomar Sky Survey—wide-field photographs filled with distant galaxies. I was awestruck at the vastness of the universe. I worked with Robert Hjellming

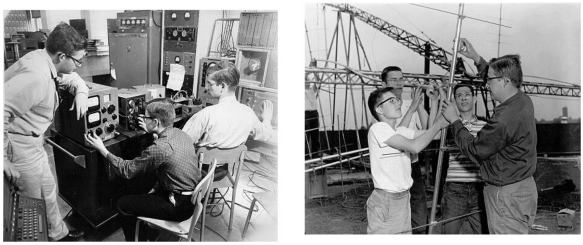


Fig. 2. Terry at the Saint Joseph High School Radio Club in 1964. *Left*: Tuning a receiver with the kilowatt transmitter behind him. *Right*: Preparing the antenna for Project Moonbounce and the swinging tower in the background.

on a computational astrophysics project to predict the ratios of transitions between highly excited states of hydrogen in large clouds of hot, ionized gas in interstellar space, created by the high-energy ultraviolet radiation from young, massive stars. This computational project led to several papers in the *Astrophysical Journal*, my first scientific publications (Hjellming, Andrews, and Sejnowski, 1969).

My most memorable adventure in college was when Mike Stimac called and asked if I could help him copilot a Cessna 210 to the West Coast. When we flew into Los Angeles, I had a few days free to visit Caltech in Pasadena, a mecca for theoretical physics. Without having an appointment, I met with Carl Anderson, chair of the Physics Department, who received a Nobel Prize for discovering the positron. He asked me whether I wanted to be an experimentalist or a theoretician. It had never occurred to me that I had to choose. I said both. He said there was someone at Caltech who did both, George Zweig, so there was an existence proof. Since then, Caltech has been alluring to me, not least because of Richard Feynman's inspirational lectures. When I applied to graduate school, Caltech turned me down, but Princeton accepted me. After I received a prestigious four-year National Science Foundation (NSF) graduate fellowship, Caltech reached out to me. By then, however, I was on my way to Princeton.

Princeton

I took a graduate course At Case Western Reserve in general relativity from Joseph Weinberg, which made a strong impression on me and on a close friend, Jonothan Logan. Relativistic astrophysics was an exotic frontier. We later reconnected in graduate school, where he worked with Mark Kac at Rockefeller University in New York, and again when he moved to Brookhaven National Laboratory on Long Island to work with Sam Goudsmit.

Relativistic Astrophysics

At Princeton, I fell under the spell of John Wheeler, a legendary physicist who coined the term *black hole*. John was fascinated by explosions. One summer, he invited me to his vacation home on High Island in Maine. At dinner one day, he lit a lady-finger firecracker that landed on my pants and burned a small hole. Wheeler had worked on the atomic bomb and the hydrogen bomb. On my first day in his lab, he asked me what a black hole would look like if one were in the middle of our galaxy. I went to the computer center and simulated the orbits of stars slowly falling into a giant black hole. When a star reaches the event horizon, it is stretched and compressed, radiating enormous amounts of energy. Wheeler was encouraging and said this might explain the mysterious signals coming

from quasars (quasi-stellar radio sources). If true, this would be a major discovery. Brandon Carter invited me to talk at the Department of Applied Mathematics and Theoretical Physics at Cambridge University in 1970. I was surprised by the audience's skeptical questions about the existence of black holes. I later overheard Martin Rees, an influential astrophysicist who later became Astronomer Royal of England, not taking black holes seriously.

Wheeler encouraged me to publish my results. I asked Lyman Spitzer to read my paper. He was the chair of the Department of Astrophysical Sciences at Princeton and influential in lobbying for space telescopes (one was named after him). He pointed out that there was no experimental evidence for black holes anywhere in the universe, and he cautioned me about starting my career working on something that most likely did not exist. So I asked Freeman Dyson at the Institute for Advanced Study, who was known for his imaginative ideas, for feedback on my paper. He told me he had ideas like this every day, but explained that he did not publish all of them. So I decided not to publish mine, which was the biggest mistake of my career, because we now know black holes do exist. There is even one in the center of our galaxy. But this experience gave me confidence in my intuitions and made me distrust expert opinions about new ideas.

I switched to another problem in relativity and worked on sources for gravitational waves. Joe Weber at the University of Maryland reported simultaneous ringing in two giant aluminum cylinders separated by a thousand miles, which looked like they might be caused by passing gravitational waves. I calculated the expected signal strengths from known astrophysical sources, such as supernovae and colliding black holes. I concluded that Weber's cylinders were a thousand times below the sensitivity of all known sources. The signals were soon revealed to be a bug in his analysis program, and I concluded it would be many decades before detectors could reach the supersensitivity needed for detection. The prospect of waiting for decades for experimental data was sobering. I contacted Kip Thorne at Caltech, a former Wheeler student working on general relativity, who thought it would take eight years and come from a supernova in the Virgo cluster of galaxies. I quoted Kip in a paper I published on possible sources (Sejnowski, 1974). Gravitational waves were finally discovered in 2015, 40 years later. I sent Kip a congratulatory e-mail reminding him of his predictions. He replied: "I was wrong about the Virgo cluster, but what's the difference between 8 years and 40 years on a cosmic timescale?"

As a graduate student in the 1970s, I sat for a weeklong general exam at the end of my first year. Every morning and afternoon, I was tested on a different area of physics: classical mechanics, quantum, mechanics, electricity and magnetism, statistical mechanics, thermodynamics, astrophysics, nuclear physics, condensed matter physics, general relativity, and, on the final day, a defense of an experimental project, a requirement

even though I was on a theory track. I obtained a highly radioactive source of positrons. When a positron annihilates with an electron, photons are emitted: most commonly, two photons, and rarely three. I measured the decay branching ratio by setting up a jerry-rigged apparatus in the attic of Palmer Physical Laboratory. I returned a month later to collect the data and calculate the branching ratio. It agreed with quantum mechanics. I thought of the experiment and carried it out independently without a faculty supervisor, much like my working volcano in grade school. In retrospect, having a highly radioactive source sitting around was problematic.

A Glimmering New Direction

I prepared for the general exam by reviewing all of physics in the attic of Joseph Henry Hall during a hot and muggy summer. For diversion, I took a summer course in French, which, along with German, counted as one of my two required languages. I also read transcripts of the Josiah Macy Jr. Conferences, a series of workshops held in the 1950s focused on cybernetics as an approach to understanding brain function (Josiah Macy Jr. Foundation, 1955). The eclectic scientists included Norbert Wiener, John von Neumann, Jerome Wiesner, Ralph Gerard, Margaret Mead, Warren McCulloch, Gregory Bateson, Heinz von Foerster, Donald MacKay, and Claude Shannon. The discussion abruptly stopped at the end of the 10th volume (1953), and I was left hanging. What was the impact, and where was the field decades later?

I was curious about brain electrical signals, so I walked into a random biology lab in Guyot Hall and asked a random person what was known about the patterns of neuronal spikes. Were they synchronous, like a clocked computer, or were they asynchronous? She did not know the answer, but told me there was a paper by David Hubel (see volume 1) and Torsten Wiesel with recordings from neurons that could answer the question. When I tracked down their 1968 discovery paper on orientation tuning in the visual cortex, I found the answer. Unlike subsequent papers that reported trial averages, they showed raw spike trains. Not only were the spikes asynchronous, but responses to the same stimulus elicited different spike patterns. On some trials, there was no response at all, just background "spontaneous activity." I was surprised, and still am.

Hubel and Wiesel inspired my first neuroscience papers. I modeled feature-detector neurons and their dynamics using neural networks (Sejnowski, 1976a, 1976b). I was also interested in synaptic plasticity. Donald Hebb predicted the conditions under which synapses should increase their strength by coincident presynaptic and postsynaptic activity. I added conditions under which synapses should decrease their strength, anticipating future discoveries (Sejnowski, 1977a, 1977b).

Transition to Neuroscience

I have always been intrigued by brains, which have an inner complexity as mysterious as the cosmos. While working on the cosmos with John Wheeler, I sat for courses from Charlie Gross (see volume 6), a physiological psychologist who worked on the visual system of monkeys; Mark Konishi (see volume 6), a neuroethologist who worked on birdsong and sound localization in barn owls; a neuroanatomy class with Darcy Kelly (see volume 12); and a mathematical biology course with Robert May on "Life Among the Schistosomes." I was accused of being a "scientific playboy" by my girlfriend's parents, but how else was I to look for new scientific challenges?

One summer, I participated in recordings from the inferotemporal cortex (IT) in Charlie's lab. He had already discovered cortical neurons in monkeys that selectively responded to faces. Bob Desimone (see volume 13) and Tom Albright, who later became valued colleagues, were recording from IT, and we were looking for neurons with other exotic responses. The chronic experiments would continue 24/7 for days. Reading the classic papers did not prepare me for the experience of listening to neurons' spikes in response to visual stimuli. We had a box of random objects that we presented in search of the neuron's preferred stimulus, a small set of trial objects from an infinite space of objects. Curiously, many IT neurons responded to a toilet brush with black bristles around a circular wire. At the time, Charlie, Bob, and Tom all had black beards around their faces that looked like toilet brushes. The IT was known to be involved in visual learning, and I suspected those IT cells were beard detectors.

I had taken a course in biophysics with John Hopfield. He developed an interest in neuroscience and was attending workshops of the Neuroscience Research Program (NRP) at Massachusetts Institute of Technology (MIT). The NRP also held annual meetings that brought together neuroscientists across the spectrum. John would tell me about the latest discoveries in neuroscience. He introduced me to Bruce Knight, who worked in the Hartline lab at the Rockefeller University. Bruce's model of the neural network in the Limulus retina, based on experimental recordings, was inspiring. I eventually received a doctorate in physics under John Hopfield's mentorship in 1978 for a thesis on "A Stochastic Model of Nonlinearly Interacting Neurons." It had the kernel of an idea that I would return to much later in my career: a second-order tier of correlated spikes riding on top of the first-order tier for fast sensorimotor actions based on rate-coded neural activity, which was the dominant experimental paradigm (Sejnowski, 1976).

I transitioned to neuroscience by taking the Neurobiology Summer Course at the Marine Biological Laboratory at Woods Hole. This in-depth, 10-week immersion taught me a wide range of experimental techniques, including electrophysiology, neuroanatomy, biochemistry, neuropharmacology, and tissue culture. I loved every minute, recording muscle action potentials, running hot-zap gels, and looking into neurons with electron

microscopy. This led to my first publication in experimental neuroscience on freeze-fracture of the ampullae of Lorenzini, ultrasensitive electroreceptors in skates and sharks (Sejnowski and Yodlowski, 1982). I was curious how neurons operating at millivolt levels could detect microvolt signals. Electroreceptors have ribbon synapses that continuously deliver vesicles to the active zone on a conveyor belt.

After finishing my doctorate, I worked for a postdoctoral year with Darcy Kelly and Alan Gelperin, who studied learning in the garden slug, *Limax maximus*. This led to a paper on activity mapping in single neurons in a *Limax* ganglion using freeze-substitution, which looked like a starry sky and appeared on the cover of *Nature* (Sejnowski et al., 1980). I was strongly influenced by the neuroethological approach of these researchers, who studied the neural basis of natural behaviors in a variety of species. Comparative studies of brains and behavior provide context that is not possible by focusing on restricted behaviors in only a few species that can be genetically manipulated.

Harvard Medical School

The Woods Hole course was taught by faculty from the Department of Neurobiology at Harvard Medical School led by Ed Kravitz (see volume 4). I stayed at Woods Hole for a few more weeks into the fall to finish the skate project. It was quite a surprise when Steven Kuffler, the father of modern neurobiology and chair of the Neurobiology Department at Harvard, called me at Woods Hole to ask if I was interested in working with him as a postdoctoral fellow. The summer course led to a two-year immersion at Harvard's Neurobiology Department. I took a vow not to touch a computer, as it would distract me from learning the lore of neurobiology. Steve and I experimented daily, voltage-clamping a peptidergic synapse in bullfrog sympathetic neurons and solving a mystery in the literature about the underlying conductance changes (Kuffler and Sejnowski, 1983).

The Harvard Neurobiology Department felt like a family. Steve Kuffler was the patriarch. When he moved from Johns Hopkins in Baltimore in 1959 to found the department, he took his postdocs David Hubel, Torsten Wiesel, Dave Potter, and Ed Furshpan. Bob Bosler, Steve's electronics technician, organized the move. Bob was an electronics specialist in the Navy and handled the logistics of the move. He also organized our annual move to Woods Hole, where he set up our rig for a summer of recording. Steve's many friends would stop by the lab during the summer, a who's who of neurobiology. When I later organized a series of computational neuroscience workshops in Woods Hole, Bob Bosler generously organized our workshop picnics at his home on Gansett Road. He was my closest friend at Harvard. I greatly miss him.

The labs in the Neurobiology Department had strengths in molecular, cellular, and systems neurobiology. Like a baby learning to swim by being thrown into a swimming pool, I learned by osmosis how to think like a biologist. All electrons in the universe are identical in physics, but nothing is identical in biology, a science based on diversity. Everyone attended evening meetings with labs taking turns presenting their recent research after dinner. Questions would continue late into the night. Scientific standards were high, and I learned how to give a biological talk. Lunch seminars by visiting faculty were lessons in how to critique scientific experiments.

I also kept up my interest in neural networks by following the research by Jack Cowan, David Marr, Christoph von der Malsburg, and many others. I met David Marr at a small ski workshop at Jackson Hole, and he invited me to visit him at the AI Lab at MIT. Marr was trained as a mathematician at Cambridge and published an influential paper on cerebellar learning. His stereo model with Tommy Poggio (see volume 8) was influential in shaping my early neural modeling papers (Marr and Poggio, 1976). At MIT, he focused on the visual system. Marr attracted an impressive group of students, a small island of brain researchers in a sea of computer scientists. I gave a talk to his group about my views on how the visual cortex is organized as a dynamical system. I later learned that Marr came to MIT on the recommendation of Sydney Brenner through his friend Seymour Papert, another South African. Marr's early death from leukemia was a great loss. I once gave the Marr Lecture at Trinity College, Cambridge, and met Alan Hodgkin, who at the time was the President of the Royal Society (see volume 1). Alan recounted Lord Adrian's advice: "Stay away from smooth muscle." This advice was mysterious until I learned about the importance of internal signaling in the heart muscle mediated by second messengers that modulated electrical responses.

A turning point came when I attended a workshop at the University of California, San Diego (UC San Diego) in 1979, organized by Geoffrey Hinton, who had the same intuition as mine—that the only proof that any problem in artificial intelligence (AI) could be solved was that nature had already solved them. His computational perspective was a new constraint on my quest to understand brain function. My training in physics and neuroscience complemented Geoff's in psychology and AI. Geoff at that time was a postdoctoral fellow in the Parallel Distributed Processing (PDP) group at UC San Diego, led by Jay McClelland and Dave Rumelhart. In 1986, this small group would produce a highly influential two-volume book in the burgeoning field of neural networks. Francis Crick had a chapter in the book on "Certain Aspects of the Anatomy and Physiology of the Cerebral Cortex." My chapter was on "Open Questions about Computation in the Cerebral Cortex" (Sejnowski, 1986). This chapter was the road map for my future research directions.

Johns Hopkins

Steve Kuffler died suddenly while we were writing up our experimental results, leaving me without a mentor. Fortunately, with help from Torsten Wiesel, who had become the department chair, I landed a job in the Thomas C. Jenkins Biophysics Department at Johns Hopkins University on the Homewood campus. I could not have asked for a better department or university to begin my scientific journey. Faculty in the department were highly supportive, and I was fortunate to attract exceptional students.

Before I arrived, I received a letter from Vernon Mountcastle, welcoming me and inviting me to visit him at the Bard Laboratory at the medical school. I had read his classic papers and was stunned to receive this invitation. When I visited him, I met the other faculty in the Bard Laboratory with whom I had an immediate affinity, including Gian Poggio, Ken Johnson, and Apostolos Georgopoulos (see this volume). After my talk, Vernon invited me into his office and told me he appreciated my theoretical approaches to neuroscience, but apologized that he was already working with another theorist, Gerald Edelman. I was speechless, but felt welcomed and benefited from my association with the Bard Lab.

Setting up a lab from scratch is daunting, especially because I wanted to pursue both experimental and computational problems. I wanted to test my hypothesis that long-term depression (LTD) should be induced when presynaptic inputs were anticorrelated with postsynaptic outputs. We did indeed find evidence for LTD when presynaptic inputs were paired with delayed postsynaptic spikes, but the finding was not robust (Stanton and Sejnowski, 1989). If we had moved the pairing closer in time, we might have discovered spike-timing-dependent plasticity (STDP). We were more successful studying other forms of synaptic plasticity. I greatly enjoyed teaching and initiated a graduate course in computational biophysics. A researcher in my class, Andreas Andreou, became an early leader of neuromorphic engineering and a lifelong friend. I also enjoyed the collegiality of the faculty. I joined a faculty lunch group that included Doug Fambrough and Richard Cone. We called ourselves the "New Vitalists" and would discuss questions like "How does a cell know how big it is?"

I was at Harvard Neurobiology when David Hubel and Torsten Wiesel won their Nobel Prize in 1981. Before receiving the Nobel, David Hubel would do all-night solo experiments, and I would hear the crackling of spikes from his lab at 3 a.m. After the Nobel, he never did another experiment. The next year, I was at the Evergreen House at Johns Hopkins for the annual reception for new faculty in October. The Nobel Prizes were being announced. I told a small group my story about David Hubel and said it was a career-ending prize I would not welcome. The person standing next to me asked for my name and told me he was Steven Muller, the president of Johns Hopkins. Many years later, Geoff Hinton would win a

Nobel Prize in Physics with John Hopfield for their contributions to AI. Geoff, too, stopped doing research and has a full-time assistant to buffer him from requests.

At the time, there were relatively few neuroscientists at the Homewood campus. Most were at the medical school. I had great discussions with researchers in the Psychology Department, including David Olton, who studied place fields and navigation in the hippocampus, and neuropsychologist Alfonso Caramazza, who studied the organization of language processing and object and action representation in the human cortex. We submitted a proposal to Steven Muller to start a small neuroscience institute on the Homewood campus, which would also benefit students. Our proposal crossed his desk at the same time as another from a group of cognitive scientists in the Psychology Department. Muller thought these two proposals might be combined and called a meeting at the Evergreen House of faculty and administrators across both campuses. He asked us two questions: Is it a good idea to integrate mind and brain within one institute, and is Johns Hopkins the right place? We all agreed, and Mountcastle got the ball rolling by stating he knew of substantial funding in the private sector. The Zanvyl Krieger Mind/Brain Institute (MBI) at Johns Hopkins was established in 1994.

I received several grants at the beginning of my career at Johns Hopkins. Geoff Hinton received a late-night call from someone who identified himself as Charles Smith from the Systems Development Foundation (SDF), which funded high-risk research unlikely to succeed. Geoff Hinton thought it might be a prank call, but it was legitimate, and he received major funding from them. The SDF funded my lab when Geoff thoughtfully told them my research was even less likely to succeed. I also received a Presidential Young Investigator award from the NSF for research that was likely to succeed. Neural network research is highly computation-intensive. I made rapid progress because, with these grants, my lab had more computing power than the entire Department of Computer Science at Johns Hopkins. After moving to San Diego, support was provided by Thomas McKenna, a program manager at the Office of Naval Research, who led a Computational Neuroscience Program that continues to this day.

Woods Hole Workshop on Computational Neuroscience

On my way back from the West Coast in 1985, I was snowed in at the Denver Stapleton airport. Steven Zucker, who pioneered an iterative constraint propagation approach to computer vision, was also stranded, and we commiserated over hot chocolate. We discussed ways to model the visual cortex and the need for a small workshop to discuss computational models with experimentalists. The 1986 Woods Hole Workshop on

Computational Neuroscience was born from this chance meeting and has met annually with 10–15 researchers. Discussion usually began with the first slide. This workshop has lasted 40 years and has served as a crucible for hundreds of researchers to explore new ways to understand brain function. The first workshops were funded by the Sloan Foundation and later by the Office of Naval Research. We only invited researchers who were early in their career. Many later became leading researchers, including Michael Stryker, Patricia Churchland, Robert Desimone, William Newsome, John Maunsell, Christof Koch, and Rodney Brooks.

Telluride Workshop on Neuromorphic Engineering

I first met Carver Mead at a small connectionist workshop near Pittsburgh in 1984, organized by Geoff Hinton. I was impressed by his silicon retina, an ultralow-energy, lightweight alternative to traditional frame-based cameras, based on the same principles as the retina in our eyes. Carver is a visionary who realized that transistors on VLSI chips near threshold exhibited dynamics similar to those of neuronal ion channels. He wrote a seminal book in 1989 titled *Analog VLSI and Neural Systems*, which I found fascinating. I had a wonderful scientific opportunity to interact with colleagues at Caltech as a Fairchild Distinguished Scholar in 1992–1993 (figure 3). Carver's lab adopted me, and I learned a lot about how analog chips are designed and communicate with each other. Christof Koch, Rodney Douglas, and I applied to the NSF to support a Neuromorphic Engineering Workshop, first held in Telluride, Colorado, in 1994 (figure 4). This workshop continues today and is the longest-running NSF-sponsored annual workshop. The Woods Hole Computational Neuroscience Workshop moved to Telluride, Colorado, in 1997.



Fig. 3. Carver Mead discussing neuromorphic chips during my sabbatical at Caltech in 1992.



Fig. 4. Participants in the first Telluride Neuromorphic Engineering Workshop in 1995. Terry is on the far left in the first row, next to Rodney Douglas and Misha Mahowald, two influential neuromorphic pioneers. Christof Koch is in the center of the last row, with Dana Ballard to his right.

Computing with Neural Networks

After the 1979 workshop at UC San Diego, Geoff Hinton and I collaborated with Dana Ballard to publish a paper in *Nature* in 1983 on "Parallel Visual Computation" (Ballard, Hinton, and Sejnowski, 1983). Dana was a proponent of active vision integrated with motor control rather than image analysis, which he called "dead-eye" vision. In contrast to the geometric methods used in computer vision, constraints were built into the neural network weights. We also organized small workshops on connectionist models during the summer. John Hopfield presented his new network model at a workshop at the University of Rochester. It occurred to me that if we added noise to "heat" up the Hopfield net and simulated annealing to slowly cool it down, we could find optimal solutions to vision-based constraint satisfaction problems. Geoff and I called it a Boltzmann machine after Ludwig Boltzmann, a 19th-century physicist who pioneered statistical mechanics and inspired our model.

I was at Johns Hopkins in Baltimore, and Geoff had a faculty position in Computer Science at Carnegie Mellon University in Pittsburgh. We drove back and forth on weekends. Finding a learning algorithm for multilayer neural networks was a long-standing logjam, and it was generally believed that none was possible. I remember an excited phone call from Geoff, who had an insight that led to an elegant learning algorithm for training the weights in Boltzmann machines (Ackley, Hinton, and Sejnowski, 1985). Our collaboration was magical, and I look back at it as the high point in my career. The Boltzmann learning algorithm broke the logjam. Neither of us could have done it on our own. Geoff and Dave Rumelhart also invented another learning algorithm, backpropagation of errors, which was much

more computationally efficient but less elegant and less biologically plausible than the Boltzmann learning algorithm.

Geoff moved to Toronto and established a burgeoning research group focused on scaling up the Boltzmann machine. He continued on this path for 30 years toward the holy grail of AI. He eventually succeeded by sequentially training a deep Boltzmann machine, layer by layer, freezing the weights and training the next layer. The development of the visual cortex follows a similar sequential path, with the critical period of plasticity closing first in the primary visual cortex early in life. The prefrontal areas are the last to mature in early adulthood. My path was toward understanding brain function. We stayed in contact over the decades and enjoyed hearing from each other. Periodically, Geoff would call to announce, "I have discovered how the brain works." And sometimes he was right. For example, he discovered that if you drop out half the units in each training epoch, performance improves by 10 percent, which could explain why synapses in the cortex are probabilistic, with most excitatory synapses not releasing any neurotransmitter most of the time. It also conserves energy.

With learning algorithms, we could create small network models that solved computational problems, giving us insight into how brains might solve them. The computers at the time were primitive compared with today's, which limited us to a single layer of "hidden" units between the input and output layers. I trained neural networks to solve problems in vision, which was the best-understood part of the cortex. In the 20th century, recordings from single neurons in the visual system revealed a multilayer hierarchy in which neurons responded to simple patterns, such as lines and edges in the first cortical layer, and more complex patterns, such as faces, in higher layers.

Shape from Shading

In 1988, Sydney Lehky, my first postdoctoral fellow, and I trained a neural network model to compute the curvature of a shaded surface. We found that the hidden units had properties of simple cells, as discovered by Hubel and Wiesel in the primary visual cortex of cats, that responded best to edges and lines, despite the absence of edges and lines in the shaded images we used to train the networks. Our paper in *Nature* introduced the term "projective field" of a neuron—its output targets—which was as important as the input receptive field in determining the function of a hidden unit in this simple neural network. Horace Barlow once gave a talk at the Salk Institute and singled out this paper as transformative. I suspect that he reviewed it.

Text-to-Speech

Of all my neural network projects from that era, the one that had the most impact was NETtalk, a network trained to translate text to speech

(Sejnowski and Rosenberg, 1987). Charles Rosenberg was a graduate student with George Miller at Princeton and visited my lab in the summer of 1986. English phonology is notoriously difficult because of many exceptions to regularities and influences from other languages. Phonologists have identified hundreds of rules, each with many exceptions and subrules among the exceptions. The network used a moving window of seven letters, a single hidden layer, and the output layer was trained to predict the phoneme of the middle letter. Remarkably, the learning process produced babbling sounds, mastered the regularities, and finally the exceptions, all of which were accommodated in a network with 200 units and 20,000 weights, tiny by today's standards. This demonstrated that a small feed-forward network could capture the complexities of phonology, a real-world problem not easily handled by handcrafted rule-based systems. In retrospect, NETtalk was a harbinger of the remarkable ability of much larger networks to perform natural language tasks, such as language translation and generative dialogue.

In 1986, Geoff Hinton and I organized the first Connectionist Models Summer School at Carnegie Mellon, which brought together a group of young, enthusiastic researchers for intensive immersion in the new neural network paradigm (figure 5). They held a skit in which they formed three lines of human processing units and simulated NETtalk, trying and failing to pronounce "Sejnowski" (the "j" is pronounced like a "y").

Protein Structure

Steven Kuffler's first faculty position in the United States was at the Wilmer Eye Institute at Johns Hopkins University. I once asked him why he didn't follow up on his classic *in vivo* recording of ganglion cell activity in cats. He said he had only done that experiment because he was at an eye



Fig. 5. The 1986 Connectionist Models Summer School at Carnegie Mellon University. Terry is in the first row with Geoffrey Hinton to his right and Jay McClelland next to Geoff. Others include Jonathan Cohen, Stan Dehaene, Richard Durbin, Mike Mozer, Yann LeCun, Dana Ballard, Michael Jordan, and Andy Barto.

institute and felt guilty about studying synapses. He had two good postdocs and encouraged them to follow the signal up into the brain—David Hubel and Torsten Wiesel. I was in a biophysics department and felt guilty about working on neural networks, so I asked a new graduate student to take on a challenging biophysics project.

Ning Qian had come to the United States from China with graduate-level training in physics and had never taken a biology course. When I asked him what was in the nucleus, he replied, "protons and neutrons." His rotation project was to predict protein secondary structure. This is a complex problem for methods based on amino acid biophysical properties. The 20 amino acids are like letters in the alphabet, and secondary structures, such as alpha helix, beta sheet, and random coil, are like phonemes. Ning used a training set from the Brookhaven Protein Data Bank, which had structures determined by X-ray crystallography.

Remarkably, the performance of his neural network model was much better than the biophysical methods. The paper was published in the *Journal of Molecular Biology* and has been cited 1,766 times (Qian and Sejnowski, 1988). There was no explicit biophysical knowledge in the model, so the problem had to be solved by pattern recognition from known structures. This may have been the first time machine learning was applied to a biophysical problem. Extracting structural patterns from structural data would later succeed in protein folding, a much more difficult problem.

Backgammon

Games were an early focus for AI. Games interest humans, but they are much simpler than the real world. Backgammon is a race to the finish between two players, with pieces that move forward based on each roll of the dice, passing through one another on the way. Unlike chess, which is deterministic, backgammon is governed by chance: The uncertainty of every dice roll makes it harder to predict the outcome of a particular move. Gerry Tesauro and I thought this would be a good problem for neural networks. He was working with Stephen Wolfram at the Institute for Advanced Study in Princeton.

Rather than write a program based on logic and heuristics to handle all possible board positions, an impossible task given that there are 10^{20} possible backgammon board positions, we taught a network to play through pattern recognition from games played by humans (Tesauro and Sejnowski, 1989). The network played at an intermediate level and would sometimes make terrible moves. Gerry went to IBM Thomas J. Watson Research and created the first backgammon program that played at world-championship levels by having the backgammon network play itself and use reinforcement learning. This same architecture was used in AlphaGo to beat the world Go champion.

Lip-Reading

Most engineers at the time believed that Minsky and Papert had proven that neural networks were a dead end and would never solve any engineering problems. Engineers at Johns Hopkins reacted this way to a doctoral project by Ben Yuhas, a student in the Electrical Engineering Department. Ben was jointly mentored by an engineering faculty member, Moise Goldstein, who studied auditory systems. Ben trained a neural network to lip-read. The shape of the mouth carries information about speech sounds, and seeing faces helps those with hearing loss. Ben trained a network to predict the audio spectrum from images of speakers' faces (Yuhas, Goldstein, and Sejnowski, 1989). His network could augment any speech recognition program in a noisy environment. When it came time for Ben to defend his thesis, the faculty objected that his claimed performance was not possible. I successfully defended Ben's thesis. This view persisted in engineering long after the evidence against it was unequivocal.

By the time I left Johns Hopkins in 1989, the tide had turned, and neural networks were on the upswing, despite many holdouts at places like the MIT AI Lab. In the previous year, I gave 50 lectures around the world and received several job offers, including two in San Diego, which I accepted. I am now on the faculty at both the Salk Institute and UC San Diego.

San Diego

When I attended the workshop organized by Geoff Hinton in 1979 at UC San Diego, I visited Francis Crick at the Salk Institute for Biological Studies. His office had a large model brain used to teach medical students, which I later inherited. Little did I know that I would be moving to La Jolla 10 years later, in 1989, and starting the Computational Neurobiology Laboratory (CNL). My immediate neighbors were Simon LeVay and David Amaral. Richard Andersen (see volume 12) had recently moved to Caltech. Tom Albright had labs nearby that were recording from a monkey visual cortex. I inherited file cabinets from Max Cowan (see volume 4), the founding editor of the *Journal of Neuroscience*, filled with reviews of old papers. A few years later, Chuck Stevens moved to the Salk Institute and was down the hall. It was an exceptional privilege to have Francis Crick as a colleague.

Tea Time

CNL has a daily afternoon tea (figure 6). We had researchers with many different backgrounds and perspectives. Francis Crick came to our daily tea and often started a discussion that lasted for hours. A portrait of Crick now hangs in our tea room (figure 7). He switched from molecular genetics



Fig. 6. Tea time at the Computational Neurobiology Laboratory at the Salk Institute in 2008. Mary Ellen Perry, my lab manager and manager of the annual Neural Information Processing Conference (NeurIPS), is sitting next to me. Tom Bartol is standing behind her. Francis Crick frequently attended tea.

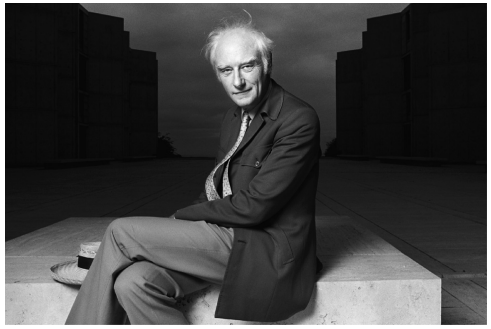


Fig. 7. Francis Crick would attend our daily tea. This portrait of Francis now looks down from a wall in the CNL tea room.

to neuroscience when he moved to Salk in 1976, and he pursued an interest in consciousness, a topic long neglected by biologists. He focused on visual awareness (Sejnowski, 2015). One day, when discussing brain modeling, Francis told us that making a brain model was not an end in itself. He said the true purpose of a model was to design a nonobvious killer experiment that would "give the game away." His famous model of DNA with Jim Watson came to mind.

Howard Hughes Medical Institute

Shortly after moving to La Jolla, I received generous support from the Howard Hughes Medical Institute (HHMI), enabling me to pursue new research directions for many decades. Max Cowan had become the scientific director and once made an unannounced visit on Christmas Eve. He had been the chief scientific officer at Salk, kept his house just north in Cardiff, and visited it on holidays. I happened to be reading an article

in the *Journal of Comparative Neurology*. He once edited the journal. I was surprised when he commented that the authors were using old-fashioned methods. I loved looking at the beautiful anatomical figures, which, for me, were neuro-porn. Before connectomics, neuroanatomy was an art form.

Institute for Neural Computation

I held a joint faculty appointment at UC San Diego, where I taught courses in neurobiology and founded the Institute for Neural Computation (INC). I felt right at home in La Jolla because the PDP group had laid the groundwork at UC San Diego, and several faculty members were already working on brain modeling. Students and postdoctoral fellows joined my lab with backgrounds that ranged from biology, psychology, and bioengineering to physics, computer science, and philosophy. I had the best of both worlds: a lab in a small, high-level biological community and an institute across the street at a university with a full range of engineering, physical sciences, and biomedical departments.

Stephen Wolfram

Stephen Wolfram began exploring complexity at about the same time that I started working on neural networks in the 1980s (figure 8). He founded the Center for Complex Systems Research at the University of Illinois in Champaign, Illinois. Stephen is a deep thinker and writes cogently on a wide range of topics. He also founded Mathematica, a widely used mathematical software platform for scientists and engineers. I once told him how important our Connectionist Models Summer School was for jump-starting neural networks. Stephen inaugurated an Annual Wolfram Summer School and the 24th took place in 2026.

Physics in the 20th century was built on the mathematics of differential equations; in the 21st century, complex systems in computer science instead use computational algorithms. Stephen explored cellular automata invented by John von Neumann. Like neural networks, cellular automata



Fig. 8. A frame from Stephen Wolfram's "The Long Story of How Neural Nets Got to Where They Are: A Conversation with Terry Sejnowski," YouTube, February 14, 2023, www.youtube.com/watch?v=XKC-4Tosdd8.

have simple processing units and update rules. He showed that even the simplest update rules led to automata capable of universal computation. This task demonstrates that the complexity we see in nature could have evolved from simple rules. Neural network learning algorithms can create complex networks that serve as universal function approximators, which partly explains their widespread success.

Matchmaking

I received a call from Stephen when I was at Johns Hopkins. I was scheduled to give a talk at UC San Diego, and he wanted to let me know that a friend of his would be in the audience. He also called his friend, Beatrice Golomb (figure 9), to alert her to my talk. Beatrice was a medical doctor and doctor of philosophy student working on a thesis in visual psychophysics. We met and went to a group dinner. Beatrice wanted to do her thesis on neural networks, but the Biology Department would not allow it.

Several years later, Francis brought Beatrice to tea and told me that she wanted to do a postdoc on neural networks. I had a database of young college students' faces and wondered whether neural networks could discriminate between male and female faces without hair and secondary sexual characteristics. Beatrice took on the project. The network was more accurate in this task than the members in my lab. When she presented SexNet at the Neural Information Processing Systems Conference in 1991, she concluded that members of my lab should spend more time at discriminating sex.

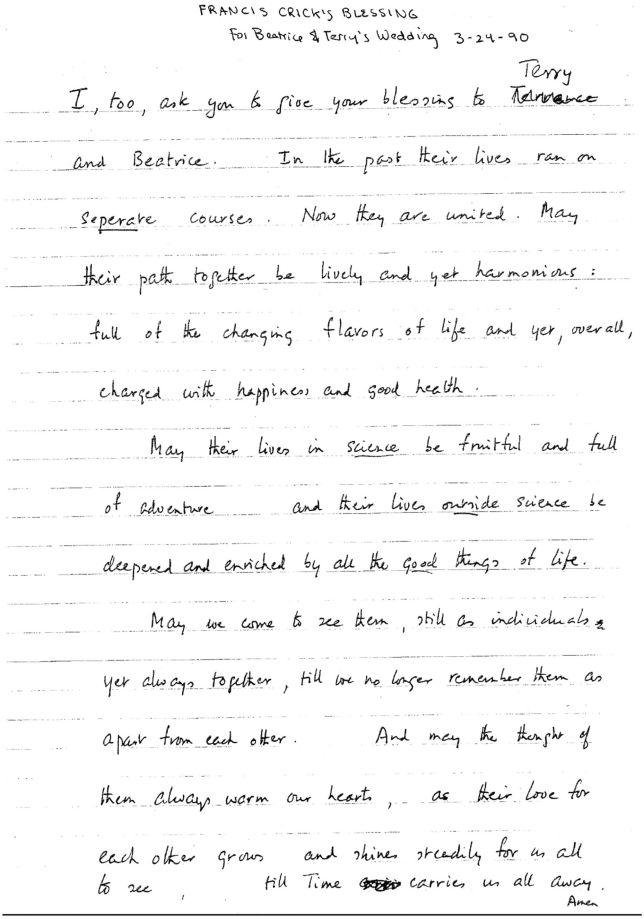
Beatrice and I became engaged. After our marriage at the Caltech Athenaeum in 1990, we went to the Beckman Auditorium for a marriage symposium, where Beatrice gave a talk ("Marriage: Theory and Practice")



Fig. 9. Terry and Beatrice visited Iceland with Cary and Marisela Staller in 2004

in her wedding gown. Stephen spoke with conviction and pride about how he had introduced us, but when Beatrice pointed out that if he was going to take credit, he also had to take responsibility, and with responsibility comes liability, he cautiously demurred.

Francis Crick knew my future wife, Beatrice Golomb, before I moved to La Jolla and gave a blessing when we were married at the Caltech Athenaeum in 1990 (figure 10). Beatrice is now a professor of medicine at UC San Diego and an accomplished medical detective. Along the way, we met Cary Staller, a member of the SUNY Board of Trustees, who became a lifelong hiking partner and a source of insight on dealing with real-world issues.



FRANCIS CRICK'S BLESSING
For Beatrice & Terry's Wedding 3-24-90

I, too, ask you to give your blessing to Terry and Beatrice. In the past their lives ran on seperate courses. Now they are united. May their path together be lively and yet harmonious: full of the changing flavors of life and yet, overall, charged with happiness and good health.

May their lives in science be fruitful and full of adventure and their lives outside science be deepened and enriched by all the good things of life.

May we come to see them, still as individuals & yet always together, till we no longer remember them as apart from each other. And may the thought of them always warm our hearts, as their love for each other grows and shines steadily for us all to see, till Time carries us all away.

Amen

Fig. 10. Francis Crick delivered this blessing at the Caltech Athenaeum, where Beatrice and I were married in 1990.

Synaptic Dynamics

Neurons may be the cellular units of brains, but synapses are the computational units. Fast sensorimotor control, cognition, and learning depend on synaptic signaling between neurons. Rapid advances in molecular biology would soon yield a comprehensive list of synaptic proteins, and biochemistry would provide insight into the dynamics of protein interactions. Tom Bartol, a research scientist in my lab, developed a computer program called MCell that uses Monte Carlo methods to integrate anatomy and biochemistry to track key molecules, including calcium and calcium-binding proteins, during synaptic activity (Coggan et al., 2005). This approach was developed by physicists to track neutrons in nuclear reactions, and we used it to model diffusion and molecular interactions in synapses.

MCell models are constrained by experimental data and validated by predictions. In the 1990s, computers could only simulate the initial steps of neurotransmitter binding to receptors and the subsequent synaptic potentials. It is now possible to model the much more complex biochemical interactions that lead to long-term synaptic plasticity. Simpler synapse models with a wide range of timescales for use in network models can be distilled from these computationally intensive models (Garcia, Bartol, and Sejnowski, 2022).

We embarked on a long-term project to model synaptic interactions at the single-molecule level in collaboration with Kristen Harris at University of Texas, Austin, a pioneer in reconstructing synapse structure at the electron microscopy level, and Mary Kennedy (see volume 11) at Caltech, who discovered important synaptic molecules and elucidated their biochemical interactions. With Mary Kennedy, we made progress in following the activation of CaMKII by phosphorylation following calcium entry through NMDA receptors, which triggers long-term potentiation (Bartol et al., 2015). Our simulations explain why large synapses are stable and the more numerous small synapses are labile. With Kristen Harris, we were the first to measure the precision of synaptic plasticity (Bartol et al., 2015).

The Computational Brain

In 1992, Patricia Churchland (see volume 13) and I published *The Computational Brain* with MIT Press, which shifted the focus from the traditional emphasis on the response properties of single neurons to distributed representations in large populations of neurons. Pat is a renegade philosopher who wrote an influential book on *Neurophilosophy*. She believed that brains and evolution might have something to do with the complex problems that philosophers ponder. It was heresy to think that how brains acquire knowledge might have something to do with epistemology. Our book was aimed at two audiences: neuroscientists who were curious about insight from neural network models and researchers with no background in neuroscience who wanted to learn more about the brain

through the lens of neural network models. It is still in print 34 years later, and we have been gratified to hear from many people over the years who said it had a seminal influence on their careers. The Levels of Investigation (figure 11) first appeared in a *Perspective on Cognitive Neuroscience* (Churchland and Sejnowski, 1988) and was reprinted in *The Computational Brain*. Most labs focus on one or two of these levels. Our goal was to build bridges between all of them.

Gain Fields in Parietal Cortex

Richard Anderson, then at the Salk Institute, and David Zipser at UC San Diego published an influential paper on a neural network model of neurons performing spatial coordinate transformation in the posterior parietal cortex (Zipser and Andersen, 1988). They combined visual inputs in retina

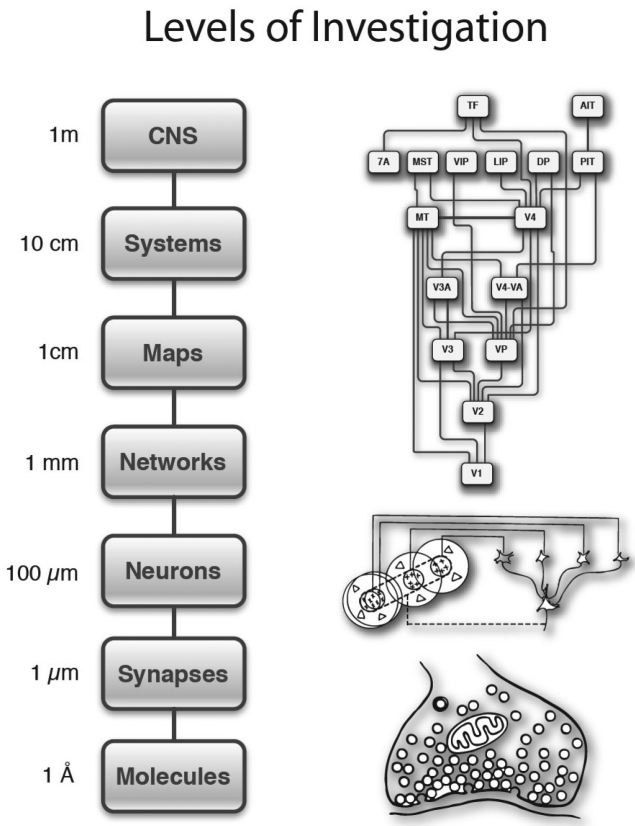


Fig. 11. Brains can be studied at many different levels of investigation. Most researchers focus on one or two levels. The CNL worked across all these levels, seeking computational principles that could bridge them and help us understand brain function.

coordinates with eye position and trained a network to transform them into head coordinates in the output layer. The units in the hidden layer developed gain fields like those seen in the parietal cortex: Their responses in retinal coordinates were modulated by eye position. The projections of these hidden units to the output units were in head coordinates. There were objections that the brain does not use backpropagation learning. The claim, however, was only that the observed gain fields in the parietal cortex were sufficient to perform the coordinate transformation.

Alex Pouget, a new graduate student in my lab, took this a step further and showed that the same transformation could be accomplished with hidden units that were joint basis functions of the inputs. Furthermore, he could replicate in his model a puzzling neurological finding that lesions of the right parietal cortex in humans led to neglect of the left side of the visual field, including object-centered hemineglect (Pouget and Sejnowski, 2001). The neuropsychological literature is filled with symptoms from brain injury that are difficult to explain. Lesions to neural networks with internal representations for tasks performed in a brain area could reproduce many of these symptoms (Hinton, Plaut, and Shallice, 2013).

Independent Component Analysis

The 1990s were a golden decade for CNL. Tony Bell and I developed an unsupervised learning algorithm called independent component analysis (ICA) that could blindly separate sources from mixtures of independent signals, such as separating speech from background music (Bell and Sejnowski, 1995). Remarkably, when we trained a network on patches from natural images, the independent components resembled the response properties of oriented simple cells in the visual cortex discovered by Hubel and Wiesel (figure 12). A small number of the independent components are needed

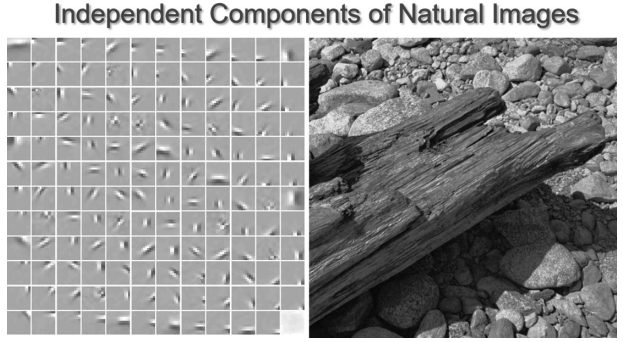


Fig. 12. ICA was given small 12×12 patches of the natural image on the right and produced 144 independent components on the left that resembled the oriented simple cells found in the primary visual cortex. A small number of weighted independent components could reproduce each patch.

to represent a patch from the image. Bruno Olshausen and David Field generated similar features starting with a sparsity assumption. Science advances when evidence converges from many different approaches. This confirmed Horace Barlow's conjecture that the visual cortex evolved to represent natural images efficiently.

ICA has many other uses and is now routinely used to analyze neural recordings and brain-imaging data. SoftMax, an ICA startup for noise cancellation applications founded by Te-Won Lee and Tzyy-Ping Jung, former postdoctoral fellows in my lab, was acquired by Qualcomm. Most cellphones now have ICA inside.

My successes with neural networks were based on abstractions of neurons that processed inputs with a single timescale, synapses with a single synaptic weight, and linear summation of inputs. Francis Crick once wrote in *Nature* that although the small neural networks we could train were far from realistic brain models, they demonstrated that it was possible to compute with brain-like networks (Crick, 1989). It is surprising what such simple processing units can compute, especially now that we have scaled them up to networks with hundreds of billions of units and trillions of weights.

I wanted to know what else could be computed with more complex synapses and dynamics. I did not have to depend entirely on other experimental labs because I had a wet lab to test models and explore brain dynamics. We also sharpened our intuition with *Neuron*, a computer program developed by Michael Hines and Ted Carnevale that simulated the electrophysiological properties of dendritic trees in neurons receiving synaptic inputs, based on the Hodgkin-Huxley equations for ion channels. We used these simulations to guide and interpret our experiments (Mainen, Joerges, Huguenard, and Sejnowski, 1995).

Computing with Spikes

There has been a long-lasting debate about whether the cortex uses a rate code—the average spike frequency—or depends on the timing of single spikes. Zach Mainen, a graduate student in my lab, performed *in vitro* experiments recording from cortical neurons and demonstrated that the timing of spikes was accurate within a few milliseconds when the same fluctuating pattern observed in *in vivo* recordings (called frozen noise) was injected repeatedly (Mainen and Sejnowski, 1995). We further used spike-triggered averaging to show that spikes followed strong hyperpolarization. The membrane potentials of neurons recorded *in vivo* rapidly fluctuate because thousands of synaptic inputs bombard them, and rapid hyperpolarization could be driven by inhibitory inputs from basket cells that contact the somas of pyramidal neurons. Analysis of *in vivo* recordings from thalamic neurons in response to fluctuating visual inputs, and

from the medial temporal (MT) motion area in response to fluctuating moving visual stimuli, also revealed millisecond precision (Buracas, Zador, DeWeese, and Albright, 1998; Reinagel and Reid, 2001). This exquisite precision in spike initiation would be unnecessary if the cortex used only a rate code (Tiesinga, Fellous, and Sejnowski, 2008).

Traveling Waves

Long before it was possible to record from single cortical neurons, scalp recordings and cortical field potentials revealed brain oscillations. Cortical 10 Hertz (Hz) alpha rhythms from the surface of the scalp were discovered by Hans Berger in 1924. Oscillations in frequency bands from 2 to 80 Hz have been reported throughout the cortex associated with attention, memory, and sleep. Their function remains elusive. Arrays of electrodes have revealed that all of these oscillations are sparse cortical traveling waves (Muller, Chavane, Reynolds, and Sejnowski, 2018). We developed spiking-neuron models with the same cortical connectivity pattern and axonal and synaptic time delays, which exhibited sparse traveling waves (Davis, Benigno, et al., 2021).

Most systems neuroscientists have ignored traveling waves despite their ubiquity, perhaps because we do not yet know their function. As spontaneous waves pass through area MT of a marmoset, the wave's phase modulates the responses of neurons to low-contrast stimuli near threshold, which is highly correlated with perception (Davis, Muller, et al., 2020). This reveals a link to behavior, but only near the threshold. Percepts above threshold can be attended to and modify behavior, but cognition does not change the percept itself. Although cognitive processing can influence behavior, it most often operates in parallel with sensorimotor processing (Muller, Chavane, Reynolds, and Sejnowski, 2018; Gepshtein, 2022).

STDP was discovered in the cortex in 1997 and depends on the relative timing of synaptic input and output spikes within the millisecond range, suggesting a possible role for precisely timed spikes (Markram et al., 1997). This raised the question, however, of how spike timing is controlled at the millisecond level in the cortex to meet the tight ±10 millisecond window for the relative timing of input and output spikes. The fluctuating membrane potential produces a highly irregular spike timing; however, the relative timing of spikes is strongly controlled by basket cells, each of which contacts the somas of thousands of pyramidal neurons with 10-15 synapses.

Traveling waves are ubiquitous in the cortex, but they have no known function. Waves generate highly structured spike patterns with distinct phases across the cortical surface. Unmyelinated long-range horizontal axons propagate the waves, some of which activate basket cells with inhibitory synapses on the somas of pyramidal cells, and produce

suppressive, nonclassical surrounds (Allman, Miezin, and McGuinness, 1985). The waves also mix delayed information from neighboring spatial locations, creating a new spacetime code (Muller, Churchland, and Sejnowski, 2024). Waves could also stimulate a backpropagating action potential through rebound inhibition, which strongly depolarizes the dendrites and triggers STDP at dendritic synapses recently activated by the traveling wave. STDP lasting hours could support long-term working memory (Ericsson and Kintsch, 1995) on a second tier of cognitive processing (Sejnowski, 2026). A new dynamic picture of global cortical activity is emerging.

Computational Models of Sleep States

I gave a talk at a New York University meeting when 40 Hz cortical oscillations were a hot topic. Wolf Singer (see volume 9) and Charlie Gray had recently reported 40 Hz oscillations in the visual cortex of cats, which was unexpected. My talk addressed the mechanisms that generated the 40 Hz oscillations. Bill Lytton, a medically qualified postdoctoral fellow, and I simulated synaptic inputs in a detailed compartmental reconstruction of a cortical pyramidal neuron with Hodgkin-Huxley ion channels. We asked whether 40 Hz excitatory or inhibitory synaptic inputs were more effective in driving the pyramidal cell. The somas of pyramidal neurons are encrusted with inhibitory synapses from basket cells. Simulations of 40 Hz inhibitory inputs on the soma were far more effective at driving pyramidal neurons than excitatory inputs on dendrites, both because of spike rebound from inhibition and their proximity to the spike-initiating zone on the axon hillock.

I received a call from Mircea Steriade at Laval shortly after the meeting, asking if I was interested in working with him on the biophysical mechanisms underlying the generation of sleep spindles in the thalamus, which receive strong inhibitory feedback from the reticular nucleus of the thalamus. Sleep spindles are 10–14 bursts of spikes that last 1–2 seconds and support the consolidation of long-term memories in the cortex. I became fascinated by sleep and collaborated with Mircea for 15 years, during which time we published 30 joint papers. In 1993, Mircea, David McCormick, and I wrote a paper in *Science* on "Thalamocortical Oscillations in the Sleeping and Aroused Brain" that has been cited more than 4,600 times, and still receives 100 citations annually, more than 30 years later. Several students and postdoctoral fellows in my lab contributed to these studies, including Alain Destexhe, Maxim Bazhenov, Arthur Houweling, Maxime Bonjean, and Flavio Froelich. Alain and I wrote a monograph on "Thalamocortical Assemblies" in 2001 that integrated all the results of our biophysical modeling of sleep states; a second edition included new and unexpected discoveries (Destexhe and Sejnowski, 2023).

Sleep Spindles

In 2016, Lyle Muller, a postdoctoral fellow, and I collaborated with Sydney Cash at Massachusetts General Hospital to study large-scale recordings directly from the human cortex during sleep (Muller et al., 2016). This was possible in drug-resistant epilepsy patients who had electrode arrays implanted for several weeks to find the sites of seizure onsets. We studied sleep spindles, which are involved in consolidating long-term memories of experiences during the day. We discovered coherent phase differences across the cortex organized as global circular traveling waves. We modeled these waves with an array of coupled oscillators with delays proportional to the distances between cortical regions (Budzinski et al., 2023). The action potentials in myelinated long-distance cortical association axons travel at around 1 meter/second. As a consequence, the wave arrives at a patch of cortex around the same time as the spindle originating in the thalamus. The timing of the 10–14 Hz spindle bursts lasting 1–2 seconds could therefore trigger STDP.

Hippocampal Replay

Hippocampal neurons have place fields in an environment, and experiences are a series of place fields as a rat explores it. The same sequences are played back during sleep. In collaboration with Bruce McNaughton, Kechen Zhang in my lab devised a spike-based Bayesian reconstruction of the path directly from the spike sequences recorded during exploration (Zhang, Ginzburg, McNaughton, and Sejnowski, 1998), which was more accurate than simply spike-averaging the place fields. The hippocampus receives sequences of inputs from the cortex during the day and replays them to the cortex during slow-wave sleep, a slow process that takes weeks for memories resident in the cortex to no longer depend on the hippocampus for retrieval (Buzsáki, 2019).

Yusi Chen, a postdoc in my lab, trained a recurrent neural network to remember and play back sequences of inputs (Chen, Zhang, Cameron, and Sejnowski, 2024). The model was a predictive autoencoder that predicted the next input. The recurrent network was a model for CA3, which projects to CA1, which directly receives the next input, and subtracts it from the recurrent network's prediction. As learning progressed, the error approached zero, as observed over many weeks when CA1 neurons were recorded in the same environment. Interpreting place fields as prediction errors is a theme that has appeared previously in the interpretation of dopamine neurons' transient responses as reward-prediction errors. The weights in the trained network had a Mexican-hat topology, with local excitatory weights and inhibitory surrounds, similar to cortical connectivity. The dynamical patterns of activity in the recurrent network were analyzed

with ICA and principal component analysis (PCA), which revealed clusters and dynamical trajectories, respectively. It is amazing what you can see inside the black box just by looking with the right analytical tools.

Similar insights were gained by modeling cortical circuits trained to perform delayed match-to-sample (Kim and Sejnowski, 2021), a gating model of the prefrontal cortex trained to perform the Wisconsin Card Sorting Task (Tsuda, Tye, Siegelmann, and Sejnowski, 2020), and a network with continuous neuromodulatory context (Tsuda et al., 2026). By constraining the architecture of the network, for example, by using separate pools of excitatory and inhibitory neurons, it was possible to compare the trained network directly with cortical recordings. Inhibitory neurons were unexpectedly important in the working memory task.

Psychosis and Interneurons

The brain is filled with oscillations at many different frequencies (Buzsáki, 2019). I was particularly intrigued by gamma oscillations in the 30–80 Hz frequency band. Abnormal gamma waves are linked to thought disorders. Schizophrenia patients have reduced amplitude gamma waves, which are associated with symptoms that include delusions and disorganized thinking. Similar symptoms appear after serial use of a party drug, Special K. After several days of raving under the influence of Special K, they present at emergency rooms with symptoms indistinguishable from a schizophrenic break. Fortunately, the symptoms go away after a few days. Special K is ketamine, an N-methyl-D-aspartate (NMDA) channel blocker and an anesthetic at high doses. At low doses, ketamine induces hallucinations and dissociative out-of-body experiences (Gil-da-Costa et al., 2013; Beck et al., 2020).

Ketamine downregulates the enzyme that synthesizes the inhibitory neurotransmitter gamma-aminobutyric acid (GABA) in basket cells. I recruited Marga Behrens to my lab as a staff scientist to work on a mouse model of human psychosis. Marga had already traced the biochemical pathways underlying the downregulation of inhibitory feedback (Behrens and Sejnowski, 2009). Monica Jadi, a postdoctoral fellow in my lab, developed a model of cortical gamma oscillations generated by a feedback loop between pyramidal neurons and inhibitory basket cells (Jadi and Sejnowski, 2014). Reduced inhibition unbalances cortical circuits, making them hyperactive, and also reduces gamma waves, as in schizophrenia patients (Jadi, Behrens, and Sejnowski, 2016). Ketamine is effective at relieving clinical depression, perhaps by boosting reduced cortical activity to normal levels. Additional experiments are needed to confirm a causal role for reduced gamma waves in psychosis.

We wondered why it took several days for ketamine's action to reverse. We thought it might be due to an epigenetic mechanism, such as DNA

methylation, which can turn off gene transcription. We teamed up with Joe Ecker, a plant biologist at Salk who is a world expert on DNA methylation. This led to the first epigenetic sequencing of the mouse and human genome across development (Lister et al., 2013). We discovered that DNA was rapidly methylated early in development, paralleling synaptogenesis. But unlike normal CG methylation of cytosine, this methylation was non-CG, a feature otherwise found only in stem cells. This did not yield new insight into the downregulation of GABA synthesis, but it did eventually lead to a major BRAIN Initiative project on a single-cell, high-resolution epigenomic atlas of the brain.

Reward Prediction Error

Peter Dayan and Read Montague, postdoctoral fellows in my lab, developed a computational explanation for the responses of midbrain dopamine neurons. Although these neuromodulatory neurons were known to be part of the reward system, what dopamine represented was obscure. We predicted that the transient responses in dopamine neurons reflected not the reward itself but reward prediction error, and that a reinforcement learning algorithm called *temporal difference learning* was implemented in brains (Montague, Dayan, and Sejnowski, 1996). This paper was published after two years of resistance from reviewers. Our hypothesis was subsequently confirmed by recordings from dopamine cells in monkeys by Wolfram Schultz (see volume 12) and by functional magnetic resonance imaging in humans (Schultz, Dayan, and Montague, 1997). This theory revolutionized our understanding of how humans respond to rewards, risks, and temporal discounting in making cognitive decisions. In economics, it is often assumed that humans optimize utility, but the decisions they make through reinforcement learning are more complex. Neuroeconomics, a new field, is built on a better foundation for understanding human economic decisions.

Facial Expressions

After our success with SexNet, we collaborated with Paul Ekman at UC San Francisco, an expert on facial expressions. There are 43 facial muscles, and Paul developed the Facial Action Coding System (FACS), which FACS coders use to annotate video frames. Each emotion activates a different set of muscles, making it an objective way to identify them, including fleeting microexpressions. Coding each video frame is labor-intensive. Javier Movellan and Marni Bartlett-Stewart at the INC, which I founded in 1990, used neural networks to win a CIA facial expression "bake-off" competition with Carnegie Mellon University researchers. Their start-up, Emotient, was bought by Apple in 2016. Marni, a former

graduate student, was the first to apply ICA to faces, whose independent components were parts of faces.

Distributed Control

Early network models of the vestibulo-ocular reflex, with Steve Lisberger (see volume 12), explored multiple sites of synaptic plasticity in the vestibular nucleus and the cerebellum. In a *Nature* paper, we showed that an alternative was a change in the time constant of vestibular inputs to the cerebellum, using a feedback loop from the vestibular nucleus to the cerebellum as a neural integrator (Lisberger and Sejnowski, 1992). Positive feedback loops are ubiquitous in brains, and dynamics depend on delays and time constants as well as synaptic strengths.

Unlike engineered systems with central control, biological control is highly distributed throughout the brain. In collaboration with John Doyle, a control theorist at Caltech, we modeled layered control, including a reflex layer in the spinal cord and a planning layer in the visual cortex (Nakahira, Liu, Sejnowski, and Doyle, 2021; Nakahira et al., 2025). In another study, feedback of motor signals to early sensory areas was used to compensate for internal time delays by predicting self-generated sensory signals (Li, Sarma, Sejnowski, and Doyle, 2023). Motor signals in the visual cortex have been observed in several labs, accounting for "spontaneous activity" previously considered noise. Control theory is giving us insights into neural signals that would otherwise be difficult to interpret.

Community Building

Helmholtz Club

In 1983, Francis Crick, V. S. Ramachandran, and Gordon Shaw founded the Helmholtz Club, a group of researchers from Southern California primarily working on vision (Aicardi, 2014). The club met monthly at UC Irvine, about an hour drive from San Diego, Caltech, UCLA, and USC. The members met for lunch, followed by two 90-minute talks, one with an invited speaker and the second with a local speaker. The session was followed by dinner. When I moved to La Jolla in 1989, I became what is called the secretary in England. I arranged the speakers, broadening the scope to include related fields such as learning and memory. Speakers were barely past their first slide when a group discussion would ensue, with interesting insight. These sessions were a master class in how to think about scientific problems. Francis would often lead the conversation, the master of the master class. One of our guest speakers told me at dinner afterwards that it was refreshing to be asked questions by researchers who really wanted to know the answers.

NeurIPS

Neural networks were a hot topic in the 1990s, and new conferences were spawned, journals were founded (including *Neural Computation*, which I founded at the MIT Press), and researchers in many fields of science, engineering, and mathematics made early contributions. The annual Neural Information Processing Systems (NeurIPS) Conference, founded by Ed Posner at Caltech in 1987, attracted researchers from all these areas. When Ed died in a tragic bicycle accident in 1993, I became the president of the NeurIPS Foundation that oversees the conference and continue to do so. HHMI generously supported my lab manager, who also managed the conference, and a computer systems manager, who ran the conference's IT. We wrote scalable software for registration and scheduling that is still used today. NeurIPS grew steadily over the years, became a leading machine learning conference, and today is the leading AI conference. In 2025, NeurIPS attracted 28,000 attendees, eclipsing the annual meeting of the Society for Neuroscience. The generosity of HHMI indirectly supported the foundation of modern AI. When NeurIPS outgrew my lab, my team started a company that also manages other machine learning and computer vision conferences.

INC

Shortly after arriving in La Jolla, I founded the INC at UC San Diego, whose faculty came from a dozen departments. INC could also hire research faculty and compete for center grants. INC is still going strong after 35 years. We have a large and dynamic research community that covers all the relevant aspects of neuroscience, computer science, and mathematics. Departments, however, tend to be silos. Institutes bring together researchers in emerging research areas, and INC brought together researchers from all these departments to work on the problem of how brains compute. I was fortunate to come to a campus where the faculty was open to crossing disciplinary boundaries, especially between the brain and cognitive sciences. In 2026, the INC became part of a new School for Data Science and Computing at UC San Diego.

Science of Learning Center

One of our center grants in the INC was an NSF-sponsored 10-year exploration of the temporal dynamics of learning, a Science of Learning Center with $35 million funding. We used an unconventional approach to organize more than 100 researchers into a "Network of Networks," with time as a crosscutting research theme. Brains exhibit activity over many timescales, from milliseconds to years. No single investigator considers all timescales, but we could as a center. By exploring neurochemical and physiological temporal dynamics at the levels of higher-level perceptual and cognitive

processing, we were able to build bridges to the behavioral, linguistic, and even educational levels. These collaborative projects had a transformative impact on my career. I learned a lot by closely interacting with my colleagues, especially Gary Cottrell in computer science, Andrea Chiba in cognitive science, and Paula Tallal in neuroscience.

Crick-Jacobs Center for Theoretical andComputational Biology

Sydney Brenner introduced *C. elegans*, a roundworm, as a new model system in biology. The title of his 2002 Nobel Prize lecture was "Nature's Gift to Science," with a footnote that he was not referring to journals. Sydney was a senior fellow at the Crick-Jacobs Center for Theoretical and Computational Biology, which I direct at the Salk Institute. We had many dinners together when he visited La Jolla on his way from England to the Okinawa Institute for Science and Technology, where he was its first president (figure 13). Sydney shaped my perspective on many biological questions that went beyond brains. When Sydney's health failed, and he could no longer travel, I would visit him in Singapore. In 2017, I was Sydney's interlocutor on a series of dialogues that ranged over his long career. He requested an audience that was predominantly younger scientists, and one theme that ran through our discussion was that young scientists should be given greater independence earlier in their careers. Our dialogue was transcribed and published in 2018: *In the Spirit of Science: Lectures by Sydney Brenner on DNA, Worms and Brains.*

Learning How to Learn

In 2013, Barbara Oakley and I launched a free, wildly popular Massive Open Online Course (MOOC) titled "Learning How to Learn" (Oakley and Sejnowski, 2019) (figure 14). Barbara introduced me to the green screen,



Fig. 13. Terry with Sydney Brenner in La Jolla, 2008, around the time that he gave a series of three lectures on reading the human genome at the Salk Institute, delivered without a single slide or prop. He did not need props to hold an audience spellbound. He said that no human had ever read the entire human genome, base pair by base pair; only computers had.

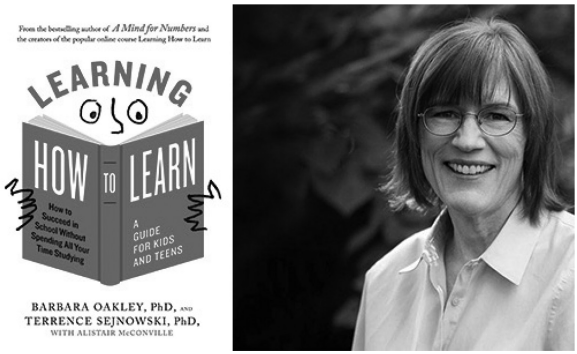


Fig. 14. Barbara Oakley, the mastermind behind "Learning How to Learn", the MOOC, and the book. It was the first of many MOOCs we launched together, including a series of three MOOCs for teachers.

which let us project backgrounds featuring giant neurons and moving zombies. This Coursera MOOC provides students with practical advice on improving their learning skills, based on our understanding of how the brain learns. Over the past 10 years, more than 6 million enthusiastic learners ages 10 to 90 have taken the course in 200 countries. In more than 40 years of classroom teaching, I have directly influenced a few thousand students. The internet has scaled education to the benefit of millions worldwide. Scaling is an important principle in biology. Evolution has vastly increased brain size, and recent advances in AI are similarly scaling up neural network models.

Cognitive Neuroscience

The James S. McDonnell Foundation and the Pew Charitable Trusts had a plan to jump-start the Centers for Cognitive Neuroscience in the 1990s. This new field would bring together cognitive psychologists and systems neuroscientists. They invited a group of researchers, including Patricia Churchland and me, to advise them on higher brain function. We traveled to major universities and interviewed leading researchers for their opinions. Eventually, six McDonnell-Pew Centers for Cognitive Neuroscience were established. I became the director of a McDonnell-Pew Center comprising the Salk Institute and UC San Diego. This brought together our neuroscience and cognitive science communities. Because the PDP group had already laid the foundation for brain-like cognitive processing, we had a head start in developing neural network models to study cognition. Cognitive Neuroscience also benefited from new developments in brain imaging, and I chaired a committee that established a Center for Functional MRI at UC San Diego in 2002. When the McDonnell-Pew Center ended, I led a 20-year training program in cognitive neuroscience funded by the National

Institutes of Health that supported a generation of graduate students and postdoctoral fellows. I was on the founding board of the Kavli Institute for Brain and Mind, which continued funding for innovative research.

Computational Neuroscience

In 1994, Hirsh Cohen at the Sloan Foundation began supporting five Centers for Theoretical Neuroscience. These centers jump-started the field of computational neuroscience, helped train a generation of postdoctoral fellows, and held annual center meetings. Most of the fellows were from physics, mathematics, and engineering, which helped them transition to neuroscience. After 10 years, Jerry Swartz, who founded Symbol Technologies, a pioneering barcode-scanning company, expanded the centers to 11. Jerry also funded a prize for computational and theoretical neuroscience under the Society for Neuroscience's aegis, thereby increasing its visibility. Jerry deserves recognition for his generous support over three decades. The Sloan-Swartz Centers anchored the computational neuroscience community, including several generations of postdoctoral fellows at the Salk Institute and UC San Diego. Jerry also funded the Swartz Center for Computational Neuroscience at INC directed by Scott Makeig, which explored brain function using electroencephalogram (EEG) and pioneered mobile high-definition EEG brain imaging. The EEGLAB software package for analyzing EEG recordings was developed at INC and is actively used by 15,000 EEG labs worldwide.

As Computational Neuroscience matured, neural models helped illuminate many brain functions, such as attractor states in the hippocampus and short-term working memory in the cortex. New conferences were established based on computational brain models, including the annual Computational Neuroscience Conference (CNS) and the Computational and Systems Neuroscience (COSYNE) meetings. The pathway from physics, mathematics, and engineering to neuroscience became a highway. A new generation of highly motivated young researchers is analyzing a cornucopia of experimental data to uncover new insight into brain function.

The BRAIN Initiative

When Barack Obama was looking for a decade-long project for his Grand Challenge in Science and Technology, the Kavli Foundation proposed scaling up recordings from neurons throughout the brain. I became involved early in this process when Myung Chun, the director of science at the Kavli Foundation, assembled a team to guide the political process in the Office of Science and Technology Policy. Our proposal ultimately became the BRAIN Initiative, announced by President Obama in 2013 (figure 15). The focus was on developing innovative neurotechnology to accelerate



Fig. 15. Terry with President Obama at the White House shortly before the BRAIN Initiative was announced in 2013.

advances in brain research, and several government agencies and foundations supported it, including NIH, NSF, and Defense Advanced Research Projects Agency (DARPA). As a member of the Advisory Committee to the NIH director on the BRAIN Initiative, I helped set the goals and milestones for computational, statistical, and theoretical brain research (BRAIN, 2025). Eve Marder (see volume 10) was a strong advocate. I also wrote a section on data sharing, which is essential for building computational models of brain circuits and systems. In the first five years of the BRAIN Initiative, most grants were awarded to teams that included an engineer, ensuring deliverable products. The progress over the past 10 years exceeded our committee's milestones and expectations.

NeuroAI

There was little overlap between AI and neuroscience in the 20th century because AI programs used symbols, rules, and logic that did not correspond with the signals flowing in neurons and neural circuits. This all changed in the 21st century when neural network models were scaled up with deep learning and made remarkable breakthroughs on many classic problems in AI, such as object recognition in images, speech recognition, and language translation (Sejnowski, 2018). Even more surprising was a new class of generative large language models that could converse with us (Sejnowski, 2024). There is now a two-way street between neuroscience and AI.

At the same time that AI was making breakthroughs, systems neuroscience advanced, enabling simultaneous recordings from tens of thousands of neurons across a dozen brain areas. New methods have been developed to analyze these recordings, and a new conceptual framework based on the counterintuitive mathematics of high-dimensional spaces is emerging. We are beginning to understand how sensorimotor behavior is generated in neural networks by analyzing the geometry of manifolds in

high-dimensional spaces. AI and neural networks share the same mathematics, accelerating progress through cross-pollination. Despite having different engineering and scientific goals, these two fields have benefited from this convergence. A new field, NeuroAI, has emerged from their overlap. These two revolutions in neuroscience and AI are advancing rapidly and leading us into a world once unimaginable. I was privileged to live through these remarkable events and contribute to both fields.

A Nobel Prize for Neural Networks

The 2024 Nobel Prize in Physics was awarded to John Hopfield and Geoffrey Hinton for their foundational research that laid the groundwork for modern AI (figure 16). I was fortunate to work with both of them. Geoff has had an enormous impact, and his scientific papers have received more than a million citations—I have had only 190,000. Geoff contributed half of his Nobel Prize to the NeurIPS Foundation to support the Sejnowski-Hinton Prize for a publication that made a significant computational contribution to understanding brain function over the past 10 years. The first award was given in 2025 to a group that showed how random feedback connections in the cortex could support backpropagation of errors (Lillicrap, Cownden, Tweed, and Akerman, 2016). This paper spurred many others who were inspired by their random alignment algorithm.

Research Horizon

I received a Pioneer Award from the NIH in 2025 to understand the neural mechanisms underlying long-term working memory, a new research direction. This is a high-risk, high-reward project that builds on my research dating back to the 1990s on spike timing (Sejnowski, 2025). In the 20th century, research was focused on fast sensorimotor tasks, but studying



Fig. 16. Physics Nobel laureates Geoffrey E. Hinton and John J. Hopfield at a Nobel Prize lecture at Stockholm University, Sweden, on Sunday, December 8, 2024.

long-term working memory lasting hours is more difficult. As you read this article, your eyes make fast, saccadic movements across the page, taking in small groups of words in your fovea three times per second. Each saccade is a snapshot that must be integrated with previous words to build a conceptual understanding of what is being conveyed. After reading this scientific autobiography, your brain will process it in the context of experiences and thoughts stored in your long-term memory, on a timescale of hours. Remembering the past and planning the future also depends on long-term working memory (Ericsson and Kintsch, 1995).

Research in my lab continues on the possibility that traveling waves trigger STDP to support long-term working memory lasting hours (Sejnowski, 2026). My doctoral thesis on a second tier of cortical processing based on correlated spike timing may have been on the right track (Sejnowski, 1978). There may be room for both rate and spike timing codes to explain cortical activity patterns and cortical function.

Understanding the neural mechanisms underlying long-term working memory could open new research directions for studying global cognitive processing in brains. At the Salk Institute, Francis Crick focused his research on the neural correlates of consciousness (figure 7). Once we understand the neural mechanisms underlying long-term working memory, the word "consciousness" will slowly fade away like the Cheshire cat in *Alice's Adventures in Wonderland*, leaving only a grin, as did the concept of "life essence" once the structure of DNA was revealed (Sejnowski, 2015). I started my career working on gravitational waves and am ending it working on brain waves—thinking about thinking (Sejnowski, 2025).

Epilogue

Neuroscience has changed enormously since I attended my first Society for Neuroscience (SfN) meeting 50 years ago in 1976. I was overwhelmed by the sheer scale of the enterprise: row after row of posters, session after session of talks, huge booths selling experimental equipment. There were very few modeling studies, and I tended to gravitate toward posters featuring Hodgkin-Huxley models. I was also drawn to systems neuroscience posters reporting intriguing properties of single neurons. The SfN meeting was like being a child in a candy shop.

At the 2025 Society for Neuroscience meeting in San Diego, I was struck by how the field has changed. Vernon Mountcastle (volume 6) once asked me what more could be learned by recording from an array of neurons that couldn't be found by recording from one neuron at a time. We now know the answer. Large-scale recordings are revealing new insight into how different parts of the brain work together. Connectomic reconstructions of the fly brain have fueled detailed models of fly vision and navigation. The SfN meeting was even more vibrant than I remembered it.

Neuroscience is still a young discipline, but this is changing rapidly. Methods have gotten much more sophisticated, and the datasets have mushroomed. One of the advantages of being a computational neuroscientist is the opportunity to conduct broad exploration through collaborations with experimental labs, which necessarily focus on single problems. In a close collaboration, I gain more than a successful project. I benefit from the deep insight of investigators who know much more than what is in the published papers. I am grateful to all of them, although only a few were explicitly mentioned here.

We are at the beginning of a new era in neuroscience, just as physics went from every boat on its own bottom to big teams tackling big problems. In 2015, the discovery paper on gravitational waves had more than 1,000 authors, and the first three were named Abbott. The International Brain Laboratory (IBL), funded by the Simons Foundation, brought together more than 20 labs to amass recordings from 100,000 neurons across the brains of mice doing the same task (Angelaki et al., 2025). They found that information about expectations and decision-making was widely distributed throughout the brain. IBL was founded by Zachary Mainen, Michael Hausser, and Alexandre Pouget. Zach and Alex were my first two graduate students at UC San Diego, and I am proud of what they have accomplished. I have attracted remarkable graduate students and postdoctoral fellows to my lab over the decades, many of whom are now senior leaders in computational neuroscience. They have many more discoveries ahead of them.

I have been offered several attractive jobs over the years. None could compete with the entrepreneurial spirit of the neuroscience research community in La Jolla. Salk and UC San Diego were founded in 1960, and I have been here for 35 years, more than half of their institutional lifetimes. My research has been generously supported by the HHMI, NIH grants for Collaborative Research in Computational Neuroscience (CRCNS), the NSF, the Keck Foundation, and the Office of Naval Research.

I have also been fortunate to have lab managers who have kept the CNL at Salk and the INC at UC San Diego afloat over this journey. Two, in particular, deserve special thanks: Rosemary Miller, who ran the NeurIPS Conference out of my lab when I moved to La Jolla, and Mary Ellen Perry (Figure 6), who oversaw its growth into what it is today. Marilee Bateman at UC San Diego helped us launch INC, which recently celebrated its 35th anniversary. My wife, Beatrice Golomb (Figure 9), has kept me in line. When I once began a talk by saying the brain is the most complex device in the known universe, she corrected me, saying it was the body, of which the brain is only one part. She is a medical doctor and doctor of philosophy, and, as always, was right.

John Hopfield had a seminal influence on many careers. He was a distinguished condensed-matter physicist who later moved to biophysics. His course in biophysics was different from the textbooks. He had original

insight into how oxygen binds to hemoglobin and how proofreading improves the accuracy of DNA transcription into RNA. John brought a physics perspective that illuminated many biological problems. I was fortunate that John took me under his wing when I transitioned to neuroscience. His model of nonlinearly interacting neurons, called the Hopfield network, profoundly influenced many physicists who followed a similar path into neuroscience.

Solomon Golomb, my father-in-law, was a mathematician and polymath who pioneered digital communications theory. With a doctorate in number theory, the purest field of pure math, he analyzed shift register sequences used in all cellphones today. He invented polyominoes, a game based on tiling the plane. He was also a biblical scholar. I relied on him for wise advice that combined uncommon sense with a deep understanding of human nature. He once told me that careers are made retrospectively. Looking back now, it all makes sense, but not at the time.

Bibliography

Ackley, D. H., Hinton, G. E., and Sejnowski, T. J. (1985). A learning algorithm for Boltzmann machines. *Cognitive Science*, 9, 147–169.

Aicardi, C. (2014). Of the Helmholtz Club, South-Californian seedbed for visual and cognitive neuroscience, and its patron Francis Crick. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 45, 1–11.

Allman, J., Miezin, F., and McGuinness, E. (1985). Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local–global comparisons in visual neurons. *Annual Review of Neuroscience*, 8, 407–430.

Angelaki, D., Benson, B., Benson, J., Birman, D., Bonacchi, N., Bougrova, K., … and Witten, I. B. (2025). A brain-wide map of neural activity during complex behaviour. *Nature*, 645(8079), 177–191.

Ballard, D. H., Hinton, G. E., and Sejnowski, T. J. (1983). Parallel visual computation. *Nature*, 306, 21–26.

Bartol, T. M. Jr., Bromer, C., Kinney, J. P., Chirillo, M. A., Bourne, J. N., Harris, K. M., and Sejnowski, T. J. (2015). Nanoconnectomic upper bound on the variability of synaptic plasticity. *eLife*, 4, e10778.

Bartol, T. M. Jr., Keller, D. X., Kinney, J. P., Bajaj, C., Harris, K. M., Sejnowski, T. J., and Kennedy, M. B. (2015). Computational reconstitution of spine calcium transients from individual proteins. *Frontiers in Synaptic Neuroscience*, 7, 17. doi: 10.3389/fnsyn.2015.00017.

Beck, K., Hindley, G., Borgan, F., Ginestet, C., McCutcheon, R., Brugger, S., Driesen, N., Ranganathan, M., D'Souza, D.C., Taylor, M., and Krystal, J. H. (2020). Association of ketamine with psychiatric symptoms and implications for its therapeutic use and for understanding schizophrenia: a systematic review and meta-analysis. *JAMA Network Open*, 3(5), p.e204693.

Behrens, M. M., and Sejnowski, T. J. (2009). Does schizophrenia arise from oxidative dysregulation of parvalbumin-interneurons in the developing cortex? *Neuropharmacology*, 57(3), 193–200.

Bell, A. J., and Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution, *Neural Computation*, 7, 1129–1159.

BRAIN (2025). A Scientific Vision. https://braininitiative.nih.gov/vision/nih-brain-initiative-reports/brain-2025-scientific-vision.

Budzinski, R. C., Nguyen, T. T., Benigno, G. B., Đoàn, J., Mináč, J., Sejnowski, T. J., and Muller, L. E. (2023). Analytical prediction of specific spatiotemporal patterns in nonlinear oscillator networks with distance-dependent time delays. *Physical Review Research*, 5(1), 013159.

Buracas, G. T., Zador, A. M., DeWeese, M. R., and Albright, T. D. (1998). Efficient discrimination of temporal patterns by motion-sensitive neurons in primate visual cortex. *Neuron*, 20(5), 959–969.

Buzsáki, G. (2019). *The Brain from Inside Out*. Oxford University Press.

Chen, Y., Zhang, H., Cameron, M., and Sejnowski, T. (2024). Predictive sequence learning in the hippocampal formation. *Neuron*, 112(15), 2645–2658.

Churchland, P. S., and Sejnowski, T. J. (1988). Perspectives on cognitive neuroscience. *Science* 242, 741–745.

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

Coggan, J. S., Bartol, T. M. Jr., Esquenazi, E. I., Stiles, J. R., Lamont, S., Martone, M. E., Berg, D. K., Ellisman, M. H., and Sejnowski, T. J. (2005). Evidence for Ectopic Neurotransmission at a Neuronal Synapse, *Science*, 39, 446-451.

Crick, F. (1989). The recent excitement about neural networks. *Nature*, 337(6203), 129–132.

Davis, Z. W., Benigno, G. B., Fletterman, C., Desbordes, T., Steward, C., Sejnowski, T. J., Reynolds, J. H., and Muller, L. (2021). Spontaneous traveling waves naturally emerge from horizontal fiber time delays and travel through locally asynchronous-irregular states. *Nature Communications*, 12(1), 6057.

Davis, Z. W., Muller, L., Martinez-Trujillo, J., Sejnowski, T., and Reynolds, J. H. (2020). Spontaneous travelling cortical waves gate perception in behaving primates. *Nature*, 587(7834), 432–436.

Destexhe, A., and Sejnowski, T. (2023). *Thalamocortical Assemblies: Sleep Spindles, Slow Waves and Epileptic Discharges*. Oxford University Press.

Ericsson K. A., and Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102(2), 211.

Garcia, J. W., Bartol, T. M., and Sejnowski, T. J. (2022). Multiscale modeling of presynaptic dynamics from molecular to mesoscale. *PLoS Computational Biology*, 18(5), p.e1010068.

Gepshtein, S., Pawar, A. S., Kwon, S., Savel'ev, S., and Albright, T. D. (2022). Spatially distributed computation in cortical circuits. *Science Advances*, 8(16).

Gil-da-Costa, R., Stoner, G.R., Fung, R., and Albright, T. D. (2013). Nonhuman primate model of schizophrenia using a noninvasive EEG method. *Proceedings of the National Academy of Sciences*, 110(38), 15425–15430.

Hinton, G. E., Plaut, D. C., and Shallice, T. (2013). Simulating brain damage. *Scientific American*, 76–82.

Hjellming, R. M., Andrews, M. H., and Sejnowski, T. J. (1969). A theoretical analysis of methods of interpreting radio-line data for H-II regions. *Astrophysical Journal*, 157, 573.

Jadi, M. P., Behrens, M. M., and Sejnowski, T. J. (2016). Abnormal gamma oscillations in N-methyl-D-aspartate receptor hypofunction models of schizophrenia. *Biological Psychiatry*, 79(9), 716–726.

Jadi, M. P., and Sejnowski, T. J. (2014). Cortical oscillations arise from contextual interactions that regulate sparse coding. *Proceedings of the National Academy of Sciences, USA*, 111, 6780–6785.

Josiah Macy Jr. Foundation. (1955). "The Conference Program." *Josiah Macy, Jr. Foundation 1930–1955: A Review of Activities.* Josiah Macy Jr. Foundation.

Kim, R., and Sejnowski, T. J. (2021). Strong inhibitory signaling underlies stable temporal dynamics and working memory in spiking neural networks. *Nature Neuroscience*, 24(1), 129–139.

Koller, D. P., Schirner, M., and Ritter, P. (2024). Human connectome topology directs cortical traveling waves and shapes frequency gradients. *Nature Communications*, 15(1), 3570.

Kuffler, S. W., and Sejnowski, T. J. (1983). Peptidergic and muscarinic excitation at amphibian sympathetic synapses. *Journal of Physiology*, 341, 257–278.

Li, J. S., Sarma, A. A., Sejnowski, T. J., and Doyle, J. C. (2023). Internal feedback in the cortical perception-action loop enables fast and accurate behavior. *Proceedings of the National Academy of Sciences, USA*, 120(39), e2300445120

Lillicrap, T. P., Cownden, D., Tweed, D. B., and Akerman, C. J. (2016). Random synaptic feedback weights support error backpropagation for deep learning. *Nature Communications*, 7(1), 13276.

Lisberger, S. G., and Sejnowski, T. J. (1992). Motor learning in a recurrent network model based on the vestibulo-ocular reflex. *Nature*, 360, 159–161.

Lister, R., Mukamel, E. A., Nery, J. R., Urich, M., Puddifoot, C. A., Johnson, N. D., Lucero, J., Huang, Y., Dwork, A., Schultz, M. D., Tonti-Filippini, J., Yu, M., Heyn, H., Hu, S., Wu, J. C., Rao, A., Esteller, M., He, C., Haghighi, F. G., Sejnowski, T. J., Behrens, M. M., and Ecker, J. R. (2013). Global epigenomic reconfiguration during mammalian brain development. *Science*, 341, 629.

Mainen, Z. F., Joerges, J., Huguenard, J. R., and Sejnowski, T. J. (1995). A model of spike initiation in neocortical pyramidal neurons. *Neuron*, 15(6), 1427–1439.

Mainen, Z. F., and Sejnowski, T. J. (1995). Reliability of spike timing in neocortical neurons. *Science*, 268(5216), 1503–1506.

Markram, H., Lübke, J., Frotscher, M., and Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science*, 275(5297), 213–215.

Marr, D., and Poggio, T. (1976). Cooperative computation of stereo disparity: A cooperative algorithm is derived for extracting disparity information from stereo image pairs. *Science*, 194(4262), 283–287.

Montague, P. R., Dayan, P., and Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience*, 16(5), 1936–1947.

Muller, L., Chavane, F., Reynolds, J. and Sejnowski, T. J. (2018). Cortical travelling waves: mechanisms and computational principles. *Nature Reviews Neuroscience*, 19(5), 255–268.

Muller, L., Churchland, P. S., and Sejnowski, T. J. (2024). Transformers and cortical waves: encoders for pulling in context across time. *Trends in Neurosciences*, 47(10), 788–802.

Muller, L., Piantoni, G., Koller, D., Cash, S. S., Halgren, E., and Sejnowski, T. J. (2016). Rotating waves during human sleep spindles organize global patterns of activity that repeat precisely through the night. *eLife*, 5, e17267.

Nakahira, Y., Liu, Q., Deng, X., Sejnowski, T. J., and Doyle, J. C. (2025) Diversity deconstrains speed-accuracy trade-offs in sensorimotor control. *Neural Computation*, 37, 1783–7180.

Nakahira, Y., Liu, Q., Sejnowski, T. J., and Doyle, J. C. (2021). Diversity-enabled sweet spots in layered architectures and speed-accuracy trade-offs in sensorimotor control. *Proceedings of the National Academy of Sciences USA*, 118, e1916367118.

Oakley, B. A., and Sejnowski, T. J. (2019). What we learned from creating one of the world's most popular MOOCs. *NPJ Science of Learning*, 4(1), 7.

Pouget, A., and Sejnowski, T. J. (2001). Simulating a lesion in a basis function model of spatial representations: comparison with hemineglect. *Psychological Review*, 108(3), 653.

Qian, N., and Sejnowski, T. J. (1988). Predicting the secondary structure of globular proteins using neural network models. *Journal of Molecular Biology*, 202(4), 865–884.

Reinagel, P., and Reid, R. C. (2002). Precise firing events are conserved across neurons. *Journal of Neuroscience*, 22(16), 6837–6841.

Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306), 1593–1599.

Sejnowski, T. J. (January 1974). Sources of gravity waves. *Physics Today*, 27, 40–48.

Sejnowski, T. J. (1976a). On global properties of neuronal interaction. *Biological Cybernetics*, 22, 85–95.

Sejnowski, T. J. (1976b). On the stochastic dynamics of neuronal interaction. *Biological Cybernetics*, 22, 203–211.

Sejnowski, T. J., (1977a). Statistical constraints on synaptic plasticity. *Journal of Theoretical Biology*, 69, 385–389.

Sejnowski, T. J. (1977b). Storing covariance with nonlinearly interacting neurons. *Journal of Mathematical Biology*, 4, 203–211.

Sejnowski, T. J. (1978). *A Stochastic Model of Nonlinearly Interacting Neurons*. PhD diss., Princeton University.

Sejnowski, T. J. (1981). Skeleton filters in the brain. In: Hinton, G. E., and Anderson, J. A. (Eds.), *Parallel Models of Associative Memory*, 189–212. Lawrence Erlbaum Associates.

Sejnowski, T. J. (1986). Open questions about computation in cerebral cortex. In: McClelland, J., and Rumelhart, D. (Eds.), *Explorations in the Microstructure of Cognition 2: Applications*, 372–389. MIT Press.

Sejnowski, T. J. (2015). Consciousness. *Daedalus* 144, 123–132.

Sejnowski, T. J. (2018). *The Deep Learning Revolution*. MIT press.

Sejnowski, T. J. (2024). *ChatGPT and the Future of AI*: *The Deep Language Revolution*. MIT Press.

Sejnowski, T. J. (2025) Thinking about thinking: AI offers theoretical insights into human memory. *The Transmitter*. www.thetransmitter.org/human-neurotechnology/thinking-about-thinking-ai-offerstheoretical-insights-into-human-memory.

Sejnowski, T. J. (2026). Dynamical mechanisms for coordinating long-term working memory based on the precision of spike-timing in cortical neurons. *Biological Cybernetics*, 120(3–4), 16.

Sejnowski, T. J., Reingold, S. C., Kelley, D. B., and Gelperin, A. (1980). Localization of [3H]-2-deoxyglucose in single molluscan neurones. *Nature*, 287, 449–451.

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

Sejnowski, T. J., and Yodlowski, M. L. (1982). A freeze-fracture study of the skate electroreceptor. *Journal of Neurocytology*, 11(6), 897–912.

Stanton, P. K., and Sejnowski, T. J. (1989). Associative long-term depression in the hippocampus induced by Hebbian covariance. *Nature*, 339(6221), 215–218.

Tesauro, G., and Sejnowski, T. J. (1989). A parallel network that learns to play backgammon. *Artificial Intelligence*, 39(3), 357–390.

Tiesinga, P., Fellous, J. M., and Sejnowski, T. J. (2008) Regulation of spike timing in visual cortical circuits, *Nature Reviews Neuroscience*, 9, 97–109.

Tsuda, B., Tye, K. M., Siegelmann, H. T., and Sejnowski, T. J. (2020). A modeling framework for adaptive lifelong learning with transfer and savings through gating in the prefrontal cortex. *Proceedings of the National Academy of Sciences*, 117(47), 29872–29882.

Tsuda, B., Pate, S. C., Tye, K. M., Siegelmann, H. T., and Sejnowski, T. J. (2026). Neuromodulators generate multiple context-relevant behaviors in recurrent neural networks. *Neural Computation*, 39.

Yuhas, B. P., Goldstein, M. H., and Sejnowski, T. J. (1989). Integration of acoustic and visual speech signals using neural networks. *IEEE Communications Magazine*, 27(11), 65–71.

Zhang, K., Ginzburg, I., McNaughton, B. L., and Sejnowski, T. J. (1998). Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells. *Journal of Neurophysiology*, 79(2), 1017–1044.

Zipser, D., and Andersen, R. A. (1988). A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature*, 331(6158), 679–684.