The History of Neuroscience in Autobiography
Volume 13

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

Robert Desimone

pp. 220–267

https://doi.org/10.1523/hon.013006

Robert Desimone

Robert Desimone

Born:

Norristown, Pennsylvania
June 12, 1952

Education:

Macalester College, St. Paul, BA Psychology (1974)
Princeton University, Princeton, PhD Psychology and Neuroscience (1979)

Appointments:

Staff Fellow, Laboratory of Neuropsychology, NIMH (1980–1987)
Research Psychologist (tenured), Laboratory of Neuropsychology (1987–2004)
Chief, Section on Behavioral Neurophysiology, Laboratory of Neuropsychology, NIMH (1992–2004)
Chief, Laboratory of Neuropsychology, NIMH (1997–2004)
Scientific Director and Head of Intramural Program, NIMH (1998–2004)
Director, McGovern Institute for Brain Research at MIT (2004–present)
Doris and Don Berkey Professor of Neuroscience at MIT (2004–present)

Honors and Awards (Selected):

Troland Prize, National Academy of Sciences (1990)
Golden Brain Award of the Minerva Foundation (1994)
Fellow, American Association for the Advancement of Science (1999)
Member, National Academy of Sciences (1999)
Member, American Academy of Arts and Sciences (2001)
Fellow, Association for Psychological Science (2007)
Helmholtz Prize of the International Neural Network Society (2010)
Patricia Goldman Rakic Prize in Cognitive Neuroscience of the BBRF (2020)
Kavli Lifetime Contribution Award of the Cognitive Neuroscience Society (2021)
Ralph E. Gerard Prize in Neuroscience, Society for Neuroscience (2021)

Robert Desimone has mapped the ventral stream for object recognition of the macaque cortex and initially established the properties of cells in many of the areas, including cells selective for the images of faces. He found that attention to a behaviorally relevant object in a cluttered field results in the filtering out of distracting information from the receptive fields of ventral stream cells. This attentional filtering results from feedback to these visual areas from cortical regions, such as the lateral prefrontal cortex, enabling people to focus on the task at hand. Attentional filtering is also accompanied by the synchronous activity of cells carrying the relevant information, increasing their impact on downstream cortical areas. He also studies the neural basis of psychiatric disorders, using nonhuman primate genetic models.

Robert Desimone

Beneath the Windmill: Family Background

I was born into a working class, Italian immigrant neighborhood in Conshohocken, Pennsylvania, in 1952. I traced my recent ancestors and their genes to Avellino Province, near Naples—a place that was also the origin of Tony Soprano, the television character, not the actor. Mom emerged from a lineage of formidable women, and she held our family together in the face of adversity. Grandma, just as steely, lived with us for a while, and I learned not to ruffle her feathers. One of my lasting memories is that when she became annoyed with my pet chicken, she lopped off its head and served it up for dinner.

Dad enlisted during World War II and returned with emotional scars. While in the army, he went through electroconvulsive therapy (ECT) after a "nervous breakdown." I am still not sure what that meant, but back then, ECT was a hammer, and every psychiatric disorder looked like a nail. I found on eBay one of the first commercial ECT machines from the World War II era and have it in my Massachusetts Institute of Technology (MIT) office as a reminder. I believe it was the fear of future psychiatric disorders in the family that first spurred my fascination with the mind and brain. It was a prescient fear.

After returning from the army, Dad was expected to work at Desimone Motors, my grandfather's used car business in Conshohocken, rather than attend college. My grandfather often boasted that he sold a car to the gangster Al Capone, but he was vague about the details. The lot was surprisingly small, its most distinguishing feature being the small Dutch windmill that served as the business office—a quirky marketing ploy that didn't exactly boost sales (*Figure 1*). Dad thought it was hilarious; I inherited his peculiar sense of humor, as well as his practical wisdom: "when someone comes onto the lot wanting to buy a pickup truck, don't insist on selling them a station wagon." I've applied this lesson when seeking funding from private supporters at MIT.

There were three of us kids; I was the eldest, followed by my sister and brother. Although Dad worked long hours in the car business and Mom worked part-time jobs in department stores to bring in extra money, we were poor—even by our neighbors' standards. Our house had indoor plumbing, but Mom was mortified that her family home in town had only an outhouse. My parents desired more money and education but rejected those who flaunted either. We did not have many luxuries, but I have happy memories of spending virtually every day with large numbers of other kids from the neighborhood, largely free from oversight. My father parked a shell of a junk car from the car lot at our house, and it served as a neighborhood spaceship, time machine, fort, or whatever else was needed in the moment. I often

Figure 1. Desimone Motors.

wonder if the epidemic of social media use today is not mostly a compensation for loss of the communal environment in which we humans evolved.

I was sent to a Catholic parochial grade school, where I discovered that I had a more focused state of attention than most other kids. When I was reading in class, I didn't hear any distractions, including the recess bell. Was it a coincidence that I ended up studying attention? One thing you couldn't ignore were the frequent duck-and-cover A-bomb drills, when we all hid under our desks in case the Soviets dropped a big one on Conshohocken. At the Catholic school, I learned to question anything that seemed unreasonable, and we were taught many unreasonable things in that school. In the third grade, the nuns told us there were only two mortal sins that God could not forgive when you went to confession: murder, and cheating on a test at school. Really? God put them in the same category? I decided to do the experiment on the next test, where I cheated on a question by looking at my neighbor's answer. I soon confessed my sin to the priest of the church attached to the school, where I expected him to sentence me to say a few

additional Hail Marys, the usual punishment for my sins. You can imagine my horror when he said, "son, there are only two mortal sins that God cannot forgive—murder, and cheating on a test." It is presumably still on my permanent record, and I never tried to replicate the experiment.

Defying his father, my dad eventually left the car business, and we moved to different cities frequently as he pursued various jobs. The frequent moving set back our education, especially for my siblings. I lied about my age to work several low-wage positions, including janitor, factory worker, waiter, substitute mailman, and taxi driver. The night shift at the taxi company in Norfolk, Virginia, was perilous at that time—tales of drivers being robbed or shot were common, and I had a few harrowing experiences myself. Everyone had a contingency plan. One older driver showed me the two guns he carried in his cab, which he justified as necessary for self-defense and for staging a believable scene should he shoot an unarmed assailant. Casual brutality was a way of life. Instead of carrying two guns, I took fighting lessons from a man who said he had been a street fighter in India. I was no John Wick, but to this day, if I ever found myself in a rough bar in a city holding an SfN convention, and you tapped me on the shoulder from behind, I would know what to do.

Beyond coping with our trying family circumstances, I was taken up with the same political-cultural issues as anyone else who was growing up in the late 1960s, in particular, the Vietnam War. Although I protested against the war, I felt guilty knowing that, unlike some of my high school classmates who would certainly be drafted, I would get a college deferment if I was accepted at a college. I decided to take my chances with the draft lottery at the time, which was based on date of birth. If you chose to go into the lottery, and the random number associated with your birthdate was lower than the cutoff that year, you were drafted regardless of college. The lottery results were reported in the newspaper, and I was profoundly relieved when my number was too high for a realistic risk of being drafted. A lucky number felt morally superior to a college deferment, and I was spared any trauma in the war.

Thinking of my father's nervous breakdown in World War II, I wanted to go to college to become a psychotherapist. The obvious path was to work to earn enough money to attend a local state college in Norfolk, while living at home. But my parents and I dreamed that I could somehow escape Norfolk. I read fiction voraciously as a child, and for many years, I believed that reading adventure and science fiction books was an escape from my childhood conditions. Isaac Asimov was one of my favorite science fiction authors, and I built my own telescopes from scraps, hoping to catch a glimpse of alien worlds. But now I wonder whether chronic reading of other worlds might also reflect an inherited personality trait, one that is poised to escape current conditions if needed. My father engaged in several famous fantasies of this sort over the years. In researching my ancestry, I found the passenger

lists of the ships that took my grandparents and great-grandparents to the United States around the turn of the last century, which revealed that they had no English, no money, and no skills. I can only imagine what personal characteristics gave them the courage to leave their homeland for an unknown world. They were probably not unlike many of the immigrants at our border today.

My own far less risky escape plan was realized when I won a National Merit Scholarship based on a standardized test that was administered in my school, and I was offered a full scholarship to Macalester College in St. Paul, Minnesota. Macalester had many attractions, one of which was its distance from Norfolk. Although standardized tests like the SAT are sometimes criticized today, the test I took served its intended purpose for me, which was to give a chance to a kid who would not have been eligible for a good college, let alone a scholarship, based on their education in schools with poor reputations. The scholarship was endowed by DeWitt Wallace, the founder of *Reader's Digest* magazine, and my mother was so thankful she kept a *Reader's Digest* subscription for the rest of her life.

Escape to Macalester

The early 1970s at Macalester College was an idyllic time for me, where I developed a love of psychology—or, as it might now be called, cognitive science. Macalester was located in the city of St. Paul, but the immediate surroundings felt more like a small, Midwest college town, with the broad Mississippi River just a short walk down the road. Although Macalester lacked the prestige of Stanford or Harvard, I often remind my children that what you learn and do in school matters far more than where you go. Macalester offered me incredible opportunities, but one life-altering chance slipped through my fingers: the invention of one of the first computer dating applications.

For the psychology department's holiday party one year, I devised an app using the only available computer at the time, an IBM machine that relied on punch cards. I have always been an early adopter of technology. I asked partygoers to complete a questionnaire in advance with items borrowed from a standard test of psychopathology like, "Do you believe that people are secretly planning to harm you?" To these I added a few of my own original questions. The app's algorithm successfully matched people for dates at the party based on their answers. The buzz after the party was so positive, I analyzed the data, including survey results, to determine which questions had the most impact. To my surprise, the most influential question was one of my own: "On a scale of 1 to 10, how likely are you to sleep with someone on the first date?" If only I'd had the foresight to launch a business. I am reminded of this whenever one of my graduating students at MIT talks about striking out on their own.

During my first year at Macalester, I participated in an eye-opening class project that required me to spend a month living on a ward at a state mental hospital. The facility housed older, chronic patients unresponsive to treatment, and the experience brought me into the realm of psychosis. Witnessing adults shuffling through the corridors, babbling incoherently, or sitting catatonic in a corner, I wondered what could have gone so terribly wrong with their brains. It started a lifelong commitment to help people like this. That state hospital was later closed, like so many others, and I wonder today what happened to those patients. I hope that they were eventually treated successfully with better medications, but I fear that they were put out on the street with other people experiencing homelessness. When I later served as a volunteer in a call-in crisis center, it became obvious to me that untreated mental illness, poverty, homelessness, drugs, and abuse were all-too-common ingredients in a stew of misery in the city.

My last encounter with mental health facilities in college came during my senior year when I squandered all of my funds for housing on a used Volkswagen (VW) Karmann Ghia, a poor man's sports car. The body was a work of art, but the VW engine had the power of a chipmunk. Used cars ran deep in our family. Taken by its beauty and naively hoping it would attract women, I quickly found myself broke. In desperation, I took a job as a live-in, nighttime emergency worker at a halfway house for young adults recently discharged from psychiatric wards. These residents, mostly stabilized with medications, needed help rebuilding their interrupted lives. The halfway house was in financial trouble and could no longer afford professional staff overnight. Consequently, I was left in charge of a large facility with an average of one suicide attempt per night.

My responsibilities included getting out of bed, calling an ambulance, and handling the crisis—tasks for which I was wholly unprepared. Yet, the residents took pity on me, shifting their suicide attempts to daytime hours when professionals were available. I had done nothing more than show up, looking young and incompetent, but the house director was so impressed by the significant behavioral change in the residents that she offered me a daytime counseling position upon my early graduation. Until that point, I had been torn between a career in brain research or psychotherapy. My experience as a counselor at the halfway house quickly revealed that talking with people all day wasn't my calling, and my best chance in helping people with mental disorders was through brain research. The director of the halfway house agreed. And so, a new chapter began.

Pontifical Cells at Princeton

In looking for graduate schools, I visited Charlie Gross's (see volume 6) lab at Princeton University, and I immediately fell in love with the lab. Everything about Charlie and the lab was unconventional, from his antiwar

stance to his work on the inferior temporal (IT) cortex, which was terra incognita at that time. Charlie's lab did basic research, but acquiring fundamental knowledge about the healthy brain seemed to be an essential first step before tackling mental illness. Although the department didn't have the money to support me as a graduate student, I was saved by another award, in this case a National Science Foundation (NSF) graduate fellowship. That award was based on a combination of grades plus performance on another standardized test, the Graduate Record Examination (GRE), which again succeeded in its purpose of providing opportunities because I did not have a conventionally strong research background in college.

I have already written some stories about Charlie and the people in his lab who had such an influence on me, including Dave Bender and Ricardo Gattass [4], and I won't repeat those stories here. Charlie valued independence most of all, which I learned on one of my first days in the lab. While cleaning up after an experiment, I accidentally dropped a monkey's hard contact lens down the drain of a deep sink in the housing room. I had never seen a monkey contact lens before, but it didn't seem like something that could be bought at Lens Crafters. I panicked and dismantled the plumbing to find it, not suspecting that the trap would be filled with truly disgusting things in addition to a lens. My resourcefulness made quite an impression on Charlie, and he told everyone afterward that he thought I would make it as a grad student. I'm not sure if I adopted Charlie's attitudes on independence in the lab or if Charlie's attitudes happened to fit my own existing personality traits, but I left his lab valuing independence and resourcefulness for myself, my lab, and everyone around me. Plumbing skills are a plus.

Although the face cells we found in Charlie's lab are what everyone remembers now, when I arrived in the lab the controversial discovery had been the first "hand cell" in the IT cortex—that is, a cell that responded best to the silhouette of a monkey's hand [6]. From what we now know, this cell was probably located in what is known as a body patch in the IT cortex. We didn't find our first convincing face cell until Charlie Bruce, a postdoc in the lab, and I were recording from the superior temporal polysensory area, or STP [1], an area adjacent to the IT that I had discovered when mapping the temporal cortex [7]. In trying to find a stimulus to drive the cell, Charlie Bruce's head passed in front of the projection screen, and the cell went crazy. Fortunately, I had previously set up a motor-drive contraption to sweep stimuli in front of the monkey, so that the stimuli could be presented under controlled conditions. At that time, the neuronal spike trains were recorded on analog tape rather than with a computer, and the trains of action potentials were photographed from the face of an oscilloscope, using a Grass Kymograph movie camera. We quickly started drawing faces on paper or putting face photos on cardboard to put on the mechanical arm, collecting data all the time. For some reason, we had a photo of the actor Al Pacino in the lab that the cell seemed to like very much. Many

people incorrectly assume that it was a picture of me. "Analyzing the data" meant developing long strips of film from the Kymograph, hanging them up to dry, and assembling them by stimulus.

STP was full of interesting cells, and the two Charlies and I had a blast studying many of them. We published a paper on the findings, including that first face cell, Al Pacino and all [1] (see *Figure 2*). The reactions on the grapevine started arriving like the unwanted reviews of critics after a bad Broadway play. People didn't believe it. While we were studying face perception in IT cortex, nearly the entire rest of the field was studying early visual cortex with bars and spots. The thinking was that with enough bar and spot detectors, you could represent practically any visual stimulus. David Hubel (see volume 1) visited the lab and told us that he thought IT cells were actually selective for very long bars (David later became an enthusiastic supporter of face cell research). I was feeling discouraged when, one day, I brightened up when a visiting Japanese neuroscientist told me that the STP face cell figure was the most famous scientific figure in all Japan. I practically fell off my chair. "People in Japan believe the

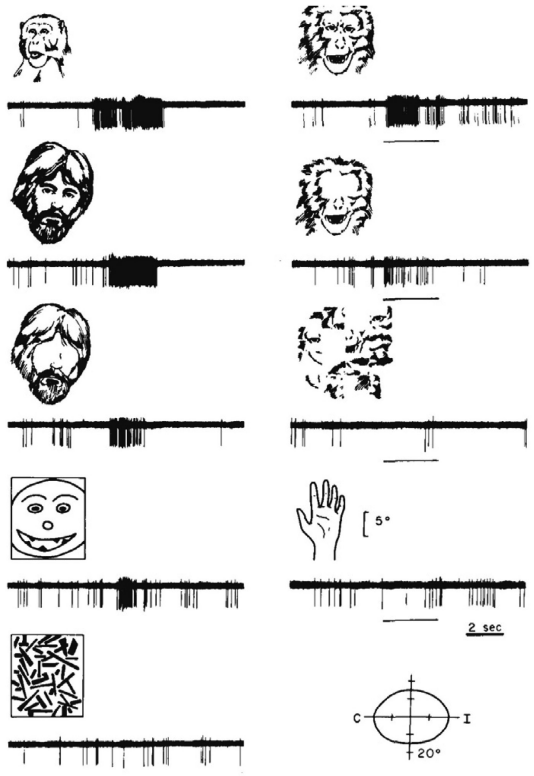

Figure 2. Adapted from [1].

face cell?" I asked. Maybe Japanese scientists were more open-minded. He looked horrified by my question. "Oh, NO," he responded emphatically. We had also bombed in Japan.

There were two general criticisms of face cells. The first implied that we were just incompetent. Nearly everyone who has had the experience of recording from cells in V1 has caused a cell to respond by accidentally putting your hand or face in front of the animal. With careful study, it usually turns out that the angle of your nose just happened to be the right angle for an oriented simple cell, not that the cell was selective for faces. Needless to say, we had the same worries, so we carefully ruled out alternative explanations for the face cells. The other criticism was that face cells implied a model for object recognition that many felt just couldn't be true. The disbelief may have arisen from Horace Barlow, who had imagined a hypothetical "grandmother cell" that would respond only to the image of your grandmother, or William James, who had earlier imagined a "pontifical cell," that would hold the experience of seeing a specific object. How could there be enough cells in the brain to have a cell dedicated to every percept or experience? Not many people imagined neural circuits for face perception in which your grandmother or the Pope would be coded by large populations of cells selective for general face features rather than individual cells dedicated to individual people. I certainly did not imagine that years later, the face cells and face patches of IT cortex would now be in nearly every neuroscience textbook.

Many people today don't realize how conservative the field was in the 1970s. In addition to face cells, another controversy was about whether spatial frequency analyses should be used to study early visual areas. As I mentioned earlier, at that time most of the field was firmly entrenched in the idea that the rich visual world was decomposed into bars, spots, and edges in areas, such as V1. Russ and Karen DeValois proposed, instead, that the visual cortex performed a local Fourier decomposition of visual scenes, and complex objects were reconstructed from their local spatial frequency components. No one grasped the idea that V1 cells operated as oriented, spatiotemporal filters that could be studied equally well with bars or gratings. The controversy was so heated that the SfN hosted a debate between the bar and edge advocates and the grating advocates, with Karl Pribram (see volume 2) serving as the moderator. Although I have enormous respect for Torsten Wiesel, he was one of the bar and edge leaders at that conservative time. Once, he visited Princeton, and after his inspiring talk, an assistant professor in the department tried to gently probe him about the spatial frequency approach. She asked him if he ever tried stimulating V1 cells with two bars, and Torsten was very excited to talk about how the cells treated double edges. Then she asked him if he ever tried three bars, and Torsten suddenly realized where the questions were headed. His response was that using gratings as stimuli would be like spitting in the street, something that he just couldn't bring himself to do.

I tried to balance studying face cells with less controversial studies. Tom Albright and I discovered the axis of motion columns in MT [8], and Julia Fleming and I used horseradish peroxidase (HRP), then a new anatomical tracing technique, to map the inputs to IT cortex [9]. With Eric Schwartz and Tom Albright, we conducted the first studies showing invariance for stimulus size and position by IT neurons [10]. For that study, we were finally able to give up photographing the oscilloscope face and instead collected the data with a PDP-12 computer, the cool new technology at the time. A green-colored monster, it stood six feet tall, weighed 900 pounds, and held 4K of memory, which was less memory than taken up by an emoji on my current phone. It generated a great deal of heat, and when it became freezing in Charlie's basement lab, someone would invariably call out to fire up the PDP-12. After searching for more than 30 years, I finally found a PDP-12 on eBay and have it on display at MIT, where it now generates more puzzlement than heat.

But with my doctorate in hand, the time in Charlie's lab came to a close. Mort Mishkin offered me a position at the National Institute of Mental Health (NIMH) once held by Pat Goldman-Rakic before her departure to Yale, and Richard Nakamura, then a postdoc in Mort's lab, led the recruitment. I had already been awarded an NIH postdoctoral fellowship to join Peter Schiller (see volume 7) at MIT. However, Peter generously encouraged me to accept the NIH opportunity, as it could pave the way to tenure. Charlie left me with a valuable piece of advice: focus on doing good science, not chasing tenure. The people who deserve tenure are the ones who don't need it.

Mapping the Cortex at NIMH

Mort's lab in the intramural program (IRP) of NIMH was in Building 9, a World War II–era building that was meant to be temporary, but it served as Mort's lab until a new research building was built to replace it next door. The building had a long history of primate research. Some older technicians told apocryphal stories of a time when chimpanzees (long gone from the building) shared cigarettes with the animal caretakers. Pat's former lab space had been commandeered by others in the lab, leaving only an undesirable room for me. The room contained a raised wooden floor that had collapsed with rot, a steel inner chamber for electrical shielding, and a two-foot gap that separated the inner chamber from the outer cinderblock walls. This gap was crammed with years' worth of neglected garbage, and light fixtures dangled precariously from the suspended ceiling.

Coming from Charlie's lab, I remained unshaken. After renovation and the removal of the rotten floor and trash, the room revealed itself to be far more spacious and pleasant than expected, with an improved aroma. Mort had secured a $25,000 "start-up package" for me, which barely covered

the CPU box of a PDP-11/34 computer, a successor to the PDP-12. By the following year, I could afford the necessary disk drives (5 Mb per drive!), and I scavenged most of the rest of what I needed from government surplus, which held everything from oscilloscopes to rocket launchers.

Mort Mishkin and Bob Wurtz (see volume 7) were really ahead of their time as lab heads at the NIH. Except for Mort and Bob's labs, most of NIH at that time operated as a kind of feudal system, with the lab heads directing all of the research, including research by several senior scientists beneath them. Although the position of a tenure-track junior scientist did not formally exist at that time at NIH, Mort treated me as one and let me do what I wanted. I initially continued to study face cells and other complex properties in IT cortex. I presented my work on face cells to the NIMH Board of Scientific Counselors (BSC), which evaluated all labs every four years, and whose recommendation was essential to keep your funding. The scientist from Harvard assigned to review me (not Hubel or Wiesel) told me privately that he didn't believe in the face cells, and even if he did, it would be the end of my career if I continued to study them. I published one final paper on face cells in IT cortex [11], which reported the properties of larger numbers of cells and received even more attention in the field, and then I switched to mapping, anatomy, and physiology of the mostly unexplored extrastriate cortex.

Initially, I was angry to receive the advice to drop the study of face cells, and I was greatly tempted to follow Charlie's advice to ignore tenure and concentrate on the science. But in retrospect, it was probably the right decision. At that time, finding a face cell in IT cortex was like finding a needle in a haystack. Without guidance from functional magnetic resonance imaging (fMRI), which was not available then, I could be recording for months before finding one. It was 13 years later when Nancy Kanwisher, Josh McDermott, and Marvin Chun found a face selective patch in IT cortex using fMRI in humans [12]. The definitive experiment in monkeys was then done by Doris Tsao, Winrich Freiwald, and Marge Livingstone (see volume 9), who localized a face patch in IT cortex using fMRI and then targeted it with electrodes for recordings [13]. The face patch was chock full of face-selective cells. Finally, there was a method for routinely locating them and studying their properties, but I had already moved on.

Leslie Ungerleider in Mort's lab was the ideal collaborator for extrastriate studies. Leslie's background was in primate behavior with Karl Pribram, but she later studied primate anatomy with Ted Jones. We mapped and/or studied the anatomy of areas MT [14–17], MST [14, 18, 19], FST [14, 19], V4t [14], TEO [20, 21], IT [22], and the posterior parietal cortex [22]. We didn't have computer programs to flatten the convoluted cortex to display the data then, so we built models of the brains out of wires and solder, and physically flattened them. They were literally wiring diagrams. Stan Schein and I also studied the properties of cells in Area V4, showing that this key

area of the ventral stream processed many types of information that was relevant for object recognition and not just color features, as others had claimed [23, 24]. By the conclusion of our studies and those of David Van Essen (see volume 9), Ricardo Gattass, Semir Zeki, John Allman, Jon Kaas (see volume 9), and others, we had at least a beginning sense of the organization and physiological properties of the dorsal and ventral streams for spatial and object vision [25, 26]. After a couple of years, Leslie and I decided to get married, and a couple of years later, we adopted a child, Matthew.

Mort's lab was the perfect environment for primate research. Just before I went to his lab, Mort had reported the first primate model for human temporal lobe amnesia, for which the patient HM was the prime example. After I joined the lab, Mort and Leslie Ungerleider published their extremely influential account of the dorsal and ventral streams for spatial and object vision, respectively, in primates. At that time, before optogenetics, the major causal method for studying brain function was the lesion method, and Mort was the legendary master of lesions. His lab grew into an intellectual epicenter for perception and memory research in the brain, with rising stars, such as Betsy Murray, Jocelyn Bachevalier, and John Aggleton, all part of the memory brain trust. Mort knew the history of the field, partly because he had worked with some of the historical figures such as Donald Hebb. Not surprisingly, a focal point for his lab was the lab library, lined with boxes that held thousands of reprints. These were the days before electronic journals. Many discussions with Mort resulted in a trip to the library reprints, which were indexed by cards held in a tall oak chest with narrow pullout drawers, similar to the ones in any large library. The system worked well, until one summer a student was assigned the job of organizing the cards alphabetically, which they did according to the first name of the first author on the papers.

The lab library was the scene of another amusing event, but first I should say something about Mort and Charlie (*Figure 3*). Both were great mentors because they loved to discuss ideas with their mentees, and both had a very high regard for each other. But their styles and personalities were very different. Charlie was iconoclastic and rejected formality and many social conventions. My memories of him are in cutoff corduroy shorts, Birkenstock sandals, and long beard. Mort was very approachable, but he also dressed and acted a little "stuffy." I don't think I ever saw him in shorts. I would say he was easily shocked, most memorably at a birthday party held for me in the lab library. On that day, we were sitting around a table eating cake, and a woman showed up from NIH procurement, to chastise me in front of the lab for having filled out the wrong form for an order. After chewing me out, the woman reached behind her back and pulled down the long zipper of her shift dress, which fell to the floor. Anticipating his reaction, everyone turned to Mort, and the horrified look on his face was priceless. Underneath, she was dressed as Wonder Woman, and she burst into song. She worked for a

Figure 3. Mort Mishkin, Bob Desimone, and Charlie Gross at a Festschrift for Mort Mishkin in 2016.

service and had been paid by Leslie. That stunt worked better on Mort than it would have on Charlie, who would have been unfazed.

Mort was a great mentor to me, but he had a troubled relationship with one of his own mentors, Karl Pribram. Karl was a prominent neurosurgeon and primate neuropsychologist, who ironically had a finger torn off by Washoe, the famous chimpanzee that spoke in sign language. It wasn't clear what message Washoe was trying to send Karl. Karl taught Mort neurosurgery, but some of his ideas had been disproven by Mort, and he resented Mort's success. One day, Karl visited the lab at NIMH and offered to give a talk in a seminar room. Mort asked if he had slides to show, which in those pre-PowerPoint days were actual two- by two-inch photographic slides loaded into a carousel in a Kodak projector. Karl declined, saying he would speak off the cuff. The lab members and I assembled for the talk, and after a few minutes, Karl said from the podium that he just remembered a slide he would like to show, and would Mort please take it from him and put it in the projector. Every minute or so, Karl would remember a new slide, and Mort would come up from the audience and take the slide to the projector at the back of the room. After a few slides, Mort asked Karl if he would put all of his slides together at once, so that he could load all of them into the carousel for the rest of the talk. Karl declined the offer, and for the rest of the talk utilized Mort as his personal slide boy. Washoe would have understood that message. Sometime later, Mort casually mentioned that junior scientists are not given independence by their mentor, they need to fight for it.

Paying Attention

I would love to say that I began my work on attention because I anticipated that a key architectural feature of the transformer networks (like GPT-4) of the future would be the attention layers. Unfortunately, the origin was more mundane. All of my recordings in the ventral stream in those early years were in anesthetized monkeys, and a major frustration was the variability of visual responses in areas like V4 and IT cortex. One moment there might be a strong response to the image of a face, and another moment there might be no response at all. I wondered whether this variability was due to the fact that the animal was anesthetized and not attending to the stimuli.

A puzzle with that explanation was that all of the work showing attentional modulation of cells in the monkey was in the superior colliculus or the dorsal stream cortex rather than the ventral stream. In fact, Bob Wurtz and Barry Richmond had searched for and failed to find positive attentional modulation of cells in IT cortex [27]. If attention was only relevant for the dorsal stream, its purpose might be limited to facilitating behavioral responses, such as eye and hand movements, rather than to facilitate object perception, which was more cognitive. I reasoned that understanding the mechanism of attention could be the gateway to understanding conscious awareness, which was more than a motor response. I thought an important clue to this puzzle might be that Bob and Barry had used a conventional diming task with simple spots and gratings when they failed to find positive effects of attention in IT cortex. Maybe it would be necessary to use a task and stimuli that engaged object recognition in the ventral stream.

Jeff Moran was my research assistant at the time, and he was looking for a thesis project to complete the doctorate he had started in graduate school. We therefore began to study the effects of attention in awake monkeys using a match to sample task with complex images, in V4 and IT cortex, which eventually became Jeff's thesis project. At the time, the field was split between the majority of labs working in anesthetized animals and a small minority of labs working in awake animals. If I had been an assistant professor in a university, it would probably not have been possible to make the switch. There would not have been the money to reequip the lab, and I probably would not have received a grant without having first trained in an awake monkey lab or at least having collected preliminary data. But the IRP was supposed to be a place to take risks, and Mort and the NIMH administration gave me the funding to buy the necessary equipment. Bob and Barry were encouraging and offered helpful advice on how to record from awake, behaving animals, which turned out to be less difficult that I feared.

When we started to record from IT cortex, we immediately faced a problem. All of the attentional studies to that point had studied attention with one stimulus placed inside the receptive field (RF) of the recorded cell and one stimulus outside, with the monkey holding gaze on a fixation spot.

One could then compare responses when the stimulus inside the RF was attended (the target) versus when the stimulus outside the RF was attended (distractor). The two stimuli were often identical. But the IT RFs were often so large it was difficult to place a stimulus outside the field. I was unable to do the intended experiment on the first cell recorded. To get around this problem, I decided to place inside the RF one "good" stimulus that could stimulate an IT neuron effectively on its own, and a "poor" stimulus, that elicited little or no response on its own. The good stimulus for one cell might be the image of a face and the poor stimulus might be a hand, for example. I reasoned that the poor stimulus would act like a stimulus outside the RF, since there would be little response. With two stimuli in the RF, the monkey was then cued to attend to one stimulus or the other.

I was very surprised to find that when the animal attended to the good stimulus in the RF the response was strong, but when it attended to the poor stimulus, the response was poor, even though both stimuli were located inside the RF at all times. If we reversed the locations of the stimuli in the RF, we got the same result. Thus, the effect of attention could be either to enhance or suppress the cell's response to the pair, depending on which stimulus was attended. Listening to the cell's response on a loudspeaker, I could hear the cell tracking the monkey's internal focus of attention as it shifted between stimuli. I felt that I was reading the monkey's mind. We proposed that the cells acted as though the effect of attention was to shrink the RF around the attended stimulus, excluding the unattended distractors, probably explaining why we had little awareness of unattended stimuli. This seemed a much more cognitive role for attention than how attention operated in the dorsal stream, where it seemed suited for facilitating motor responses. As I will explain later, I may have been too quick to dismiss any role for motor responses in attention in the ventral stream, but at least we had established the cognitive component.

We found similar results in V4. The paper reporting these results is my most cited empirical paper, and the results were much more readily accepted than my initial studies on face cells. The role of attention in the ventral stream eventually became a model system for studying a basic element of cognition and even conscious awareness. As a side note, the neuronal responses in awake, behaving animals were much more reliable than in anesthetized animals, so I made a permanent switch to awake animals. Jeff received his doctorate and went off to Pat Goldman-Rakic's lab for postdoctoral work.

One conclusion that I got wrong in that first study was that in area V4, it first appeared that there was no effect of attention on the response if only one of the competing stimuli was located in the RF. In retrospect, that was likely because the task was too easy for the animal and the stimuli were very high contrast. Shortly after that first study, Hedva Spitzer in my lab did find that attention could enhance the response to the RF stimulus under those conditions, if the task was made more difficult [28]. The effects of attention

were much smaller than when there were two competing stimuli in the RF, but they were nonetheless significant. John Reynolds in my lab also later showed that the magnifying effects of attention on the response to a single stimulus in the RF were stronger when that stimulus was lower in contrast, suggesting that the effects of attention were to enhance neuronal sensitivity to the attended RF stimulus and not just magnify responses [3, 29]. It wasn't until years later when I was at MIT that Ethan Meyers, working with Tommy Poggio (see volume 8) and myself, formalized how attention solved the problem of how objects can be decoded when embedded in clutter [5].

Biased Competition

I was very fortunate to have the British psychologist John Duncan conduct his sabbatical in my lab in the early 1990s. John was well known for a dispute with Anne Triesman (see volume 8). Anne believed that the purpose of attention was to bind together all the features of a single stimulus at a time (e.g., to bind the orientation of a bar with its color), whereas John believed that attention was just a type of bias that could shift visual processing from one stimulus or feature to another. We decided that the neurophysiological work in my lab supported John's views, and we collaborated together on a model of attention we termed biased competition, that combined psychological and neurophysiological arguments [30].

The biased competition idea received more attention, so to speak, than either of us could have imagined, and was widely cited. The basic idea was that because of the large size of RFs in the ventral stream, multiple objects in the visual field were typically located within these RFs, resulting in reduced information about any individual object. When John Reynolds was a postdoc in my lab, he had shown that, in the absence of attention, the response to two stimuli in the RF was close to the average of the response to either stimulus alone rather than the sum, which was evidence for competition [3] (see *Figure 4*).

Top-down feedback from structures important for attentional control, such as those in prefrontal cortex (PFC), biased that competition in favor of the most behaviorally relevant stimulus at the time. The bias was evident from the enhanced response to a single attended stimulus in the RF, and the effects on competition were evident by the effects of attention shifting the cell's response from one stimulus to another, when there were two competing stimuli in the RF. In a collaboration with Leslie Ungerleider and Sabine Kastner, then Leslie's postdoc, we found in human subjects results that were compatible with the animal work, including the influence of RF size [31, 32], using fMRI. Sabine then took a faculty position at Princeton, where she is one of very few people who successfully combine human and monkey work on attention.

Our biased competition account of attention offered an alternative to the idea that attention bound together objects or features—we saw feature binding as an indirect consequence of filtering distractors, and thereby

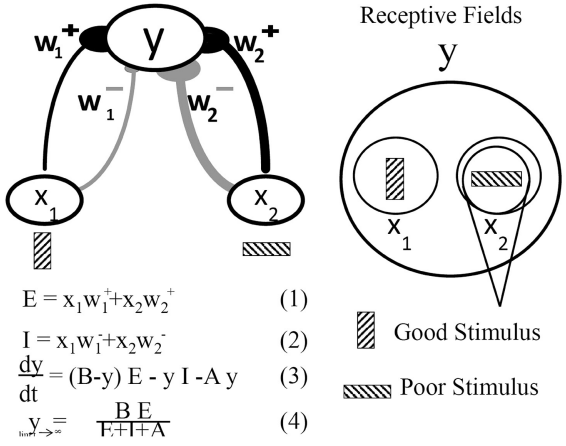

Figure 4. A quantitative account of biased competition. Left: The oval on top represents the neuron being recorded, whose firing rate is designated by the variable y. The two ovals below represent populations of "input" neurons that respond to the good and poor stimuli and that project to the upper neuron. The average responses of the input populations are designated x1 and x2. Black lines represent the excitatory projections from each input population to the measured cell, and gray lines indicate the inhibitory projections, which are assumed to depend on inhibitory interneurons (not shown in figure). The variables w1 and w2 stand for the magnitudes, or weights, of the excitatory projections from the two input populations, while w1 and w2 stand for the weights of the inhibitory projections. The equations are described in [2] and [3]. Right: Illustration of large receptive field (y) derived from inputs from smaller receptive fields (x), modeled on the left. The cone and circle represent the focus of attention on X2. This increases the weights from X2 to Y (right). With attention directed outside the receptive field, the response is close to the average of the response to each stimulus alone. The role of attention to a stimulus is to increase the weights from one input population, so that the cell responds primarily from that input. Adapted from [3].

possible incorrect feature pairings, out of the RF [2]. It also argued against the idea often derived from work in the dorsal stream that the role of attention is mainly to facilitate motor responses to a stimulus.

The idea that attention biases competition between neurons, or object representations, has now been realized in computational form as a normalization model of attention, where the competition between object representations is a type of normalization, and attention is mediated by feedback to visual cortex that biases responses toward one input or another. The first quantitative model was developed by John Reynolds in my lab [33] (*Figure 4*). John came to the lab from a purely computational background, but he wanted to combine modeling with neurophysiology. Joined by Tania Pasternak, then taking her sabbatical in my lab, he tried to understand the finding that the effect of attention could be to either increase or decrease the response to a pair of stimuli in the RF, depending on which one was attended. The key insight was that every input to a cell derived from a stimulus must have both excitatory and inhibitory components and that the response of the

cell to multiple inputs must be due to the weighted integration of all these excitatory and inhibitory inputs, or a type of normalization process. This is why, in the absence of attention, when there were two competing stimuli in the RF, the response of the cell to the pair was typically the average of the response to each stimulus alone. The role of attention was to bias the weights of the inputs from one stimulus to another. When Leonardo Chelazzi was a postdoc in the lab, he extended these findings to visual search in V4 and IT cortex [34–36]. John Reynolds developed straightforward formulas that gave a quantitative basis to the biased competition idea. Later, he and others [37] expanded on the normalization models, and they remain the most prevalent mathematical accounts for the role of attention in visual cortex today. The biological basis for those mathematical models remains unknown.

When Beth Buffalo was a postdoc in my lab, she took the attentional studies in a more causal direction, directly testing whether lesions of V4 or TEO would reduce the filtering of distractors by downstream IT neurons, as the biased competition account predicted. Richard Saunders, who learned surgery under Mort, did the lesion. In collaborative studies with Leslie, Beth, Peter DeWeerd, and Guiseppe Bertini, we found that the V4 and TEO lesions not only caused a behavioral impairment when the distracting stimuli were nearby to the target (within the range of a V4 or TEO RF) but also reduced attentional filtering by IT neurons using the same stimulus configuration, as we had predicted [38–41]. The combination of a causal manipulation coupled with both behavioral tests and neurophysiological recordings became the template for many of my future studies. I am proud to say that Beth eventually became the chair of her department at the University of Washington.

The Rhythm Method

Although I was comfortable with these accounts of attention in visual cortex based on firing rates, Wolf Singer (see volume 9) often asked me about the possible role of synchrony, or rhythmic neural interactions, in explaining the effects of attention. Wolf was well known for the idea that gamma rhythms between neurons responding to different stimuli or to features of the same stimulus might bind together these different representations. No one in my lab was willing to test this high-risk idea in the context of attention, but one day I received a note from Wolf that an extremely talented former student of his, Pascal Fries, wanted to come to my lab to do just that sort of experiment.

Pascal turned out to be one of the most efficient people to ever work in my lab. With help from John Reynolds, he set up the experiment and finished the data collection in record time. Contrary to the expectations of everyone in my lab, including me, he found a clear enhancement of gamma synchrony with attention to a stimulus in the RF of neurons in V4 [42–45]. Our thinking was that this gamma synchrony between cells would enhance their impact on downstream neurons simply due to the limited integration

times of postsynaptic cells. An increase in synchrony would be the equivalent of an increase in firing rates. If you are in a room full of people talking randomly, one good way to be heard is to join a group and sing synchronously.

A puzzling later result from Beth Buffalo and Pascal Fries in the lab was that low frequency synchrony was reduced by attention, and this low frequency synchrony was prevalent in the lower layers of the cortex, whereas gamma synchrony was more prevalent in the upper layers of the cortex, suggesting that synchrony played very different roles in different layers. Yasaman Bagherzadeh trained human subjects to control their alpha with neurofeedback and found that it was negatively correlated with attention [46]. Most recently, in a collaboration with Earl Miller's lab, Diego Mendoza-Halliday, Andre Bastos, and Alex Major found a common spectrolaminar pattern across several cortical areas in several species, from mice to humans, with gamma in the upper layers and alpha-beta in the lower layers [47]. This difference between layers may have something to do with their differential roles in feedforward and feedback connections.

Although no one denies the prevalence of synchronous interactions, because they are now found everywhere that anyone looks, I find that there is still a split in the field over their mechanistic interpretation. I find it odd that many people believe that the only important way for cells to communicate is through changes in average firing rate, as though firing rates and synchrony were mutually exclusive. I am encouraged that the field is becoming more receptive to the role of synchrony in neural computation, and I have no doubt that better tools, including high-density recordings and voltage imaging, will settle these questions.

Memory: Recordings of Things Past

The other cognitive line of research I pursued at NIMH was memory, initially with my postdoc Earl Miller, who also received his doctorate in Charlie's lab. He was both Charlie's scientific child and grandchild. Earl loves to tell stories of his desk in the decrepit basement of Building 9, which was located under an asbestos covered steam pipe and reached by traversing former rabbit warrens. The service elevator to the basement had an old brown stain on its wooden floor, which was rumored to have come from someone losing their leg in the iron gate that served as a door.

Earl and I published a paper in *Science*, in which we reported that repetition suppression served as a mechanism for working memory in IT cortex [48]. Repetition suppression was the name we gave for the phenomenon that neuronal responses to repeated presentations of the same stimulus were typically suppressed. Afterward, we added a new behavioral condition to the task, which revealed that our previous conclusion about memory was probably premature. We found a new mechanism, based on enhancement rather than suppression of IT responses to repeated stimuli, but only when

the repeated stimulus was actively held in working memory. Repetition suppression was a passive mechanism, but the enhancement was active. With this new result, it seemed that repetition suppression had properties more suited for perceptual learning and a general sense of familiarity. We published a second paper in *Science* reporting the new results [49] in IT cortex and another one on PFC [50]. I fondly remember that my longtime friend and mentor Larry Squire (see volume 11) was incredulous that *Science* would publish a paper attacking a previous *Science* paper, by the same authors. These papers were part of Earl's phenomenal record in the lab, which helped him land a faculty position at MIT, where he is now a distinguished professor.

Wendy Suzuki came to my lab as a postdoc after receiving her doctorate with Larry Squire. Larry and Mort had enormous respect for each other, but they pursued a friendly rivalry over the years, with different ideas about the organization of the medial temporal lobe system for memory. In my lab, Wendy found that the memory mechanisms based on response enhancement in IT cortex also extended into the entorhinal cortex, but these entorhinal responses were enhanced not only for objects that matched items held in memory but also for the locations of objects that matched locations held in memory [51]. These neuronal findings gave us a possible explanation for the behavioral studies showing the importance of entorhinal cortex for object and place memory, and we were planning to continue this work to understand how the dorsal and ventral streams might converge in the medial temporal cortex.

Cindy Erickson, another postdoc in the lab, also made the very interesting observation that cells in the perirhinal cortex of IT began to respond similarly to successively presented stimuli, based on simple temporal contiguity of the stimuli over time, a type of unsupervised learning of associations in the cortex [52]. For example, if an IT cell responded well to the image of a fork but not a spoon, repeatedly pairing the two stimuli might cause the IT cell to respond well to both. This seemed to be an essential element of an associative memory. The work on memory was picking up steam; however, I was soon faced with vastly increased administrative responsibilities, as described next, and I decided that something had to give. That something was my lab's work on memory, which I put on ice for many years. Fortunately, Earl, Wendy, and Beth all continued to pursue memory mechanisms in their own labs. I am happy to say that Wendy eventually became a dean for the School of Arts and Sciences at New York University.

Mental Illness

Steve Hyman (see volume 10) became the NIMH institute director in the mid-1990s, one of a cadre of innovative institute directors recruited by then–NIH director Harold Varmus. The institute directors are responsible

for both the intramural (IRP) and extramural programs of the institute. Steve was a psychiatrist, but he preferred to have a basic research scientist as scientific director and head of the IRP. The IRP had previously made the transition from a psychoanalytic clinical orientation to a more biological orientation, and Steve wanted an even firmer neuroscience foundation. A previous scientific director, Seymore Kety (see volume 1), was a key figure in the early transition, and he told the story of being approached by IRP psychiatrists, who said that like all previous scientific directors, he should undergo Freudian psychoanalysis. He responded that he couldn't afford it, and that seemed to solve the problem for the moment. But the psychiatrists returned, saying that they had taken up a collection, and they would pay for his psychoanalysis. Kety responded that if they had taken up a collection to remove one of his kidneys, he wouldn't take them up on that offer either.

Steve failed twice in recruiting a new scientific director, mostly because of some of the unattractive features of government service. Given the demoralizing failed searches, Steve offered me the position. The IRP was the largest mental health research program in the world, with more than 65 principal investigators (PIs), a budget of more than $145 million a year at that time, many staff scientists and psychiatrists, and a major clinical program that included psychiatric wards with patients spanning children to the elderly. Although the position was intimidating, I felt an obligation to support an institution that had supported me for so many years. I figured that I was being given the opportunity to shift from only taking satisfaction from the success of my own lab, to taking satisfaction from the success of the people I supported. But I also saw that this position would bring me full circle back to working on mental illness, which is why I had started in the field.

I was sent to interview with Harold Varmus, a Nobel Prize winner who was a very informal and down-to-earth NIH director. My parents would have liked him. Harold himself had little administrative experience when he became NIH director, so I didn't think my lack of experience would faze him. When I went to see him, he had just come back from playing squash, and he asked if he could interview me from the other side of the door of the bathroom attached to his office, as he needed to take a shower. I am sure that wasn't in the NIH human resource (HR) handbook. Harold's lack of any pretentions served him well in his relationships on Capitol Hill, where he was well-trusted. His message to me that day from behind the bathroom door was that the public deserved much more progress than it was getting from the money invested in the IRP and that the NIH needed me to take on this challenge. It was time to revitalize the IRP.

Moving from lab chief to IRP head turned out to be more difficult that I imagined, and it was only possible because of the enormous help I received from Su Koester, my deputy director; David Rubinow, the NIMH clinical director; and Barry Kaplan, the training director. It also helped tremendously

to have Story Landis occupying a parallel position as the scientific director of NINDS. Like me, although she had a doctorate, she was nonetheless responsible for a clinical program. She was also learning on the job, and I like to think that we served as role models for each other.

I knew nothing about managing a large program, and after my first couple of staff meetings, the staff bought me both a gavel and a video on how to hold a meeting, narrated by the comic actor John Cleese. The tape was funny, but the administrative problems were no joke. I eventually paid for management consultants to review the IRP and give me some administrative advice. One piece of advice I remember was to listen to the people in the program and not ride into town like a new sheriff in the old west, to clean things up. I like to think I learned to listen more, but when they finished, they said I had moved from a C to a B−. Much room for improvement.

The first big challenge in my new position came when I was brought a stack of clinical protocols and was asked to sign off on their scientific merit. As a basic scientist, I didn't feel fully qualified to judge whether or not the clinical protocols really had merit, but I was skeptical about some of them. In one of the first reviews by the BSC of a clinical lab I attended, the PI presented the results from the use of bright lights to treat depression, compared with the placebo control of sitting in front of an "ion box." When the clinical results turned out to be equivalent, he suggested that the results were not actually negative because they revealed that the ion box was actually a good antidepressant because of some unknown property of the ions. I called in an outside committee of top psychiatrists to evaluate all of the protocols in the IRP, so we could start with a fresh slate.

The result was depressing. The external committee recommended suspending more than half the existing protocols in the program because they just did not seem novel or important enough to proceed. What I learned was that the clinical research pipeline in the IRP, and really the country as a whole, was running dry of new ideas, which was profoundly discouraging. It was also discouraging to our IRP psychiatrists, who were dedicated to helping their patients. That dedication is probably what led to some wishful thinking about their research findings. With help and support from Steve Hyman and Dave Rubinow, I recruited Dennis Charney to head a new clinical branch, and together with Su Koester, we worked as a team to turn over more than a third of the PIs in the IRP over six years, freeing up lab space, money, and positions for new recruits. Although I tried to treat everyone with compassion and respect, and made all decisions following peer review by the BSC, reassigning so many people and encouraging them to move on was the most difficult part of my job.

Hiring new people was the best part of the job, and one of our new recruits was Carlos Zarate, a young psychiatrist. I can't claim to have foreseen it then, but Carlos, along with Husseini Manji, Dennis Charney, and others achieved probably the greatest success in the revitalized clinical

program by starting the first randomized trial of ketamine for depression, which we funded. This was the first novel mechanistic approach in psychiatry in many years, and the success of ketamine alone probably justified the cost of the new clinical program. Ketamine is also replacing ECT for some applications, so it has personal significance for me.

The profound impact of serious mental illness on patients and their families was brought home to me, when our young teenage son, Matthew, became mentally ill. He became delusional and was hoarding knives, and his psychiatrist recommended hospitalization because he was a danger to himself and others. It was devastating to be unable to help him even though I was the head of the largest mental health research program in the world. The day that Leslie and I sent him to the psychiatric hospital was the worst day of our lives, up to that time. He was later transferred to a residential treatment center for juveniles that was attached to the hospital, but after more than a year, it couldn't help him or even control him, so they simply released him.

For the next several years, Leslie and I went through some of the turmoil that other families have gone through when handling a mentally ill child, including the crisis of the week, the numerous therapy programs, numerous medication changes, a SWAT team raid on our house, and an unwanted journey through the juvenile justice system when things got out of hand. What struck me in the group family therapy sessions was how commonly psychologists assumed that if a child was having serious problems, it must be the result of some bad behavior of the parents. Indeed, some of the parents in therapy did seem to have serious problems of their own, including depression and substance abuse. But my view was that these parents were struggling with many of the same genetic vulnerabilities as their children. The entire families seemed like victims of mental illness to me.

We were trying (unsuccessfully) to recruit Ken Kendler to the IRP at the time, and Ken was the foremost psychiatric genetic epidemiologist in the country. At the recruitment dinner, I asked Ken what he thought of *The Nurture Assumption*, a controversial book by Judy Harris. Judy Harris claimed that parenting, or the shared family environment, had little influence on how children turned out, including personality, talents, mental problems, academic success, and so on. She argued that both genetics and the peer group had larger influences. Ken was adamant that he had the actual heritability data from twin and adoption studies on the contribution of the family environment to the child outcome on multiple measures, and that Harris was wrong when she said that parenting had no role. "So, Ken," I asked, "what is the number? How much of the variance in how kids turn out is due to the parenting?" He stated authoritatively that 4 percent of the variance was due to the shared family environment, not 0 percent like Harris claimed. So, there! Parenting counted for 4 percent. Judy Harris would be gratified to hear that. Ken seemed taken aback when the table erupted in laughter. It is still not settled how

much parents can effectively guide their children, but new research supports the idea that polygenetic factors play an important role together with the overall environment in shaping personality and cognition. When I tell this story, I often find that parents who have highly successful children think I am deranged, but parents of children who have had serious problems often feel relieved and even grateful. I ran into Ken recently and asked him if, after 20 years, there was update on the percent variance explained by the shared family environment. He said the number was still 4 percent.

After a few years, Matthew's life was saved for at least a time by Ray DePaulo, then chair of psychiatry at Johns Hopkins Hospital. Ray agreed to take Matthew into the Hopkins juvenile psychiatric ward, where he took him off all medications to start with a fresh baseline. That had never been done with him before, and I considered it a critical element of the scientific method, in contrast to the many physicians who seemed to pile on one medication after another chasing symptoms and side effects. When Matthew was clean, Ray prescribed lithium, the original bipolar medication which had not been tried before because valproic acid, a newer medication for bipolar disorder, had seemed ineffective. After about two to three weeks on lithium, a miracle occurred. Matthew seemed restored to the wonderful boy we knew before all of his troubles. Bipolar disorder seemed like the right diagnosis, although I no longer believe that psychiatric diagnoses have much mechanistic validity. We did not live happily ever after (both Matthew and Leslie died too young), but lithium at least brought relief to Matthew and indirectly to Leslie and me for several years. However, as frequently happens in families who have gone through trauma, Leslie and I divorced.

In one of the final chapters of that era, the NINDS director Gerry Fischbach (see volume 9), Steve Hyman, Story Landis, and I initiated a proposal to build a new neuroscience building, the Porter Building. Su Koester and I wandered around the NIH campus and found the perfect spot. Steve and Gerry convinced the NIH leadership to persuade Congress to fund the building, and we worked with the other neuroscience institutes and the brilliant architect, Rafael Vinoly, to design it. Rafael designed a stunningly beautiful building. Unfortunately, it ran over budget, and we were forced to go through what was euphemistically called "value engineering." Value engineering meant removing from the architectural plans many things that were really nice, including many of the restrooms. Seeing the building being occupied with new people was like witnessing the birth of a baby. With the building completed, I became focused on how to fill the clinical pipeline with new ideas that could help families.

McGovern Institute at MIT

In late spring of 2004, I was approached by Bob Brown, then the provost of MIT, about the directorship of the McGovern Institute for Brain Research (MIBR). Phil Sharp was the founding director, and although he had done

a brilliant job recruiting a beginning group of faculty members, he was a Nobel Prize–winning molecular biologist, not a neuroscientist, and he had only agreed to stay as director for up to five years, which was coming to an end. The mix of faculty grant funds and private donations at the McGovern Institute was really attractive, as was the fact that MIT is one of the greatest universities in the world, with academic departments such as brain and cognitive sciences and computer science, which we did not have in the IRP. Private funds could be used for risky, early-stage research, which could lead to NIH grants for the faculty if the work was successful. MIT also encouraged faculty to work with industry, and even start their own companies, which I felt would be critical for bringing lab discoveries into the clinic. It was nearly impossible for an NIH PI to start a company at the time.

I could help build basic research in neuroscience at MIT, which would then help fill the pipeline with the new ideas that clinical researchers desperately needed. It helped that my former postdoc, Earl Miller, was a professor at MIT, as mentioned earlier, so I would start with a friend and ally. The administrative burden would be far less than I had at NIMH because I would not be responsible for the scientific reviews of the PIs, and the faculty was spectacular—no turnover was needed. Conversely, one of the challenges of a scientific leadership position in academia compared with the government or the corporate world is that it rarely comes with much authority. You lead by building consensus.

I had previously met with Pat and Lore McGovern at an SfN meeting, and they shared my vision of supporting basic research while keeping an eye on practical applications in the future. Although many donors understandably contribute funds to research because of some particular family connection to disease, Pat and Lore had no such connection at the time. They simply believed in the value of brain research and that their gift would help humankind. They were also willing to put in the time to help the McGovern Institute grow and help raise additional money for research.

I also really liked Bob Brown, who was the provost at that time. About a year before Bob approached me about the position at MIT, I had been invited by Steve Hyman, then provost at Harvard, to visit Harvard to explore the possibility of a recruitment. Steve took me to see Larry Summer's impressive office, located next to his, whose walls were adorned with the original artifacts of the Lewis and Clark expedition. I was duly impressed. Later, when I visited MIT, I met with Bob Brown in his provost office, which had a very large Monet painting on the wall. When I complemented Bob on the painting, he smiled and said it was a perfect digital reproduction. MIT seemed like the right place for me.

Days after I accepted the position at MIT in September 2004, all hell broke loose at intramural NIH, where several intramural investigators were charged with the failure to report money received from pharmaceutical companies for giving talks. It was time for another congressional investigation, and the scientific directors of the institutes would be in the hot seat. You would be damned if you knew about the problems and damned if you

didn't. At the same time, the IRP research budget became constricted. When I was called by congressional investigators, they were crestfallen when I told them I no longer worked for NIH. Many of my friends half-joked that I must have had insider information about the impending meltdown and got out in the nick of time. Actually, it was pure luck.

I had remarried when I accepted the position at MIT in late 2004, and my wife, Chen Chen, gave birth to the first of our two children before we moved to the Boston area the next spring. Having lived in a townhouse in a sprawling Washington suburban housing development near Dulles airport, we decided to go for the old Boston experience, buying a Victorian house built during the Gilded Age. It had the original box on the wall to call for servants anywhere in the house, but it seemed to be broken because no servants appeared when we pressed the buttons.

In fall of 2005, most of the neuroscientists at MIT moved into the fabulous new neuroscience building designed by Charles Correa, the iconic building with the train running through a tunnel in the middle. After we moved in, Pat and Lore McGovern funded events such as the opening celebration, and we put on one of the most sensational opening events ever held, if I do say so myself. When I was at the NIH, there had been a scandal involving a meeting of government employees someplace (not NIH) where the muffins cost $100, or something like that. After the obligatory congressional muffin hearings, it became impossible to serve coffee or cookies at any of NIH events, even if they were paid for with private funds. So, having events at the McGovern Institute, where guests were not only served dinner but also a creative artist like Nick Cave could direct a performance of dancing condoms (or so they appeared), was a welcome change.

As I described earlier, my position as NIMH scientific director was feasible only because of the key people I worked with. The same has been true at the McGovern Institute. Guoping Feng, the McGovern associate director, has been a real partner. The senior McGovern staff—including Gayle Lutchen, the administrative director; Kara Flyg, the director of development; Donna Wells, the financial officer; Jill Crittenden, the scientific adviser; Meagan Jalbert, the human resources director; and Julie Pryor, the communications director—have all been dedicated to the mission of MBIR and have made the Institute feel like a family. I think people considering whether to take scientific leadership positions should think seriously about the key staff, because they are often the people who will make the difference between whether you are a success or failure. It also helped that MIT leadership has consistently been supportive. However, even with the greatest staff and university support, I soon discovered a key difference between the leadership skills I developed at NIMH and those required at MIT. At NIMH, I was an expert at spending money, whereas at MIT, I needed to become an expert at raising money. They are nonoverlapping skill sets. Fortunately, Kara Flyg, the MVP of development at MIT, has made me look good.

I was incredibly lucky that coming to MIT coincided with major developments in science and medicine that accelerated the pace of research. The revolution in genetics not only started revealing the genetic mutations that cause vulnerability to many brain disorders but also led to the development of many genetic-based tools, ranging from optogenetics to single-cell RNA-sequencing to CRISPR methods for altering the genome. The technology for acquiring and analyzing brain imaging data from humans also improved over the years, and we are currently undergoing a revolution in artificial intelligence and machine learning, which is leading to better models of higher brain function and powerful tools for analyzing the immense quantities of data being acquired from the brain. There has never been a better time to be in neuroscience, although I must admit that rapid change can understandably make many young people in the field anxious. The key to success is flexibility.

Of course, MIT was not always a walk in the park. I hesitate to say anything about my first traumatic event at MIT, but sometimes not saying anything is more of a statement than saying something. The very public trauma not long after I arrived resulted from trying to recruit Alla Karpova to the McGovern Institute, with a faculty appointment in the Biology Department. The appointment was opposed by the head of the Picower Institute at MIT at the time. I remain grateful for all the support I received from Pat and Lore, the McGovern faculty (especially Bob Horvitz), and from Phil Sharp in the conflict. The most important thing to say about this was that Alla was the innocent victim of this conflict. She was caught in the crossfire of a battle between opposing visions of neuroscience at MIT, and mistakes were made all around. The best I can say is that I never threw her under the bus. Fortunately, Alla ended up in a position at the HHMI Janelia Research Campus, where she has been doing sensational research, and the MIT neuroscience community healed when Mark Bear and then Li-Huei Tsai later became heads of the Picower Institute. It is a fantastically collegial environment today.

Journey to the East

In 2001, while at the NIH, I made my first trip to China to attend a conference organized by my good friend Lin Chen, and I started relationships that accelerated when I moved to MIT. Pat McGovern's company, IDG, made a great deal of money from its investments in China, and he and Lore wanted to give back something to China by helping it grow its neuroscience. They were joined by Hugo Shong, then head of IDG China investments and now the founding partner of IDG Capital. We all shared the view that brain disorders were global problems that would require global solutions, including from China. Pat and Lore decided to donate funds to start a McGovern Institute in China, organized loosely after the one at MIT. They announced a competition and invited several institutions to submit proposals for an institute, as they had done before they decided to found the McGovern Institute

at MIT. My good friend Bai Lu gave us helpful advice. That all said, many people wonder how there came to be three IDG/McGovern Institutes, all in Beijing.

In 2012, after at least a year of negotiations with the competing universities, Pat, Lore, and Hugo had decided on Tsinghua University for the new institute. We dined with the Tsinghua president to finalize details and then a driver drove us back toward our hotel. As we were en route, Hugo received a call telling us that we were being diverted to Zhongnanhai in Beijing, the secretive compound near the Forbidden City that housed the innermost center of the Chinese government. None of us had ever been there before. Madame Liu Yandong, then a state key counselor and one of the most powerful leaders in China, met with us, and in a magnificent hall suitable for a state visit, she gave an eloquent speech about the importance of the McGoverns and neuroscience in China. She ended by pointedly suggesting that perhaps the McGoverns should also consider Peking University and Beijing Normal University as sites for McGovern Institutes.

As we were driven back to the hotel from the meeting, we all pondered the significance of what had transpired. Beyond the not-so-subtle pressure, it actually made sense to have three McGovern Institutes, because each of them had different strengths, with Tsinghua stronger in molecular neuroscience, Peking University stronger in systems neuroscience, and Beijing Normal University stronger in child education and development. Thus, the three IDG/McGovern Institutes were formed, and it turned out to be a much better decision than having a single institute. The IDG/McGovern donation, the IDG/McGovern Institute branding, and the occasional joint symposia with the McGovern Institute at MIT all attracted a great deal of additional funding and attention to the three institutes, including additional funding from Hugo and IDG Capital. They have grown spectacularly, each with their own directors, new neuroscience buildings, and first-rate new hires. Science is not a zero-sum game—the new faculty members at the three institutes became colleagues rather than competitors. After Pat died, his ashes were placed at the base of a tree at each of the three McGovern Institutes in China and the one at MIT. The trees are all growing beautifully.

Although the three IDG/McGovern Institutes have been very important for neuroscience in China, the most intensive research relationship I personally developed in China was my collaboration with the Shenzhen Institute for Advanced Technology in Shenzhen (SIAT), part of the Chinese Academy of Sciences. The director of SIAT, Jinping Fan; two of my neuroscience friends, Liping Wang and Zhonghua Lu; and my former postdoc, HuiHui Zhou, had started a program dedicated to research in macaque genetic models for brain disorders. The newly constructed state-of-the art animal facility featured zoo-like group housing surrounded by trees, and it was accredited by AAALAC, the international accrediting organization that inspects and accredits animal facilities around the world, including MIT.

My colleague at MIT, Guoping Feng, and I strongly believed that primate models held the key to unlocking the next generation of therapeutic breakthroughs in brain disorders. Although mice have and always will play a crucial role in fundamental neuroscience research, treatments developed for brain disorders in mouse models have often fallen short when tested in humans. We were convinced that treatments developed in primates would stand a better chance at success in humans. Macaques with SHANK3 mutations were developed by Andy Peng at Sun Yat-Sen University and Shihua Yang at South China Agricultural University. They were developed in a breeding farm the size of a small city—I thought of it as Cambridge for monkeys. The founders were transferred to SIAT for study. SHANK3 is the gene mutated in Phelan-McDermid syndrome, which is characterized by severe intellectual disability and autism in humans. Guoping is one of the world's experts on the SHANK3 gene, and we jumped at the chance to collaborate in studying them.

As part of the collaboration, we found that the macaques with SHANK3 mutations had alterations in social behavior, cognition, and functional connectivity measured by fMRI that were reminiscent of some of the symptoms in Phelan McDermid syndrome [53]. We were fortunate to secure grants to study the SHANK3 monkeys through the U.S.-China Cooperative Biomedical Research Program, created during a time when China and the United States cooperated in programs of joint interest. As part of the program, MIT received funds from NIH, and SIAT received funds from the Chinese National Science Foundation. Although the amount of funding was modest, we were grateful that both the NIH and the Chinese NSF had carefully reviewed and approved our collaboration on the basis that it was of mutual benefit to the two countries.

Our partnership was forged just before a time when many U.S. scientists, almost all of Chinese origin, were facing scrutiny or even persecution from the NIH, federal law enforcement, and some universities because of their collaborations with Chinese scientists and institutions. Although everyone accepts the importance of national security and the need for full disclosure of activities, this seemed more of a witch hunt. It stemmed from the "China Initiative" started during the Trump administration. These researchers were charged with violating various regulations related to foreign collaborations, but many of the charges were ridiculous, and the interpretations of these regulations changed so often that even some NIH institute directors confided in me that they could not fully comprehend them. Nearly all of the affected people were conducting basic research with results that were freely published in international journals. Furthermore, if the plan was to restrict any transfer of scientific information to China, the plan backfired because it caused many Chinese-origin scientists to return to China where they established labs.

After many lives were ruined, the rules for collaboration were eventually clarified, the China Initiative was terminated, and the witch hunt

seemed to diminish. Ironically, some of the scientists caught in this dragnet had previously endured the turmoil of China's Cultural Revolution. They never imagined that they would experience such persecution in the United States. As I reflect on my experiences in China and my collaboration with SIAT, I remain convinced that the surest path to better treatments for brain disorders is through global scientific partnerships, including the United States and China.

King of the Brain: My Unexpected Journey into China's World of Super Brains

Among my countless adventures in China, my participation in the Super Brain television show stands out as the most unexpected and remarkable (*Figure 5*). This wildly successful game show showcased astonishing intellectual abilities and captivated the nation, amassing a staggering 250 million viewers who tuned in to watch achievements of the human mind. By comparison, the very popular American Idol show in the United States had an audience of less than nine million viewers.

My involvement began in 2014, during the show's first season, when a colleague from China who was advising the show called me because of public concerns about the authenticity of the contestants' remarkable skills. He asked me to join the panel of discussants on the stage of the show and carefully scrutinize the competition for any signs of fraud. I found that the contestants' exceptional talents were genuine, and the show was teeming with human drama and excitement, reminiscent of an intellectual Olympics. I was hooked, and I returned every season except during the pandemic. I usually arrived at the competitions jet-lagged, once falling asleep onstage for 45 minutes. The accommodating crew simply filmed around me.

The entertainment quality of the contests was over the top. A Rubik's Cube contest would not just require solving cubes, but solving them underwater, holding one's breath, and blindfolded, all at the same time. A memory contest would not consist of memorizing numbers or letters, but memorizing the features of walnut halves and then finding the companion halves among hundreds of walnut halves on a wall. A comparable contest involved a wall of lips. I still have some of the walnuts and lipstick marks in my office, in a display case of mementos from the show. In later years of the show, the problems could no longer be solved with narrow talents, such as memory or perception. They required very high degrees of intelligence to solve unique puzzles, such as observing the motion of many planets and moons in an alien solar system set up on stage, and predicting their locations on a future date. Once, one of the contests was so complex I could barely understand the rules myself, and it needed to be performed at lightning speed. The host of the show asked one of the contestants if he was worried about how he would perform,

Figure 5. Bob Desimone with some of the cast and contestants of the Super Brain Show in 2019. To the immediate left of Desimone is Jiang Changjian, the host, and to the right are Kunlin Wei, the leader of the Chinese team, and Yigong Shi, president of Westlake University.

and he answered that it didn't seem so hard because it resembled in some ways the game of "Go." Surprised, the host asked him if he was an expert at Go, and the contestant said that he had never played Go, but he had read the rules and it didn't seem so hard. A Go champion on the panel was quite upset when this very smart novice won the contest, beating actual Go players.

The program's popularity was further boosted by the frequent appearances of famous pop stars, comedians, and high-visibility academic leaders, such as Yigong Shi, president of Westlake University, on its panel. Yigong is one of the most accomplished scientists in China, and he is revered by many of the families with children watching the show. The very popular host of the show is Jiang Changjian, who is also a distinguished professor of political science at Fudan University. One of his talents is setting the contestants at ease and getting them to talk about their life. It was important that the audience care not only about the contests but also care about the very real people competing, always under intense pressure. The audience was introduced to obscure talents, such as Alex Cummings, the international memory champion and U.S. medical student, and Rinne Tsujikubo, a 10-year-old girl from Japan who was a calculation prodigy. The Chinese contestants ranged from children of farmers straight from the countryside to top university students and the son of a famous billionaire. The audience was live and enthusiastic, and the pace and editing were heavily influence by K-Pop.

It was the perfect mix of intellectual talents and entertainment, so it is perhaps not surprising it was so popular. One year, as I was exploring a Mayan temple in the Yucatan jungle with my family, a Chinese couple recognized me and asked about the show. My children looked at me in awe. It seemed that my participation in Super Brain had made me a recognizable figure among Chinese viewers, and it was often the first topic of conversation when I met with business and academic leaders in the country. I had to wonder whether any television show centered on intellectual talents would be so popular in the United States.

The highly entertaining contests were the brainchild of a creative team that included my friend Jia Liu, now a psychologist and AI researcher from Tsinghua University. I recruited some of the international contestants, and I persuaded other neuroscientists like Tom Sudhof, Marge Livingstone, Dora Angelaki, Tom Albright, and Mike Gazzaniga (see volume 7) to appear as guest panelists. Many people in the United States have assumed that I must speak Chinese to be on the show. I am almost ashamed to say that after more than 22 years traveling to China, I never learned to speak Chinese, although my wife and children do. On the show, I have a speaker in my ear with real-time translation of everything said. By the end of 12 hours or more of filming each day, I have the strong illusion that I actually understand Chinese.

The winner of the domestic competition phase of the show was crowned the king of the brain, and the king and many of the other contestants often became famous in China. Some became internet "influencers" and others leveraged their fame to start successful businesses. Soon, I became the international team leader for the international competition. For several years, the leader of the Chinese team was Kunlin Wei, a psychologist from Peking University. Unfortunately, the Covid pandemic interrupted my appearances on the show, but I am pleased to say that on our first post-Covid show, I led the international team to victory. It was hard to give the team much specific advice because we were informed about the exact nature of the competitions just a few days in advance, but the one piece of general advice that was always relevant was to stay calm. Often, the winner was the person who could best hold up under pressure. I think it is a relevant lesson for many of our current students and postdocs.

Throughout my career in neuroscience, my focus had been on understanding the brain to help people with disorders and disabilities. However, I soon realized that the exceptional minds on Super Brain could also hold invaluable insight into the brain. One example of this occurred during an after-show party, when Mike Gazzaniga and I questioned a Japanese calculation expert from our international team, Hiroaki Tsuchiya. We had just watched him solve complex problems with astonishing speed, such as dividing an 18-digit number by an 11-digit number, with no delay between seeing the numbers and writing the answer. Surely, Hiroaki Tsuchiya had a rare

inborn talent for calculation, but it was honed by practice. He revealed that all calculation champions initially practiced with an abacus, eventually giving up the physical device in favor of manipulating an imagined one, and finally abandoning even the mental beads. Eventually, he explained, the answers simply came to him, digit by digit, as soon as he saw the problem. My guess is that the learning with the abacus engaged the motor system of his brain, which has structures, such as the basal ganglia and cerebellum, that have the capabilities to perform many calculations in parallel, which are normally needed to throw a ball or ride a bike. Even I can throw a ball without a delay to calculate the joint angles. I think Hiroaki Tsuchiya learned how to use these different parts of his brain for math calculations, and I have no doubt we have much to learn from these abilities.

Dissecting Cause and Effects in Complex Systems

I think many people who have worked in neuroscience for a long time have had the experience of publishing a few papers in a new area and suddenly finding that there are aisles of posters on the same topic at the SfN meeting. This was my experience more than 15 years ago at the SfN meeting, when I started to see many great posters reporting effects of attention on responses in many visual structures, and these posters were followed by many great papers, of course. At that point, the attention field seemed crowded with great studies, and maybe it was time to jump the shark. I decided that I should shift my own attention away from the phenomenology of attention to understanding the biological mechanisms for these attentional influences on cells. In other words, shift from correlation to causality. Given that there really were not good tools for understanding the microcircuitry of the cortex in awake macaques, the best approach seemed to be to study the interactions among the different elements of the circuits distributed across different cortical areas and subcortical structures. If you want to study the interactions between different cells in awake primates, it really helps if those cells are located in different structures.

The most important tool for studying causal interactions across brain structures in rodents has probably been optogenetics, but it had not been tried in primates. One of the many great advantages of MIT is its proximity to wonderfully creative people, such as Ed Boyden, who are likely developing whatever tool one needs. Ed and I quickly formed a collaboration to test in monkeys the opsin Arch, which can suppress neural activity when stimulated with blue light. The project was led by a joint postdoc, Xue Han, now a professor at Boston University, who succeeded in manipulating neurons using optogenetics for the first time in a primate [54]. Leah Aker, then a graduate student working with Ed and me and now an anesthesiologist and professor at Duke, later succeeded in suppressing neurons in the frontal eye fields (FEF) with the opsin Jaws, developed by Ed, which was better suited

for the large primate brain because its sensitivity was shifted into the red wavelengths and was less absorbed by the blood in the brain [55, 56]. Jaws is now routinely used in primates. Ed and I have continued these collaborations in technology, now with a joint postdoc, Nava Shmoel, working on signal reporters for imaging.

Using a variety of causal methods, including optogenetics, muscimol, lesions, magnetoencephalography (MEG) recordings in humans, and analytic techniques, such as synchrony and the latency of attentional effects on cells, I sought to understand how the attentional effects on cells in visual cortex came about. For spatial attention, the FEF was already known to play a key behavioral role, and Georgia Gregoriou in my lab (now a professor in Crete) found that the effects of spatial attention on cells in FEF occurred with a very short latency—much shorter than in V4 [57]. The latencies were short enough for FEF cells to be the source of attentional feedback to cells in V4, and the cells also went into gamma frequency synchronization with V4 cells, with the phase of their activity leading that in V4 by around 10 milliseconds, around the estimated transmission time from FEF to V4. Georgia and Andrew Rossi (now a program officer at NIMH) found that lesions of PFC, including FEF, impaired the animals' ability to attend to a target in the presence of distractors, and reduced the filtering of distractors by IT neurons [58, 59]. Leah Acker found that optogenetically suppressing FEF cells at virtually any time during a spatial attention task impaired performance in the task [55]. By contrast, when HuiHui Zhou (now a PI at the Peng Cheng Laboratory in Shenzhen) and Bob Shafer in my lab tested the role of the lateral pulvinar of the thalamus in attention, they found that deactivation of the pulvinar left the animal almost blind in the affected portion of the field, and it greatly reduced the responses of cells in V4 to sensory stimuli, whether or not the animal paid attention to them [60]. The FEF performed like an attentional control structure, whereas that part of the pulvinar seemed to have more basic visual functions. Incidentally, Bob Schafer left my lab to start his own company testing humans instead of monkeys. He later became the head of research and then CEO of Luminosity, the cognitive training company. The tech industry is increasingly an employer of some of our best postdoctoral researchers, which I think is a very good opportunity for people who are willing to consider alternative career tracks.

In another line of work in the lab, we undertook studies that questioned the very basis of many studies of spatial attention, from my own and other labs. Neurons in many brain structures play a role in oculomotor control and respond in association with saccades into their movement field or receptive field. To study spatial attention independently of these oculomotor responses, we and most other people in the field used the popular covert attention paradigm, in which animals (or people) were trained to fixate on a spot and then hold fixation on it while they attended to an object in the extrafoveal field and ignored distractors. By cuing the animal to attend to

one object or another, we could test the effects of attention, independently of eye movements, or so we thought. What we didn't know at the time was that the animals were often making small eye movements, termed microsaccades, toward the attended object and then back to fixation, two to four times per second, or roughly delta/theta frequencies. Studies in humans had shown that the direction of a microsaccade is often associated with the direction of spatial attention.

When Eric Lowet joined the lab, we decided to investigate this relationship. In many labs, the eye position signal from a typical infrared eye tracker is often too noisy to even detect small microsaccades. It is easy to ignore what you don't measure. However, Eric had received his doctorate while working on microsaccades, so he knew how to carefully measure them. Joined by Khartik Srinivasan in the lab, Eric initially found that it didn't make a great deal of difference whether or not a trial in a covert attention task included a microsaccade. The key insight came later, when he separated the intervals following microsaccades toward the attended stimulus from the intervals following microsaccades back to the fixation spot. As soon as that was done, it became clear that the effects of attention on the response to a stimulus in the RF were predominantly in the interval following a microsaccade toward the attended stimulus and not the opposite direction [61].

Later, Karthik, Eric, and Bruno Gomes found that cells also conveyed more information about the stimulus in an interval following a microsaccade and had sharper tuning curves [62]. It appears that attention to a stimulus is more strongly engaged when an animal makes a microsaccade toward it. Although this finding does not invalidate any of the work that has been done on spatial attention, it does put it in the context of other elements of cognition with motor associations. It brings the research on the covert attention paradigm closer to natural vision and visual search, where the eyes dart around to significant stimuli in the world, sampling them one to four times per second with saccades. In the covert attention paradigm, the same rhythmic sampling of the scene seems to occur, but with attention coupled to small saccades that often escape detection.

Focusing on Features

Feature-based attention has always seemed more intriguing to me than spatial-based attention. For spatial-based attention, it is easy to imagine that high-level areas with a visuotopic organization, such as FEF, send visuotopically organized feedback projections to areas such as V4. Cells in FEF with RFs containing a behaviorally relevant stimulus would have greater activity than cells with other RFs (the notion of a priority map in FEF), and this greater activity would be fed back to V4 to bias their responses toward the same stimulus. Synchrony between FEF and V4 could magnify these effects. But for feature-based attention, it is less clear how this feedback

might work, because FEF has little or no feature selectivity, and there was no obvious map for features elsewhere in the PFC.

Imagine searching for your pen on a crowded desk. It is commonly thought that you would use your memory for what your pen looks like, and then use some or all of the stored pen features (what is known as the attentional template) to bias the neural representations of objects resembling pens on your desk so that you attend and orient toward the most likely candidates. For example, you would probably be more likely to attend first to a pencil on your desk rather than to a coffee cup. This bias in favor of searched-for features is known to operate globally across the visual field, in contrast to spatial attention, which normally operates on one or two locations at a time. Narcisse Bichot in my lab showed that during a visual search task like this, cells in V4 give enhanced responses to stimuli matching the searched-for object anywhere in the visual field. But where is the attentional template stored, and how does it bias features or object representations throughout the visual field in areas, such as V4?

It was known that during visual search, cells in FEF have greater responses to objects matching the features of a searched-for object, or the target, than to nontargets, similar to what Narcisse Bichot found in V4 [63], and HuiHui Zhou in my lab found that the latencies of these FEF cells for targets versus nontargets are earlier than in V4 [64]. Could FEF cells be the source not only of top-down feedback for spatial attention but also of top-down feedback to V4 for the features of searched-for objects? That seemed unlikely. How could FEF cells with limited receptive fields sensitize V4 cells to respond best to red targets anywhere in the visual field, to take just one example? How would FEF compute the locations of stimuli matching the attentional template when FEF cells show little or no selectivity for object features?

The answer to this puzzle was that the source of feedback for attended object features was not in FEF, but rather in an area located just anterior to FEF, which we termed the ventral prearcuate region, or VPA. Narcisse Bichot led a team in the lab to record in VPA and FEF while animals performed a visual search task, similar to finding a pen on a desk. The animal was shown a cue stimulus at the start of the trial, which varied from trial to trial. After a blank delay period, an array of stimuli appeared throughout the central visual field, and the animal was rewarded for finding and holding fixation on a stimulus that matched the cue (termed the target). The animal had free gaze and could make any number of saccades to find the target. The saccades in this case were big and easily measurable.

Narcisse found that cells in VPA often had feature or object selectivity in their responses to the cue, like in IT cortex, but also some had RFs the size of RFs in FEF [65], so the cells combined spatial and feature information, as Earl Miller had previously found in PFC. Furthermore, cells in VPA gave enhanced responses to stimuli matching the target features during visual search, like in FEF, but the latencies of these effects were even earlier than in V4. By

contrast, when he computed the latency of the effects of spatial attention in VPA and FEF, the latencies were shorter in FEF. These results suggested that the locations of stimuli matching the target might be computed in VPA, with information about the location of likely targets communicated to FEF.

This idea was supported by the effects of deactivating VPA with muscimol and measuring the effects on cells in FEF. Unilateral VPA deactivation impaired search in the contralateral field, suggesting that VPA was important for search. It also eliminated the effects of feature attention on the responses of cells in FEF. For example, if there was a red stimulus in the RF of an FEF cell, following deactivation of VPA, it no longer made a difference to the cell's response if the animal was searching for a red stimulus versus searching for a green stimulus. By contrast, following VPA deactivation, FEF cells showed normal response enhancement depending on whether or not the stimulus in the RF was the target of a saccadic eye movement. In sum, the results suggested that VPA and FEF functioned together, with VPA the source of top-down biases for feature-based attention and FEF the source of top-down bias for spatial attention and saccades.

One of the lessons I take from the work on FEF and VPA is that initial appearances (from recordings) can be deceiving. When recording from both areas during visual search, the response properties will initially seem very similar. In fact, cells throughout most of the PFC seem to have similar properties. But with careful dissection of properties coupled with causal manipulations, one can tease apart their differential roles.

What about the role of extrastriate areas, such as V4? Certainly, extrastriate areas must be the source of information about the features and locations of all the stimuli in the visual field, which is communicated to areas in PFC such as VPA and FEF at short latencies. However, our results suggested that VPA and FEF computed the locations of behaviorally relevant stimuli, with FEF then directing attention or the eyes to likely targets. If so, what was the purpose of the feedback from VPA about attended features to extrastriate cortex, if VPA and FEF could function to orient the animal to behaviorally relevant stimuli on their own?

One likely possibility is that extrastriate areas are involved in more complex, lengthy visual tasks than simply orienting to a relevant stimulus. We found a potential example of this in a study of feature-based attention in humans using MEG and fMRI [66]. Daniel Baldauf led a study in the lab in which subjects were show a sequence of faces and a sequence of houses, spatially overlapping with each other. Subjects were cued on each trial whether to attend to the faces or the houses, and to press a button if either two successive faces or two successive houses matched each other, depending on the cue. He found that the inferior frontal junction (IFJ) in PFC went into synchronized oscillations with the fusiform face area if the subjects were attending to faces and with the parahippocampal place area if the subjects were attending to houses. Furthermore, the IFJ oscillations led the ones in the temporal cortex by around 20 milliseconds, which was the

expected transmission time. IFJ seemed a good candidate for the human equivalent of VPA, but in this case, the task required the sustained involvement of the face area or place area in temporal cortex to process the relevant stimulus type so it could not be solved by simply selecting a face and orienting to it. Other complex tasks probably require sustained interactions between VPA and other parts of extrastriate cortex.

How did the IFJ (humans) or VPA (monkeys) have this effect on sensory processing areas? Rui Xu and Narcisse Bichot in the lab set out to study this by measuring what we called the "connectome" of lateral PFC [67]. They used electrical stimulation of PFC applied during fMRI scanning to measure all of the brain areas activated by the stimulation, based on methods developed by Nikos Logothetis and others. By applying the stimulation to a dense grid of stimulation sites, they arrived at a very fine-scale mapping of PFC connectivity. The results showed that lateral PFC, including VPA, had connections with five of the major processing domains of the cortex, and that within each of these processing domains, the connections with PFC were topographically organized on a fine scale. If one considered an orderly progression of stimulation sites in PFC, for example, they were connected to an orderly progression of connected sites in most or all of the processing domains. Because at least some of these processing domains had a hierarchical organization, such as the ventral stream for vision, this meant that PFC was in a position to integrate information and coordinate processing at comparable hierarchical levels in nearly all major processing systems.

This method worked so well to reveal the connectome of lateral PFC in the monkey, it occurred to us that the same method might be used to map human brain neuroanatomy. Some epilepsy patients have electrodes temporarily implanted in their brains to identify the epileptic focus for later surgery. In a collaboration with the MGH neurosurgeon Mark Richardson and his colleagues, MRI physicist Atsushi Takahashi, and McGovern faculty members Ev Federenko and Nancy Kanwisher, we began a year of safety testing, and we are now ready to start the mapping the connections of these electrodes in patients with epilepsy who volunteer. This not only will give us unprecedented information about human brain connectivity but also help neurosurgeons identify components of the epileptic network in each patient, which could serve as a treatment target. It is just another example of how basic research in animals can ultimately help clinical practice.

Working Memory Revisited

I was also able to return to studying memory, which I had given up when I became NIMH scientific director. Diego Mendoza Halliday (now on the faculty at the University of Pittsburgh) in the lab used an optogenetic manipulation to dissociate the role of VPA in working memory from attention [68]. Although I had lectured in my classes for years that the neural

circuits for working memory had proven impossible to dissociate from those for attention, so they were probably the same mechanism, I had never tested this idea myself. In a task requiring feature attention to direction of motion, Diego found that optogenetic deactivation of VPA impaired the animal's performance if the deactivation occurred during the attentional selection phase of the task but not if it was deactivated during the working memory phase. Likewise, deactivation during the attentional selection phase reduced the effects of attention on cells in areas such as MST and LIP, but there was little impact on the effects of working memory on cells if the deactivation took place during the working memory phase. It looks like there is a dissociation between the circuits for attention and working memory, with VPA much more important for attentional section. I will have to change my class lectures. Although the role of prefrontal feedback in attentional selection seems well established, our study adds to a growing literature questioning what we thought we understood about the mechanisms of working memory.

We have barely scratched the surface in understanding the neural circuitry of attention, but I feel like the playbook we are following is working. Anatomy, physiology, behavior, and causal manipulations are together revealing at least the major structures involved, and the general principles of how they work together. The next major challenge will be to develop good computational models of how it works at the cellular level.

Marvelous Marmosets

My McGovern colleague, Guoping Feng, and I envisioned a two-pronged approach to developing better therapeutics for brain disorders using primate genetic models. The first approach was gene therapy to correct the genetic mutation, which Guoping had already demonstrated in mouse models of brain disorders. This might be the best solution for disorders with single gene mutations with large effects, but it wasn't clear how gene therapy could be applied in animals with the more common polygenic forms of disorders, in which many genetic variations might each contribute a small degree to the disorder. The second approach was to try to understand the neural circuit dysfunction in a primate genetic model, and correcting it at the circuit level rather than with gene therapy. We hoped that some of the circuit dysfunctions we identified in the animals with monogenic mutations of large effect would be shared with the more common, polygenic forms of the disorder. For example, if we identified a circuit dysfunction in the striatum that led to stereotypies in one of the monogenetic models for autism spectrum disorder (ASD), a drug or some type of genetic medicine targeted to the striatum to correct the stereotypy might also be effective in other, polygenic, forms of ASD. Based on what the field has learned from the study of mouse genetic models for brain disorders, understanding circuit dysfunctions would likely require a very large effort, utilizing several neuroscience methods.

We reasoned that this need for a large neuroscience effort to understand the primate genetic models would be facilitated if we had genetically modified animals to study at MIT as well as at SIAT. The problem was that macaque studies such as these would require large groups of animals, and it would require too much space and too much money to do the studies at MIT. Although we had no experience working with marmosets, these Brazilian primates were much smaller than macaques—indeed, they were not much larger than rats and lived in multigeneration family groups. Their small size made it feasible to have a marmoset genetics program at MIT, which has limited animal space. Fortunately, we learned that the New England Primate Center was donating marmosets to a small number of national primate centers and universities. Harvard was closing the center for financial reasons, but looking back, this was a poor decision given the explosive need for primates for vaccine development during Covid. With help from then-NIMH director Tom Insel, the Primate Center gifted us a very small number of marmosets, which became the starting point for a breeding colony. The MIT Division of Comparative Medicine, first under Jim Fox and later under Kelly Pate, became our partner in developing the facilities and expertise needed for the marmosets.

We soon became desperate to obtain more marmosets to increase the genetic diversity and the rate of breeding. The only site that had available marmosets was a breeding colony in Japan, but no international carriers would transport primates from Japan to the United States at that time. Lore McGovern asked a friend who owned a fashion empire if we could fly them on their private plane, but filling a luxurious plane with rat-size marmosets posed daunting clean-up challenges, not to mention snakes-on-a-plane scenarios. Fortunately, we found a Mission Impossible–like team headed by a former Navy SEAL who seemed unfazed by transporting anything anywhere. Lore's generous friend paid for the chartered 757 cargo plane and the Navy SEALs needed to fly the marmosets directly to a quarantine facility in the United States, and our U.S. marmosets soon had Japanese friends.

I have found the international community of marmoset researchers to be much like the marmosets themselves—highly social and cooperative. I don't know if working with marmosets caused the researchers to develop these qualities, or whether marmosets attract people with those qualities. In either case, marmoset researchers from around the world, including Japan, England, and Brazil, as well as universities in the United States invited us to learn from them as we got started. Charles Jennings in MIBR helped us obtain a Mass Life Science Center grant to buy the key equipment for the genetic engineering platform. Guoping and his team headed by Qiangge Zhang and Martin Wienisch used CRISPR methods developed by Feng Zhang at MIBR and the Broad Institute to create the genetic mutations.

The first grant application Guoping submitted to NIH for the marmoset project was turned down because he had not already shown that we could

create the genetic models. So how could we raise the money to get started? Fortunately, a private research institute and private donors committed to research on psychiatric disorders stepped in. For privacy reasons, I will not name the donors, but they gave us the critical gift funds we needed to get the program off the ground. Now, the program is supported with a mix of public and private funds. Guoping is leading a large gene therapy project in marmosets with the mutation in the SHANK3 gene found in Phelan-McDermid syndrome, and several labs at MIT, Harvard, and other cooperating universities are studying circuit function in primate genetic models to develop better therapies. Guoping's SHANK3 mini-gene is now licensed to a company to start clinical trials. More than half my lab, including Haoran Xu, Beizhen Zhang, Florence Liang, Veronica Su, Claudia Valenzuela, Mark Fossesca, and Frederico Azevedo, is now engaged in studying social cognition in marmosets. A team led by Haoran has used ecog electrodes to map a network of areas concerned with the perception of faces, bodies, and social interactions. They work with Will Menegas, Jitendra Sharma, Erin Corbett, Ruoyang Chai, Yefei Chen, Minqing Jiang, Gina Liberti, and others in Guoping's lab on the SHANK3 project. When new therapies for brain disorders are developed with the primate models and tested successfully in people, I will consider it a dream fulfilled.

A Thousand Flowers Bloomed

Beyond the research in my own lab, my proudest accomplishment at MIT is the role I played in developing the McGovern Institute into the research center it is today. The faculty have free will and choose their own research directions, but the McGovern funding and other major gifts allowed us to create the environment to attract the best people and to give them the support they needed to succeed, including a building (MIT committed to the new building in association with the McGovern gift), core facilities, the best administrative support, many graduate fellowships, and some significant seed funds for pilot projects, which are critical for obtaining the preliminary data needed to apply for federal funds. Pat and Lore's commitment made more than 20 years ago made all this possible.

The formal mission of MIBR is very broad, encompassing neuroscience and many related fields, but one of the more controversial decisions of MIBR was to recruit an equally broad faculty from the very beginning. Phil Sharp recruited a faculty with interests ranging from molecular genetics, to computation, to human cognition. This beginning group included Emilio Bizzi (see volume 6), Martha Constantine-Paton, Jim DiCarlo, Michale Fee, Ann Graybiel, Alan Jasanoff, Nancy Kanwisher, Bob Horvitz, and Tommy Poggio. Martha and Emilio later retired. This diversity was criticized by some people at the time because they believed that neuroscience centers like the McGovern Institute should have a narrow focus to succeed. When I arrived at MIT, one of our distinguished board

members at that time suggested that the McGovern Institute should focus on one topic, like oculomotor control, in the future. Rather than becoming the oculomotor center of the world, I felt our goal should be to have a broad impact on brain health. This would require a multidisciplinary approach, recruiting people who could communicate with each other across levels.

Facilitating such breadth required not just freedom but also a culture of collaboration. We focused on hiring people who were the best in their fields, of course, but who were also driven to work closely with others. The result was an atmosphere in which weekly faculty lunches and periodic retreats became more than social gatherings—they were idea incubators, in which conversations sparked research projects and seeded long-term collaborations.

The additional full or associate faculty members currently at MIBR include Polina Anikeeva, Ed Boyden, Ev Federenko, Guoping Feng, Ila Fiete, John Gabrieli, Mark Harnet, Hugh Herr, Mehrdad Jazayeri, Rebecca Saxe, Nidhi Seethapathi, Fan Wang, Robert Yang, and Feng Zhang, who are amazingly diverse in their interests but very interactive across levels. Most have academic appointments in the Brain and Cognitive Science Department but some have appointments in Material Science, Biology, Biological Engineering, Electrical Engineering and Computer Science, and Media Arts and Sciences. I have never worked with a more friendly and collaborative group of people. Of course, the real test of the interactive philosophy is the research progress, and in my view, it has also been outstanding. Our faculty have advanced our fundamental knowledge of the mind and brain at all levels, from gene regulation, to the biophysics of neurons, to the neural circuitry of perception and attention, to motor control, learning and memory, language, navigation, the relationship between the brain and the gut, the experience of pain and anxiety, child development, and theory of mind, with many or most advances resulting from collaborations. Being at MIT, we also have a strong interest in developing new technology, which is freely shared with the neuroscience community.

MIBR provides a very fertile field, but the faculty grow the crops, of course. For this, our faculty have benefited tremendously not only from government grants and foundation support but also from visionary private supporters, some of whom endowed research centers that are now supporting innovative research in many areas. Many of our supporters share with me and other faculty members a history of neurodevelopmental or psychiatric disorders in the family, and we all understand the desperate need for more research and better therapies. They don't just donate money—they stay engaged with the science in many ways, including serving on one or more of our boards. Lore McGovern developed an interest in addiction and has supported a new addiction initiative. Jim and Pat Poitras funded the Poitras Center for Psychiatric Disorders Research, the first major research center established at MIBR. After hearing about my plans for translational research in psychiatry soon

after I arrived at MIT, they stepped forward and asked how they could help. It has supported collaborations between John and Susan Gabrieli and several clinical groups around the Boston area, including at McLean Hospital, as well as many preclinical studies of pathophysiology in psychiatric disorders, and groundbreaking gene therapy methods. The Poitras Center was so successful that it became the prototype for five later centers. Lisa Yang and Hock Tan established an Autism Center and a Molecular Therapeutics Center, and Lisa Yang independently established the Brain-Body Center (headed by Polina Anikeeva), a Bionics Center (focused on new types of prosthetics with better neural integration, headed by Hugh Herr and Ed Boyden), and an integrative computational neuroscience center (ICoN, headed by Ila Fiete).

The centers established by Lisa Yang and Hock Tan now make up the Yang-Tan Collective, which is a research accelerator. The faculty supported by the Collective span several departments and different schools at MIT. Although the Collective covers several diverse areas of translational neuroscience, joint retreats, monthly talks by fellows, and common social gatherings have led to the creation of many bridges between projects. We also benefit from frequent interactions with the Tan-Yang Autism Center at Harvard, headed by Mike Greenberg. More than 20 postdoctoral fellows and students are supported by all of the Collective Centers, and they have all become part of a highly interactive community that has forged many new collaborations across labs. The ICoN Center of the Yang-Tan Collective only supports postdoctoral fellows who have at least two faculty mentors, for example. The Yang-Tan Collective also funds postbaccalaureate scholars who are from disadvantaged groups, giving them the research experience that they need to apply for graduate schools. All of the fellows and scholars benefit from regular personal interactions with Lisa Yang, who encourages their research. The magnitude of the gifts is extraordinary, and all of our Centers have transformed research in MIBR and MIT.

It has been a very long road since I decided as a graduate student to pursue research that I thought would ultimately help people, but I am seeing some light at the end of the tunnel. I am deeply indebted to all of the students, staff, and postdocs who did all the research in my lab over the years. I am sorry that I could not describe the work of all of them in this short autobiography. But my biggest role may turn out to be the help I have given to others to achieve this goal, not simply directing the work of my own lab. The CRISPR discoveries of Feng Zhang, many of them funded by Jim and Pat Poitras and Lisa Yang, are transforming medicine and advancing into clinical trials. The primate genetic studies headed by Guoping Feng, also supported by Lisa, Jim, and Pat, are on a solid path toward gene therapies and other therapeutics. Our faculty have cofounded numerous companies and have licensed their discoveries to an even larger number, which are developing clinical applications and starting clinical trials. I have become a cofounder of two companies myself, which are developing products for clinical use.

I believe this research progress validates the reason that I left NIMH, which was that to make clinical progress, we needed to go back to the bench and develop the foundational knowledge of the brain and behavior that would ultimately fill the clinical pipeline with new ideas. And the fundamental knowledge we have acquired is just the start—there are amazing discoveries about the mind and brain still to be made. It brings immense satisfaction to know that I have played a role in supporting so many faculty members, staff members, and students in realizing their research dreams. I feel even better when I contemplate how their work will ultimately benefit humanity. I have no doubt that Pat McGovern is also smiling someplace, happy with the seed that he planted. In a way, my broader role in helping people also serves as a token of gratitude for the aid I've received throughout my career—from generous scholarships and fellowships to the invaluable guidance of my mentors, Charlie Gross and Mort Mishkin.

I am also thankful for my wonderful wife and our two wonderful children, who, remarkably, are considering careers in biomedical research. This scientific autobiography is really written for them. The journey thus far has been amazing, and I see no reason to decelerate just yet. As Pat McGovern used to say, the best is yet to come.

References

1. Bruce, C., R. Desimone, and C.G. Gross, Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *J Neurophysiol*, 1981. 46(2): 369–84.
2. Reynolds, J.H., and R. Desimone, The role of neural mechanisms of attention in solving the binding problem. *Neuron*, 1999. 24(1): 19–29, 111–25.
3. Reynolds, J.H., L. Chelazzi, and R. Desimone, Competitive mechanisms subserve attention in macaque areas V2 and V4. *J Neurosci*, 1999. 19(5): 1736–53.
4. Desimone, R., Welcome to the revolution: Stories about Charlie Gross and his lab. *Prog Neurobiol*, 2020. 195: 101923.
5. Zhang, Y., et al., Object decoding with attention in inferior temporal cortex. *Proc Natl Acad Sci USA*, 2011. 108(21): 8850–55.
6. Gross, C., C. Rocha-Miranda, and D. Brender, Visual properties of neurons in inferotemporal cortex of the Macaque. *J Neurophysiol*, 1972. 35(1): 96–111.
7. Desimone, R., and C.G. Gross, Visual areas in the temporal cortex of the macaque. *Brain Res*, 1979. 178(2–3): 363–80.
8. Albright, T.D., R. Desimone, and C.G. Gross, Columnar organization of directionally selective cells in visual area MT of the macaque. *J Neurophysiol*, 1984. 51(1): 16–31.
9. Desimone, R., J. Fleming, and C.G. Gross, Prestriate afferents to inferior temporal cortex: an HRP study. *Brain Res*, 1980. 184(1): 41–55.
10. Schwartz, E., et al., Shape-recognition and inferior temporal neurons. *Proc Natl Acad Sci USA*, 1983. 80(18): 5776–78.

11. Desimone, R., et al., Stimulus-selective properties of inferior temporal neurons in the macaque. *J Neurosci*, 1984. 4(8): 2051–62.
12. Kanwisher, N., J. McDermott, and M.M. Chun, The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci*, 1997. 17(11): 4302–11.
13. Tsao, D.Y., et al., A cortical region consisting entirely of face-selective cells. *Science*, 2006. 311(5761): 670–74.
14. Desimone, R., and L.G. Ungerleider, Multiple visual areas in the caudal superior temporal sulcus of the macaque. *J Comp Neurol*, 1986. 248(2): 164–89.
15. Ungerleider, L.G., and R. Desimone, Cortical connections of visual area MT in the macaque. *J Comp Neurol*, 1986. 248(2): 190–222.
16. Ungerleider, L.G., et al., Subcortical projections of area MT in the macaque. *J Comp Neurol*, 1984. 223(3): 368–86.
17. Ungerleider, L.G., and R. Desimone, Projections to the superior temporal sulcus from the central and peripheral field representations of V1 and V2. *J Comp Neurol*, 1986. 248(2): 147–63.
18. Boussaoud, D., L.G. Ungerleider, and R. Desimone, Pathways for motion analysis: cortical connections of the medial superior temporal and fundus of the superior temporal visual areas in the macaque. *J Comp Neurol*, 1990. 296(3): 462–95.
19. Boussaoud, D., R. Desimone, and L.G. Ungerleider, Subcortical Connections of Visual Areas MST and FST in Macaques. *Vis Neurosci*, 1992. 9(3–4): 291–302.
20. Boussaoud, D., R. Desimone, and L.G. Ungerleider, Visual topography of area TEO in the macaque. *J Comp Neurol*, 1991. 306(4): 554–75.
21. Distler, C., et al., Cortical connections of inferior temporal area TEO in macaque monkeys. *J Comp Neurol*, 1993. 334(1): 125–50.
22. Baizer, J.S., L.G. Ungerleider, and R. Desimone, Organization of visual inputs to the inferior temporal and posterior parietal cortex in macaques. *J Neurosci*, 1991. 11(1): 168–90.
23. Desimone, R., and S.J. Schein, Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *J Neurophysiol*, 1987. 57(3): 835–68.
24. Schein, S.J., and R. Desimone, Spectral properties of V4 neurons in the macaque. *J Neurosci*, 1990. 10(10): 3369–89.
25. Desimone, R., et al., Contour, color and shape analysis beyond the striate cortex.*Vis Res*, 1985. 25(3): 441–52.
26. Desimone, R., and L. G. Ungerleider, Neural mechanisms of visual processing in monkeys, in *Handbook of Neuropsychology*, E. Boller, Grafman, J., editors. 1989, Elsevier: Amsterdam, 267–99.
27. Richmond, B., R. Wurtz, and T. Sato, Visual responses in inferior temporal neurons in awake Rhesus monkey. *J Neurophysiol*, 1983. 50(6): 1415–32.
28. Spitzer, H., R. Desimone, and J. Moran, Increased attention enhances both behavioral and neuronal performance. *Science*, 1988. 240(4850): 338–40.
29. Reynolds, J.H., T. Pasternak, and R. Desimone, Attention increases sensitivity of V4 neurons. *Neuron*, 2000. 26(3): 703–14.
30. Desimone, R., and J. Duncan, Neural mechanisms of selective visual attention. *Annu Rev Neurosci*, 1995. 18: 193–222.

31. Kastner, S., et al., Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science*, 1998. 282(5386): 108–11.
32. Kastner, S., et al., Modulation of sensory suppression: implications for receptive field sizes in the human visual cortex. *J Neurophysiol*, 2001. 86(3): 1398–411.
33. Reynolds, J.H., and R. Desimone, Interacting roles of attention and visual salience in V4. *Neuron*, 2003. 37(5): 853–63.
34. Chelazzi, L., et al., A neural basis for visual search in inferior temporal cortex. *Nature*, 1993. 363: 345–47.
35. Chelazzi, L., et al., Responses of neurons in inferior temporal cortex during memory-guided visual search. *J Neurophysiol*, 1998. 80: 2918–40.
36. Chelazzi, L., et al., Responses of neurons in macaque area V4 during memory-guided visual search. *Cereb Cortex*, 2001. 11(8): 761–72.
37. Reynolds, J.H., and D.J. Heeger, The normalization model of attention. *Neuron*, 2009. 61(2): 168–85.
38. Bertini, G., et al., Visual responses to targets and distracters by inferior temporal neurons after lesions of extrastriate areas V4 and TEO. *Neuroreport*, 2004. 15(10): 1611–15.
39. De Weerd, P., et al., Loss of attentional stimulus selection after extrastriate cortical lesions in macaques. *Nat Neurosci*, 1999. 2(8): 753–58.
40. Buffalo, E.A., et al., Impaired filtering of distracter stimuli by TE neurons following V4 and TEO lesions in macaques. *Cereb Cortex*, 2005. 15(2): 141–51.
41. De Weerd, P., R. Desimone, and L.G. Ungerleider, Generalized deficits in visual selective attention after V4 and TEO lesions in macaques. *Eur J Neurosci*, 2003. 18(6): 1671–91.
42. Fries, P., et al., Modulation of oscillatory neuronal synchronization by selective visual attention. *Science*, 2001. 23: 1560–63.
43. Womelsdorf, T., et al., Gamma-band synchronization in visual cortex predicts speed of change detection. *Nature*, 2006. 439(7077): 733–36.
44. Womelsdorf, T., et al., Modulation of neuronal interactions through neuronal synchronization. *Science*, 2007. 316(5831): 1609–12.
45. Fries, P., et al., The effects of visual stimulation and selective visual attention on rhythmic neuronal synchronization in macaque area V4. *J Neurosci*, 2008. 28(18): 4823–35.
46. Bagherzadeh, Y., et al., Alpha Synchrony and the Neurofeedback Control of Spatial Attention. *Neuron*, 2020. 105(3): 577–87 e5.
47. Mendoza-Halliday, D., et al., A ubiquitous spectrolaminar motif of local field potential power across the primate cortex. *BioRxiv*, 2022.
48. Miller, E.K., L. Li, and R. Desimone, A neural mechanism for working and recognition memory in inferior temporal cortex. *Science*, 1991. 254(5036): 1377–79.
49. Miller, E.K., and R. Desimone, Parallel neuronal mechanisms for short-term memory. *Science*, 1994. 263(5146): 520–22.
50. Miller, E.K., C.A. Erickson, and R. Desimone, Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J Neurosci*, 1996. 16(16): 5154–67.
51. Suzuki, W., E. Miller, and R. Desimone, Object and place memory in the macaque entorhinal cortex. *J Neurophysiol*, 1997. 78(2): 1062–81.

52. Erickson, C.A., and R. Desimone, Responses of macaque perirhinal neurons during and after visual stimulus association learning. *J Neurosci*, 1999. 19(23): 10404–16.
53. Zhou, Y., et al., Atypical behaviour and connectivity in SHANK3-mutant macaques. *Nature*, 2019. 570(7761): 326–31.
54. Han, X., et al., Millisecond-timescale optical control of neural dynamics in the nonhuman primate brain. *Neuron*, 2009. 62(2): 191–98.
55. Acker, L., et al., FEF inactivation with improved optogenetic methods., 2016. 113(46): E7297–E7306.
56. Acker, L.C., et al., Large volume, behaviorally-relevant illumination for optogenetics in non-human primates. *J Vis Exp*, 2017(128): 56330.
57. Gregoriou, G.G., et al., High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science*, 2009. 324(5931): 1207–10.
58. Rossi, A.F., et al., Top down attentional deficits in macaques with lesions of lateral prefrontal cortex. *J Neurosci*, 2007. 27(42): 11306–14.
59. Gregoriou, G.G., et al., Lesions of prefrontal cortex reduce attentional modulation of neuronal responses and synchrony in V4. *Nat Neurosci*, 2014. 17(7): 1003–11.
60. Zhou, H., R.J. Schafer, and R. Desimone, Pulvinar-cortex interactions in vision and attention. *Neuron*, 2016. 89(1): 209–20.
61. Lowet, E., et al., Enhanced neural processing by covert attention only during microsaccades directed toward the attended stimulus. *Neuron*, 2018. 99(1): 207–14 e3.
62. Srinivasan, K., et al., Stimulus representations in visual cortex shaped by spatial attention and microsaccades. *BioRxiv*, 2023.
63. Bichot, N.P., A.F. Rossi, and R. Desimone, Parallel and serial neural mechanisms for visual search in macaque area V4. *Science*, 2005. 308(5721): 529–34.
64. Zhou, H., and R. Desimone, Feature-based attention in the frontal eye field and area V4 during visual search. *Neuron*, 2011. 70(6): 1205–17.
65. Bichot, N.P., et al., A source for feature-based attention in the prefrontal cortex. *Neuron*, 2015. 88(4): 832–44.
66. Baldauf, D., and R. Desimone, Neural mechanisms of object-based attention. *Science*, 2014. 344(6182): 424–27.
67. Xu, R., et al., The cortical connectome of primate lateral prefrontal cortex. *Neuron*, 2022. 110(2): 312–327 e7.
68. Mendoza-Halliday, D., H. Xu, and R. Desimone, Dissociable neuronal substrates of visual feature attention and working memory. *BioRxiv*, 2023.