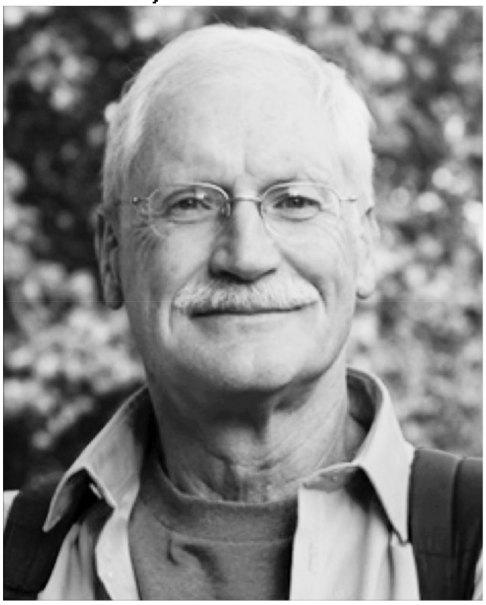
Wolfram Schmidt



Wolfram Schultz

Born:

Meissen, Germany
August 27, 1944

Education:

Dr. Med. University of Heidelberg, Germany, 1973
Habilitation (PhD) in Neurophysiology, University Fribourg, Switzerland, 1981

Appointments:

Postdoc, MPI Biophysical Chemistry, Göttingen, Germany (1973)
Postdoc, State University of New York at Buffalo, United States (1975)
Postdoc, Karolinska Institute, Stockholm, Sweden (1976)
Lecturer–Professor Neurophysiology, University Fribourg, Switzerland (1977–2001)
Wellcome Trust Principal Res Fellow, University Cambridge, United Kingdom (2001–present)
Professor of Neuroscience, University Cambridge, United Kingdom (2002–present)
Visiting Professor, Tamagawa University, Japan (2002)
Fellow, Churchill College, Cambridge, United Kingdom (2004–present)
Visiting Associate, California Institute of Technology, United States (2004–present)
Visiting Professor, Tohoku University, Sendai, Japan (2018)

Honors and Awards (Selected):

Ellermann Prize, Switzerland (1984)
Brooks International Lecture, Department of Neurobiology, Harvard University (1993)
Theodore-Ott-Prize, Switzerland (shared) (1997)
Golden Brain Award (2002)
Pfizer Lecture, Society of Neuroscience, San Diego (2004)
Ipsen Prize for Neuronal Plasticity (shared) (2005)
Fellow, Royal Society, United Kingdom (2009)
European Journal of Neuroscience Award—FENS Award (2010)
Kavli Presidential Lecture, Society of Neuroeconomics (2010)
Zülch Prize, Germany (shared) (2013)
Brenda Milner Lecture, Montreal Neurological Institute (2014)
Qi Zhen Global Lecture, Zhejiang Univ, Hangzhou, China (2014)
EBBS/Behavioral Research Prize Lecture, FENS Forum, Copenhagen (2016)
Plenary Lecture, 39th Annual Meeting, Japan Neuroscience Society, Yokohama (2016)
Brain Prize (shared) (2017)
Caltech Chen Distinguished Lecture (2017)
Erlanger Distinguished Lecture, San Diego (2018)
Volker Henn Lecture, University of Zürich (2018)
Gruber Prize (shared) (2018)
Waelsch Lecture, Columbia University (2018)
Karl Spencer Lashley Award (2019)
AV Hill Lecture, Cambridge, United Kingdom (2021)

Wolfram Schultz has combined behavioral and neurobiological techniques to investigate the neuronal mechanisms of reward processing and economic decision-making. He has used behavioral concepts from animal learning theory and economic decision theories to study individual neurons and global activity of the dopamine system, striatum, orbitofrontal cortex, and amygdala. He and his group discovered the dopamine reward prediction error signal that indicates how much a reward is better or worse than predicted. They also described other reward, risk, and decision signals in the striatum, frontal cortex, and amygdala.

Wolfram Schultz

Reward is in my genes. My maternal great-grandfather Carl A. Beegen demonstrated this fact by running a wine shop and spirit distillery in Meissen, Germany, many decades before I was born there. The genes were passed on to his son, my maternal grandfather, who took over the booze shop. He initially emigrated to the United States on the occasion of the 1904 World Fair in St. Louis. But his mother could not accept the absence of her beloved son and lured him back by claiming a serious health scare. When that health scare turned out to be a fake, my grandfather became unhappy. But by now he had met a woman with whom he ultimately fathered 12 children, only six of whom survived their first year, and one of whom became the mother of my younger brother and me. But my grandfather's new wife would in no way emigrate. So, given the fake health scare, my grandfather coerced his father into selling him the shop for the money he had made in St. Louis. He then built a flourishing business, importing wine by the barrel on ships and river barges from Bordeaux all the way to Meissen, which borders the Elbe River. He also distilled spirits in the basement of the shop for his own brands, all legal. He was very social, had a small tasting corner in his shop, and offered free wine to the many unemployed workers in 1920s and 1930s Germany. My paternal grandfather was slightly different. He was a colonel and also into reward but not the liquid variety. He married a much younger woman who also happened to be the daughter of a Royal Saxonian Major General whose hunting ground was next to that of Frederick Augustus III, King of Saxony (i.e., before the monarchy was toppled in 1918). The two noble gentlemen occasionally shot bucks on the other's range and then moved them slightly over into their own territory. With his young wife, the colonel fathered half a dozen children most of whom survived beyond childhood. One of them was my father, Robert.

Ultimately, I was born in 1944, also in Meissen. At the time, my father had followed the footsteps of his military pedigree and was captain of the infantry and was mostly stationed in Russia during the Second World War. He was caught by the Russians at the end of the war in Germany, transported back to Russia, interned in Siberia, and sent back home in 1946 because his abdominal scar from a skin transplant for a lost finger was mistaken as a potentially troubling stomach problem. He joined his wife Herta, my mother, said hello to me, and fathered my brother who was born in 1947 and is now a retired physics professor and former head of a large metal physics research laboratory at the University of Dresden. My father had been trained before the war as an accountant with Siemens, the big electro company. I inherited his attention to detail and perseveration,

which served me well in my later laboratory life. My mother, encouraged by her socially responsible, wine- and booze-selling father, was one of the few female students of German, English, and history at the Universities of Leipzig and Vienna; there she occasionally heard about gender discrimination but was overall happy. She became a high school teacher before and during the war but afterward was a housewife to raise my brother and me. She was surely a lost talent, and very supportive, but staying home for a married woman with children but without servants was necessary at a time when cars, supermarkets, central heating, washing machines, and refrigerators did not exist or were unaffordable.

My parents were fed up with having another dictatorship and with the poor educational prospects of their non-proletarian offspring in communist Eastern Germany and moved to Western Germany. So, in October 1950, I was thrown over a two-meter-high fence into West Germany at a moment when the armed fence patrol had reliably passed on their hourly rounds. There was mention of border guards having shot dead another refugee a couple days earlier, but we were not deterred. We made it without problems, like thousands of others, took a train to a small town in northern Bavaria (Bad Neustadt) and settled where my father had been hired by his former boss in his former company, Siemens.

Getting Started

Early Learning

I was neither a great pupil nor a proficient high school student. The stuff was just too much routine and boring, and my performance was borderline. My mother, as a former high school teacher, was shocked but nevertheless installed in me a desire for academic work. When I was in danger of not surviving the final high school exam required for university, she persuaded me to think further and continue the boring school. To address my weakest point, she arranged for instructions in Latin from a private and expensive teacher. Her leverage was my interest in becoming an automotive engineer, a popular job in car-crazy Germany, which at the time required a university education and thus a high school diploma, and for that I needed a pass in Latin. At the final exam, after eight years of studying the language, the examiners let me succeed after I had promised to not become a Latin teacher; they were sympathetic but did not want to compromise their teaching reputation. I soon lost interest in car engineering, but I am eternally grateful to my mother for having kept me on a path that kept all education options open; she had no easy job. While my father often expressed shock about my undisciplined behavior, he was liberal enough to let me explore the world instead of doing boring homework (he himself had not been a great high school student, so that is also in my genes).

Indeed, the German high school at the time was fantastic, because it ended at noon or shortly thereafter. If you could escape homework without completely blowing school, you could live a decent life full of self-driven activity and exploration. I found plenty of attractions on the sports grounds, where I learned persistence in the face of failure: I tried very hard, for several years, to run the 100-meter dash faster than 12.5 seconds but never succeeded. My biggest sports achievement was to take second place in the 200-meter breaststroke in a regional swimming championship, which eliminated another potential career, that of sports teacher. But the benefit was that I became resilient and enduring. Learning to do something for its own sake, and trying hard, irrespective of expected glory, helped me throughout my life to undertake seemingly useless tasks that at some point might pay off.

In the winter, I was attracted to metal constructions after we had found an abandoned and incomplete but large Märklin metal construction set. I constantly was constructing something, including a huge tower with blinking lights that were visible through the window from the street. On visits to my grandparents, who had remained in Eastern Germany, my grandfather bought me and my brother model trains that, unfortunately, were of the poor quality of the time and hardly ever worked (we never had money to buy the more reliable Märklin trains). They required constant mechanical and electric adjustments, which taught me mechanics, electricity and, maybe most important, to never give up. Ultimately, I succeeded in constructing an operation of several trains running automatically in parallel and alternatingly (when they worked).

I never got rid of my obsession with exploration and novelty-seeking, and I kept avoiding repetitive tasks with prescribed procedures, unnecessary detail, and slow progress. I was good in math, which was easy, but I lost interest because of its repetitive and procedural nature. When my math grades declined, I worked on math for a few days, and my mother was worried about my monomaniac insistence; I lost interest again when my marks had recovered. Except for sports, basically all other disciplines bored me. Despite all this jittering, I could focus very well, with great persistence, when I liked tasks that were interesting and captivating, like the eternal tinkering with the never-working model trains, and I would not stop until I was content with the outcome. I understood later that the combination of these behaviors may inadvertently expose one to rare events and lead to novel experiences. It can be formalized as exploration and positive skewness seeking. The epsilon-greedy exploration model of reinforcement learning of Sutton and Barto (1989) might apply; rather than always seeking the best-known option, a small percentage (epsilon) of choices are made randomly, which occasionally leads to unpredicted and unknown outcomes that are much better than the best-known option and are worth the many failures. Anyway, that was my not-so-great school life until I ran into the idea to do a year of high school exchange.

A Healthy Interruption

After the war, various organizations brought German high school students for a year to the United States. For me, this was another opportunity to experience the unknown and nonrepetitive. I was selected by Youth for Understanding, for which Ulrich Irmer, the later German and European parliamentarian, did the screening in the American Cultural Mission (*Amerika Haus*) in Munich and in his own, slightly bohemian student flat. Thus, I left Europe in the summer of 1962 on a four-engine Super Constellation propeller plane from the Flying Tiger charter line (reportedly owned by the CIA). We flew from Amsterdam for almost 24 hours to Detroit, with intermediate landings in Ireland and Newfoundland. I entered a welcoming family of a gas station owner (the father) and special education teacher (the mother) in Mayville in the "thumb" of Michigan. The father was very practical and had built, with his own hands, six or seven private houses, his own included, and the mother with her patience and knowledge was very reassuring. This was real life; a small village with a main street that was deserted during the day because everybody was working, a run during the morning school break to the dime store for sweets, work at the Mobile gas station on Wednesday afternoons and all day Saturday, and the oldest daughter married to a dairy farmer (what a nice guy, actually both of them). A number of lakes nearby allowed me to practice water skiing in the summer and to go ice skating and watch motorcycle races on ice in the winter. I also enjoyed American football and track, although with my usual moderate success in sports.

The U.S. experience provided quite a contrast to postwar Germany with its painful recovery from misleading demagogic warfare and industry-level murder. But what amazed me most, and would become the most lasting influence, was the unpresumptive attitude and openness. By attending the local Methodist church with my host family, while not being religious at all, I met the preacher, Reverend Bill Lutz, and struck a great friendship. We discussed the state of the world on many Sunday evenings, starting with his sermon and widening into the history of the United States and the benefits of democracy. It was just the right moment, when everybody was optimistic, everything seemed possible if you had enough initiative, and no political murders had yet taken place. Reverend Lutz's broad intellect and profound knowledge, in the middle of the countryside, was impressive. At the end of the school year, he suggested that I take a trip through the United States. He picked up the phone a few times, called up fellow Methodist and Presbyterian preachers throughout the west and south of the country, and arranged a tour for me through Wyoming, Oklahoma, California, Arizona, Texas, and Louisiana, using a Continental Trailways bus ticket one of my German uncles had donated. Afterward, I went back to my old high school in Germany and finished the last year in 1964 with new motivation and moderate success.

Another Healthy Interruption

With my broad and unfocussed interests, I had no idea about a future profession. The fresh high school diploma and my mother's guidance toward academia both suggested that I should attend university. But what discipline? The decision was not imminent because of mandatory military service (which was only abandoned many years later by Angela Merkel's government). The standard duration was 18 months, but I received a regular salary for doing 24 months that could supplement my future government studentship. The military produced another welcome delay for my professional indecision. I passed a comprehensive entrance exam with more points than anyone else had scored in the past several years and then was thoroughly interviewed by a suspicious psychologist. I could only explain that my performance might be due to my extracurricular interests, which convinced him that I was not nuts. Thus, I went for 24 months to the Air Force. Rather than flying, I served in the ground troops because of the short duration and my less-than-perfect eyesight. I was stationed in the Netherlands (on a NATO base), Munich, northern Germany and Bavaria again, regions I had not known before. I found that the boring military routine removed the idiosyncratic routines, unconscious prejudices, and reduced motivation that I had acquired in high school. The biggest benefit of all was that I met a lovely, and ultimately irresistible, girl in a club in Munich in 1966; she had cancelled a midsummer party in the Bavarian mountains because of rain and thus had run unintentionally into me in that club. Gerda was to become my wife a few years later.

University Studies

My interest in careers spanned from sports teacher (which is nice because of the exercise in fresh air and swimming pools but not realistic given my mediocre sports performance) through foreign diplomatic service (compatible with my desired international living but incompatible with my lack of patience in the face of malfunction) to law (more serious and promising, as I had done an excellent job in the law course at the German Air Force Officer's Training School). I also annoyed my parents and everyone else by proposing to become a politician, which was a true joke, almost surreal, and incompatible with my dislike for slow progress. I sometimes shudder when thinking how these options could have gone wrong. At the end, I decided on medicine, as some relatives were medical doctors, and as I wanted a job with a large spectrum of possibilities. I found biology interesting and wanted to help people; however, I was not quite so sure, as I had been public enemy number one for my high school biology teacher, who was back from an American two-year postwar de-Nazification camp and still showed remnant behaviors. And of course, medical school required excellent school grades.

But luck does exist. On one of my last trips home from the military, I had bought a newspaper and read that the University of Hamburg was crediting former soldiers one half mark for every half year of service, because the military had helped the city of Hamburg a few years earlier to recover from a surge of the North Sea and a massive flooding from the Elbe River. With my two years of service, I seemed to be eligible, and that opportunity finally made the decision for me. I was accepted by Hamburg University on the second reserve list two weeks before the term started in the fall of 1966.

The two-year military service had left my brain drained. I was even more bored in the military than in high school and was longing for something truly useful and engaging to do.

Thus, at Hamburg University, I soaked up, frantically, like a dry sponge, any knowledge and ideas I could get hold of. Others talked about parties and girls, but I was already busy with Gerda in the early stages of becoming her future husband. I attended every possible and impossible lecture in medicine, and some in psychology and philosophy, in particular the small ones with good interactions with the lecturer. I just could not get enough. I went to Carl-Friedrich von Weizsäcker's lectures on Friday evening at the end of a busy student week. He read Kant's *Critique of Pure Reason* in an overflowing auditorium of 2,000 students sitting everywhere. This was a cultural, and partly morally redeeming, event; Weizsäcker had been a nuclear scientist who had reportedly delayed his work on nuclear fission to avoid building an atomic bomb for Hitler. After the war, he had become a philosopher. His Kant lectures taught me to question dogmatic beliefs and the limits of knowledge as a prelude to Horkheimer and Adorno's critical theory, which I encountered later in Heidelberg.

I found the fundamental biology in undergraduate medicine absolutely wonderful. It was so easy to understand. Almost everything makes perfect sense, falls into place naturally, and works the way it should, as if it had been tuned to optimal function. I lived in a nice student housing and had many discussions that helped us bring everything down to the essentials, including physics that I had forgotten from high school and, most interestingly, physiology that was just beautiful if you could boil it down to the essential laws. As Einstein says, "if you cannot explain it in simple words, you have not understood it well enough." My teachers and fellow students had understood it. Then my "career" in neuroscience began in 1968 when I asked Professor Duncker about working in his neurophysiology laboratory. He immediately suggested that I do my medical thesis, which is like a master's for medics but conveys the title of medical doctor and allows one to do research without a further doctorate. While exploring my possibilities, I also had tried a lab stunt in the biochemistry department of the natural sciences faculty, but medical students were considered inferior to natural science students, and the lecturer told me to come back when I had done my formal lecture course (which turned out to be super simple and all written

out clearly in an excellent textbook). Thus, I became a neurophysiologist and not a molecular biologist.

After my medical undergraduate exam, I wanted to move to a less sober environment and went to cozy Heidelberg. It was the only university, apart from Hamburg, that would accept my high school grades for clinical medical studies. The city is absolutely delightful, a classic center of poetry, literature, and philosophy, like the German Romanticists around Hölderlin. There is the philosopher's path on the mountain overlooking the Neckar River, a pretty old town (infested by tourists but beautiful on rainy days when nobody is around), and a classic university that later turned out to be partnered with Cambridge on the basis of its many similarities. Besides my clinical medicine, I enrolled in undergraduate mathematics because I figured that the brain is so complex that math could provide reductionist formalizations, and I attended a few select philosophy lectures in the old town. There I met Harald Pilot who was a senior philosophy student and who served as a tutor for younger students like me. His simple and profound explanations of the philosophy of science and critical theory, which continued informally in the pub after his evening sessions, left a profound desire to search for unifying concepts underlying the complexity of nature. Without him and the Heidelberg philosophy environment, I probably would not have used the formal theories of economics for 20 years in my neurophysiology research in Cambridge. I also learned from these philosophy seminars how to run a contemplative discussion that mixes formal issues with exploration and allows one to pursue unexplored arguments and to make mistakes, which can be quite fruitful. These experiences became very helpful for the interactions in my later research group and prevented me from running sterile report-to-the professor meetings. I was also somewhat engaged in the political student activities of the time where we read volume 1 of Karl Marx's *Capital*, learned to doubt the unmerited authority of people and concepts, and used critical theory to search for the big picture, to get to the bottom of problems rather than simply understanding and superficially fixing them, and to question and reflect on the usefulness, ethics, and wider implications of our actions, all of which helped me avoid many unnecessary steps in life and research. Studying medicine in Heidelberg was also a cultural enterprise; I was attracted by the small evening lectures of Professor Tellenbach who showed us the relationship between psychiatric disease, culture, and history. It was a truly great time! Of course, things go well when one is in resonance with the world: I did the best undergraduate medical exam in Hamburg and the second best in Heidelberg out of several hundred students; I never realized the value of it until I saw applicants to my lab boasting about such marks. In the end, I did a practical year in which I spent several months in neurosurgery and neurology and found the textbook by the renowned neuropsychologist Klaus Poeck to be very stimulating and beautifully written. I wrote long and detailed patient reports after

hours of examination and was close to becoming a neurologist and neuropsychologist myself. But I wanted to try neurophysiology first.

Postdoc Life

Without much connection to real science, I asked fellow Heidelberg students with lab experience for advice and was pointed to Otto Creutzfeldt, the son of the neurologist who defined Creutzfeldt-Jakob disease. Otto Creutzfeldt had done pioneering intracellular recordings in the visual cortex at a Max-Planck Institute in Munich and was in the process of moving to a new MPI in Göttingen. During a brief visit with him in Munich, he asked how many years I had taken to study full medicine and undergraduate math (I took only the minimum time required for medicine and did math in parallel), and he accepted me as a postdoc in his new department. I was amazed by his neuroscience perspective and his international horizon. He had a whole Max-Planck department in a new building overlooking the city and a dozen groups working on all forms of sensory processing—very up to date. I worked with Christian Hellweg between 1973 and 1975 on intracellular (whole cell) recordings in the cat somatosensory cortex and natural mechanical vibrissae stimulation. To learn such a sophisticated technique, I pestered colleagues in the lab until it worked and am much indebted to Klaus Albus and Uli Kuhnt for their tolerance of this impatient newcomer. In the course, Sir John Eccles visited and was looking for a postdoc for his lab in Buffalo. Creutzfeldt suggested me, we talked, and off we went to Buffalo with Gerda and our one-year-old son Johannes in the snowy spring of 1975.

In Eccles' lab in Buffalo, I worked with Gary Allen on connections between the cortex and the cerebellum in anesthetized monkeys, using electrical stimulation of the motor cortex and single-neuron recordings in the dentate nucleus of the cerebellum. This was connectionist work very similar to current experiments that use optogenetic stimulation, except that we had no idea about the transmitter identity of the stimulated cortical and recorded cerebellar neurons. But the results showed that the anatomically demonstrated pathway from cortex via pons and inferior olive was functional. Gary had bought a primate chair for trying out experiments on awake behaving monkeys but never got around to using them because he became a Christian missionary. The monkey chair together with the ensuing discussions stimulated my interest in this elegant work. Watching single neurons in real time while a monkey performs in a controlled behavioral task seemed the most direct and precise way to watch the brain in action. I did a further experiment on the cerebellum with Erwin Montgomery in which we stimulated the cerebellar nuclei and investigated the resulting muscle contractions in anesthetized monkeys. We discovered an output pathway from the dentate nucleus (which was a part of the neocerebellum and primarily linked to the cortex) that bypassed the cortex and went through the brainstem, suggesting

an evolutionary ancient cerebellar output to behavior. Eccles closed his lab during this period, and I needed to find another postdoctoral opportunity before starting my own research. Having experienced a wonderful introduction into classical neurophysiology, I was attracted to the new field of neuropharmacology that defined brain function on the basis of neurotransmitters instead of electrical brain activity. I contacted Urban Ungerstedt in Stockholm, who obtained a fellowship for me to introduce neurophysiology into his lab, and that is how I got acquainted with the dopamine system.

We arrived with Gerda, Johannes, and our few weeks' old son Thomas on a sunny winter morning in 1976 in beautiful and serene Stockholm. I conducted a few straightforward neurophysiology experiments in Urban's lab that would complement his pioneering histological and behavioral work on rats. The neurophysiology of dopamine neurons was just starting to be explored, and its sophistication did not compare to that of the well-developed motor control field. I recorded from striatal neurons in rats whose dopamine system had been lesioned at different time points. The striatal neurons showed elevated activity a few days after the lesion but no longer a year later. The literature was very limited, and I was able to follow the complete literature on dopamine electrophysiology, anatomy, and a few extra things, like neuronal plasticity and adaptation. This seems unbelievable now when interesting paper after interesting paper appears each week, and it is quite impossible to keep up with fields outside one's own perspective. In any case, Urban and his Stockholm lab gave me the opportunity to get fully acquainted with the state of the dopamine field at the time and thus provided me with an excellent start for one of my major scientific avenues.

The Dopamine Reward Signal

My Plans

With my Stockholm introduction into neurophysiology based on neurotransmitters rather than anatomy, I wanted to start my own work and moved in 1977 into a junior faculty position at the university in the lovely city of Fribourg in Switzerland. From my medical studies, boosted by the short stunt in neurology, I was familiar with Parkinson's disease that arises from the degeneration of midbrain dopamine neurons. From my postdoctoral stay with Gary Allen in Eccles' Buffalo laboratory I was keen to enter motor neurophysiology and monitor single neurons while monkeys performed well-controlled movements. This research interest had been boosted by Eccles' almost daily lunch discussions about Free Will in the cafeteria of the Amherst campus at Buffalo, where he emphasized the importance of studying self-initiated movements in which free will might be expressed in an objectively measurable way. With the cortex-centered view of neuroscience at the time, Eccles' focus was not on dopamine neurons but on the supplementary motor area (SMA) where he thought such neuronal activity

should first arise, and Ben Libet's readiness potential in humans supported this notion (1978). But deficits in spontaneously initiating movements also constituted a substantial symptom of Parkinson's disease. To address these aspects in combination, I thought to first establish recordings from single dopamine neurons in awake behaving monkeys and to see how their activity might be related to the Parkinson's deficits; then I thought to take a stab at the SMA to find activity during spontaneous movement initiation.

However, trying to link dopamine signals to specific behaviors was not without risk. Sure, the function of dopamine neurons in movement was much less known than that of motor cortex, cerebellum, and much of the basal ganglia, and any data in this respect would be interesting. But Parkinson's patients got better by administration of dopamine drugs that cannot replicate subsecond neuronal changes. So many knowledgeable people reasoned that dopamine was simply a chemical that was necessary to prevent Parkinsonian symptoms but might not show a relationship to specific behaviors. It might be a necessary ingredient of the brain's chemistry, something like a "rain on the brain" or a "fuel" that was necessary for other brain structures to control behavior, a neuromodulator that does nothing on its own and only enables other processes.

Building My Own Lab

To get started in Fribourg, I visited Uli Büttner and Volker Henn at the University Neurology Clinic in Zurich in 1977. They studied oculomotor nuclei and explained to me in one day how to record from individual neurons in the midbrain of monkeys. We are still using some of their techniques today, and I am eternally grateful for their help and was much honored to be invited to present the Volker-Henn-Lecture in 2018 in Zurich. I went back to Fribourg and started my new monkey lab in an empty room with little money but with the help of two young workshop members named André Gaillard and Edouard Regli. They were absolutely wonderful, ingenious, super precise, and supportive people who liked to create something interesting. To celebrate the new lab, I designed a little demonstration of the work on reward I was planning to do. I put a bottle of brandy into the liquid-dispensing bottle, programmed the liquid valve to just fill just one small brandy glass with a single key touch, and off we went to a very happy (human) party that became the talk of the place for many years to come.

Searching for a Dopamine-Specific Neuronal Discharge

From my teenage mechanical and electrical tinkering, I had gained plenty of hands-on experience, including resilience from repeated failure. As good commercial metal microelectrodes were unavailable at the time, I learned to make my own glass-insulated tungsten recording electrodes under the

microscope, which was a feast on quiet evenings in the lab. These electrodes turned out to be crucial for recording from dopamine neurons; the impedance-reducing platinization of their tips was important for reducing the background noise that made neuronal discharges more recognizable. There are certain things in life that you have to do yourself, and electrodes with micrometer precision tips are good examples. Tungsten is the English word for a hard metal whose chemical element name is Wolfram (W, atomic number 74, melting point 3,422 C); some of my friends call me Tungsten (as did a Swiss databank website of scientific talents). So, I thought I might have a native relationship to tungsten electrodes and thus have used them for all 40 years of my work in neurophysiology.

While working in Stockholm, I had run into the work of George Aghajanian in the 1970s who had reported that rat midbrain dopamine neurons discharge extracellularly recorded action potentials with a distinct waveform (Guyenet & Aghajanian 1978). I tried this while still in Stockholm and confirmed on the first recording day the wonderful positive-negative-positive waveform with a little break in the upgoing flank and the characteristic blop-blop-blop in the loudspeaker. Without optogenetic identification at the time, the waveform would be crucial to identify dopamine neurons in awake monkeys in which we needed to test antidromic stimulation or systemic pharmacological injections.

To become familiar with dopamine recordings in monkeys, I implanted a recording chamber that allowed me repeated electrophysiological testing under anesthesia without worrying about the animal's behavior. The first challenge came when locating dopamine neurons at 30 millimeters below the cortical surface. Stereotaxic coordinates in macaque monkeys are notoriously inaccurate, which we overcame by locating the face and head area of the ventromedial somatosensory thalamus that are positioned above the lateral region of dopamine neurons. The thalamic neurons had fantastic strong and precise responses to the touch of single facial hairs or a single tooth or the gingiva; the responses were a beauty for any electrophysiologist and a reason to call in departmental colleagues to illustrate what dry textbooks claimed. From then on, on every recoding day, we would pass our electrode through the thalamus before going for lateral dopamine neurons below.

After several months of weekly recording in the anesthetized monkey, I had developed the electrodes I needed to confirm that I could see in the recordings all the essential characteristics: the characteristic waveform with a systematic polarity change associated with electrode movement, the blop-blop-blop sound, a reduction of impulse rate by systemic apomorphine, as is typical for dopamine neurons, and identification of striatum-projecting dopamine neurons by antidromic responses. Then I tested the same monkey sitting awake in a monkey chair and confirmed the striking electrophysiological characteristics. Now I was ready to record from neurons that used

the neurotransmitter dopamine during well-controlled behavioral tasks in macaque monkeys, taking advantage of their exquisitely controllable behavior. While I always remained anxious about the identity of our dopamine neurons, recent optogenetic identification of monkey dopamine neurons revealed the same waveforms (Stauffer et al. 2016). My moment of gratitude came on a psychopharmacology conference in 2007 in Boca Raton, Florida, where I happened to sit opposite George Aghajanian in the hotel bus and could tell him how much of a game-changer his papers on identified dopamine recordings had been for me.

I was working on my own in Fribourg and had all the time in the world to explore what dopamine neurons were doing during behavior. Nobody else was doing similar recordings, so there was no pressure or competition, and hardly anybody took notice of my work at conferences. I benefited from the low-level but permanent faculty position and the modest but pretty reliable funding from the Swiss National Science Foundation. Great, leisurely moments with lots of time for imagination and trying things out, a pleasant and ultimately very productive period.

Surprise 1a: Phasic Movement-Related Dopamine Activity

Before going any further, the first tests of dopamine neurons in the awake monkey had to address the potentially discouraging issue of "rain on the brain," according to which there might not be any phasic, subsecond changes in dopamine neurons related to behavior. Even before perfecting my electrodes, I designed an experiment based on the known Parkinsonian motor and cognitive deficits. Together with my first graduate student, André Ruffieux, who later became a respected neurologist in Fribourg, we trained monkeys in a then-state-of-the-art delayed go/no-go task that included many sensory, movement, and cognitive processes appropriate for wide screening of behavioral relationships. Indeed, variable fractions of dopamine neurons showed diverse, excitatory, and inhibitory changes with reaching or mouth movements, during reward reception, and even during the whole task (Schultz et al. 1983). Thus, dopamine neurons were more than "rain on the brain," and we might even have found a positive correlate with Parkinsonian deficits. We immediately submitted a paper to *Science* that was, just as immediately, rejected. This was fortunate, as Mahlon DeLong, a pioneer and authority on the motor neurophysiology of the primate basal ganglia, and a friend, reported an absence of dopamine changes in his more controlled and concise arm movement task (DeLong et al. 1983). Their result fit well with the "rain on the brain" notion of dopamine function. How embarrassing: The first findings on movement relationships in a novice laboratory were not reproducible in a world-class laboratory!

Being nevertheless convinced of the validity of our dopamine movement relationships, the only explanation we had was that Mahlon's study used

much more constrained arm movements, whereas our monkeys performed a large arm reaching movement that engaged more than 35 hand, arm, and shoulder muscles plus several muscles on the back, as our widespread muscle recordings had shown. We had already cautiously interpreted our data as reflecting behavioral activation rather than motor control, and that still seems to be the best explanation. We then used the muscle recordings to refine our tasks and to reduce muscle activity, and now we rarely found movement-related dopamine changes. We also tested a wider variety of movement situations, including eye movements and, separately, more sluggish arm movements following the offset of a visual stimulus rather than its onset (Schultz & Romo 1990). None of these tests showed any dopamine activation. Now we had what Mahlon had seen: Dopamine neurons, despite their crucial role in Parkinson's akinesia, are not activated during well-controlled movements. Still, the irreproducibility of our earlier movement activation remained a sensitive issue.

Then Karl Deisseroth, Ed Boyden, and their colleagues developed optogenetics around 2010. They used rodents because of their suitability for molecular biology, but rodents are not monkeys and have more difficulty controlling their difficult-to-monitor movements. Also, rodents are often tested during whole body movements in boxes, or during locomotion in T-mazes, or when running on treadmills or wheels, all of which engage hundreds of muscles and sensory receptors in the skin, joints, and muscles and lead to substantial behavioral activation. Although these might be straightforward tasks for behaviorists, they are confusing for behavioral neurophysiologists. Even seemingly simple licking movements can be accompanied by uninstructed movements and unintended neuronal activity, as Anne Churchland's laboratory at Cold Spring Harbor showed (Musall et al. 2019). When testing neurons in rodents, with or without optogenetics, an early voltametric study (Stuber et al. 2005)—and later several high-profile neurophysiological studies—reported time and again dopamine activations during these movements (Berke 2018; Engelhard et al. 2019). This was a relief: Finally, someone confirmed our movement-related activity, which we ourselves had difficulty doing when studying better controlled movements. Dopamine neurons would still not code individual muscle contractions or movements but would get all excited when the animal became active. Could this be what was causing Parkinsonian hypokinesia: not specific motor control but rather a lack of behavioral activation by dopamine neurons?

Surprise 1b: Phasic Dopamine Stimulus Response

The movement-related changes were statistically significant but were nevertheless quite modest, resulting rarely in more than 20–50 percent increase of activity. I just was not convinced that this was all. Maybe the task in our study, and also in Mahlon's study, contained too many

stimuli and movements. At a conference in 1978 in Denmark, I expressed my worries to Susan Iversen who suggested a more natural, scaled-down task to get to the bottom of things. Instead of the stimuli from the previous task, I used a new custom-made food box whose front door (40 by 40 millimeters) would open rapidly upward through an experimenter-operated release cable (400 degrees/second); then the animal released a touch-sensitive resting key, reached into the food box, retrieved a raisin or a morsel of apple, brought it to its mouth, and ate it—all quite intuitive and natural. Infrared photo beams detected the door opening and the animal's hand entering the box.

On my first dopamine recordings with this task, something happened that neither Mahlon nor we had ever seen before: Less than 100 milliseconds after the food box door opened, dopamine neurons showed a sharp, phasic, stereotyped increase in activity that lasted less than 200 milliseconds (Schultz 1986b). The increase consisted of just a few extra dopamine impulses that were sharply time-locked to the precise time of the door opening. Well more than half of the dopamine neurons in the substantia nigra with all their fancy electrophysiological characteristics showed this response, and so did more than 90 percent of the dopamine neurons located in the ventral tegmental area. The response of only a few extra discharges was not much in itself, and became an object of occasional ridicule, but the high proportion of dopamine neurons responding with similar latency and duration made the response a substantial population signal that was likely to get noticed by postsynaptic neurons. The response occurred with the opening of the food box and not when the animal received the reward. The response could reflect many things: general alert, salience, and attention; the sensory sight of the opening door and the food and/or the door opening sound; the initiation of an arm and hand movement into the box; or maybe the prediction of obtaining a reward inside the box. We had much more work to do before putting a name on it.

Nevertheless, the response was a surprise: The simplified task with limited stimuli had yielded a clear, phasic, rapid, consistent, and substantial subsecond signal in most dopamine neurons. Now I had two separate phasic dopamine changes: the rather sluggish and heterogeneous changes with motor activation during various task epochs seen before, and now the sharper, faster, stronger, stereotyped, and more substantial response to the food box stimulus. I drew a scheme that separated the two responses (Schultz 1986b), which despite many refinements and better specifications stood the test of time (Engelhard et al. 2019). These two dopamine changes contradicted the notion of unmodulated dopamine activity, including Mahlon's lack of task modulation: Dopamine neurons were more than stupid "rain on the brain." Sometimes when I woke up in the early hours of morning, I wondered whether I could convince the community about the existence of a subsecond dopamine signal.

Surprise 2: Reward!

Then Ranulfo Romo joined the laboratory in 1985, and we took up the idea of spontaneous movements, which were better related to Parkinsonian deficits. We covered the food box with a shield and allowed the monkey to reach into the open box from underneath to fetch a piece of apple at a time of its choosing and without a triggering stimulus. Ranulfo was great in patiently training the capricious monkeys. Indeed, some dopamine neurons showed changes with these self-initiated movements (Romo & Schultz 1990), which was a nice confirmation of the movement activity observed earlier (and which also removed the lump in my stomach caused by Mahlon's negative results). Still, when we opened the food box door in control trials, many more of the same dopamine neurons showed the familiar, much sharper response.

With Ranulfo's enthusiasm, we asked whether this phasic response to the food box door opening might reflect more than simple alertness or stimulus-driven attention, as someone suggested politely. To start, I went home and asked my wife. After my lengthy description, Gerda suggested a German word for incentive (*Anreiz*): As the food box opening allowed the animal to capture a piece of food, the neuronal response might simply reflect the incentive for getting a reward. Based on Gerda's common-sense suggestion, we looked at Bindra's description of "incentive motivation" (1968), which combines "incentive" (to do something) with "motivation" (the reason for doing something, like getting a reward). Another possible term is "incentive salience," which captures the alert and attention function expressed in our earlier minimal-assumption interpretation; we later identified an initial attentional component of the phasic dopamine response (Schultz 2016). As it turned out, both "incentive" interpretations were superseded by our findings on reward unpredictability and reward prediction error (RPE) coding described below in sections "Surprise 4: Reward Unpredictability" and "From Reward Unpredictability to RPE."

To address the issue of "incentive," we wanted to see what happened when the animal encountered the food that provided the incentive. So, we stuck the food onto a touch-sensitive wire that told us precisely when the animal touched the food. As the shield prevented the animal from seeing into the box, any dopamine response could not be linked to seeing the food, and there was no stimulus that triggered the movement and could elicit a difficult-to-interpret dopamine response. Now, Eureka, the dopamine neurons responded to the touch of the food (Romo & Schultz 1990)! This was not a somatosensory touch response: On occasional test trials without food, the response was gone, or replaced by depressed activity, even though the animal touched the wire and moved its hand in the empty box (into which it could not see). Thus, there was a sharp dopamine signal that was linked to the reward and not to visual or somatosensory stimulation nor to hand movement.

This reward response was surprise 2 (after the phasic dopamine changes) and a reward for us that needed to be properly acknowledged at the end of a recording morning. We had bought a bottle of rum from a nearby supermarket and fetched a coke from the machine in the basement; we drank a bit coke off the top of the bottle, replaced it with rum, and drank a Cuba Libre to the discovery of a dopamine reward signal. As rum is lighter than coke, the two did not mix, and we drank almost pure rum, which inevitably finished the morning session in good spirits.

New data bring new problems. At the time, the definition of reward function was too unclear for precise neurophysiology. I was told about reinforcement (which includes punishment) and operant or goal-directed behavior (which might be too narrow). For the "man on the street," reward is a bonus for having done something special (also too narrow). There is also happiness (maybe the essence but hopeless to test in animals). All this was inadequate for the millisecond precision afforded by electrophysiology. The same Volker Henn whose laboratory members showed me their neurophysiology recording techniques in awake monkeys emphasized the importance of reducing the complex notion of reward to measurable variables. We had to generate not only our own microelectrodes but also the concepts for understanding the data.

It was first Roy Wise who helped a lot. I met Roy during a summer conference in 1989 in the Sardinian sea resort of Capo Boi. While the wind blew softly through the open shutters of the lunch place, we kept discussing, nonstop and repeatedly, the definitions of reward functions and their suitability for controlled testing of fast neuronal signals. It was unbelievable how his sharp mind combined with common sense dissected these inscrutable notions and came up with straightforward suggestions. We became good friends and had many more such discussion, including one during a conference organized by Ranulfo Romo in 2010 in Mexico DF. While visiting the different buildings around the great pyramid of Teotihuacan, we intermittently mixed Aztec culture with neuronal reward signals.

Back in Fribourg, I put together our own definition—rewards have three global functions: learning (as Pavlov has it), approach behavior (straightforward and intuitive) along with value for economic decision-making (requiring further concepts), and happiness (unavoidable but untestable in animals). These definitions, with many ramifications and sophistications, have proven to be very helpful to guide our work through the years.

But there were also words of caution. We had occasional visitors working on motor control whose expertise I valued highly. I told them that I wanted to test reward processing by neurons. But I was very clearly told, repeatedly, by prominent people in the field, that reward is exactly what everybody tries to factor out because it messes up their data interpretation. True, larger rewards shorten reaction times and speed up movements, so

you want to hold reward constant to keep perfect control over the movement. Fortunately, I did not listen.

But a dopamine reward function is nothing new. A researcher at the Imperial College of Tropical Agriculture in Trinidad had shown that 3,3-dihydroxyphenylethylamine, now referred to as dopamine, is involved in keeping the banana healthy (Griffiths 1959). If dopamine is oxidized, the banana gets brown, and when the enzymatic reaction is prevented and dopamine remains detectable, the banana stays nice and yellow. One might call it the "Rotten Banana Model of Dopamine Reward Function." Thus, all we were doing beyond this discovery was to associate the known dopamine reward function with neurons.

Surprise 3: Reward Prediction

While it was nice to see a response to the reward, the response to the opening of the food box still required interpretation. We found that the same dopamine neurons that responded to the food touch during spontaneous movements also responded to the door opening (Romo & Schultz 1990). Maybe the door opening response was related to the reward the animal would find in the box? That was all fine, but something was funny about these responses: The food touch response was gone when the door opening triggered the animal's reaching into the box. How can that be a neuronal reward signal when the neuron fails to respond to reward in some situations? Maybe the dopamine response to the reward had simply been transferred to the door opening, as if the neuron had done enough by responding to the door opening and no longer cared for the reward itself? It seemed most likely that the door opening was a reward predictor, in which case we would have surprise 3 (after the phasic dopamine changes and the reward response). To substantiate this assumption, we needed better controlled tests.

This is where Tomas Ljungberg came in. Tomas had done a PhD with Urban during my postdoctoral stay, at which time we had many discussions about the lack of knowledge about neurophysiological dopamine signals. He came to Fribourg 1987 with his experience in animal conditioning, which was the occasion to address the reward and prediction issues. Tomas wanted to understand stuff and was not content with our simple and intuitive tests that had been crucial for the initial detection of phasic dopamine signals. He suggested that we use concepts from animal learning theory and try a decent conditioning experiment in which we had control over the animal's behavior and could monitor the learning progress. We would use an arbitrary, intrinsically neutral stimulus and associate it with reward to make it a reward predictor.

We replaced the food box with a light-emitting diode and a lever, and the food with apple juice whose quantity was perfectly controlled by a computer-controlled electromagnetic valve that delivered drops from a spout in front

of the animal's mouth. Thus, like the more intuitive food box opening, the small light-emitting diode would become a conditioned stimulus that predicted the juice and elicited the reaching, and the animal received the juice for touching the lever. Sure enough, most dopamine neurons began responding to the light shortly after it had been presented together with the juice, around the same time the animal began to touch the lever that produced the reward. In comparison tests, the same dopamine neurons responded to the food box opening, and both responses were neither time-locked to the reaching movement nor to the muscle activity in the arm, shoulder, and back that we monitored via implanted and inserted electrodes (Ljungberg et al. 1992). How nice to reproduce our own data! Thus, our previous dopamine responses were not specific for the food box opening. Dopamine neurons did not seem to care about the kind of reward the stimulus predicted—food morsels or liquid drops.

But what about the response to the reward that we saw with Ranulfo when the animal touched the food during spontaneous movements without preceding stimuli? I was brooding over this question during a quiet weekend and asked Tomas on Monday morning whether he had ever referenced the dopamine responses to the moment of juice delivery. He came back a couple hours later with a bright smile on his face: Indeed, at the beginning of conditioning, dopamine neurons did respond to the juice reward. But with advancing learning trials, the reward response disappeared and reappeared at the stimulus that had come to predict the reward and elicit the movement (Schultz et al. 1993). The process resembled the response transfer we had seen with Ranulfo between the touch of a food morsel without a preceding stimulus and the stimulus of the food box opening (Romo & Schultz 1990). Now we had a response that engaged most dopamine neurons and transferred from reward to reward-prediction stimuli.

These experiments told us also something fundamental about reward processing. The response to the food morsel inside the food box (without any predictive door opening stimulus) occurred outside any learning episode. Maybe the reward response did not depend on an explicit learning process? Here, Jacques Mirenowicz came in who had joined us from Paris for a PhD. He was reading philosophy texts while supervising a monkey's task training (which is admittedly boring) and thus had a more reflective attitude. He figured that the simplest reward test would be a drop of juice reward delivered outside any task or learning episode, out of the blue. Sure enough, dopamine neurons responded strongly to such a "free reward" (Mirenowicz & Schultz 1994). The reward may well serve to condition all stimuli present at that moment, but the test was remote from any learning. It looked like the dopamine reward response reflected a more fundamental process still obscure to us.

The response transfer from reward to reward-predicting stimuli occurs not only in the described operant tasks but also during Pavlovian

conditioning (Waelti et al. 2001). Of course, reward predictors are Pavlovian conditioned, not operantly conditioned. Any stimulus in an operant procedure becomes inadvertently a Pavlovian conditioned reward predictor through its association with the reinforcer, as long as the animal is awake and attentive.

Together these results were funny in several respects. Dopamine responses do not categorically distinguish between liquid and food reward nor between primary and higher-order rewards. The response transfer occurs also between reward-predicting stimuli, and it is not restricted to learning. And the response occurs with artificial visual and auditory stimuli and with more natural food box opening when they predict reward. Thus, the dopamine response looks like a basic and fundamental reward signal. Our later economic approach would tell us that this is exactly what one would expect from a general reward value signal, which we then specified as coding formal economic utility (Stauffer et al. 2014).

Surprise 4: Reward Unpredictability

The response transfer was a funny thing, but we had no idea how to approach it. I remember a late morning in the summer of 1989 when Tomas Ljungberg and I were getting tired of searching for another dopamine neuron after several hours of recording and we started telling each other jokes. They were pretty good, but they were distracting to the point that we missed mistakes the monkey made. Of course, you don't get reward for making mistakes, and monkeys are no exception. When the predicted juice drop failed to occur because of a mistake, the dopamine neurons did not simply stop responding. Rather, they shut down their activity completely for a couple hundred milliseconds, time after time, exactly at the moment at which the reward would have occurred had the animal not made the mistake. We saw the dopamine depression clearly on the monitor of our computer, which was programmed to display, among other things, behavioral errors.

Well, the animal should not control the experiment, so we stopped chatting and occasionally withheld the juice deliberately even when the animal performed perfectly well. We saw the same brief dopamine depression when the reward would have occurred had we not interfered (Ljungberg et al. 1991). And there was no depression on those trials on which we did not omit the reward and gave it to the animal as predicted (which were most trials). Instead of going home for family lunch, we continued recording and found time and again the same depression with reward omission in all further dopamine neurons we recorded before stopping for the day. Thus, while we were joking, we had found that dopamine neurons were depressed when a predicted reward was omitted. The depression reminded us of an earlier result with Ranulfo in which dopamine neurons were activated when the animal touched food inside the food box into which it could not see; when the food box turned out to be empty, dopamine activity was depressed (Romo &

Schultz 1990). The experiment had been too uncontrolled to give the result much thought, but it had presaged the depression in the better controlled task and now made perfect sense. We later saw the same dopamine depression during well-controlled learning when the animal failed to get the reward because of making the wrong choice (Hollerman & Schultz 1998).

A dopamine depression with reward omission contrasts with intuition and neurophysiology: Neurons do not respond to absent events. Something else was going on, and its persistent and reproducible nature made us suspicious. A reward that is predicted by a stimulus and then does not occur is more than just an absent event. There was not nothing but something—that is, there was a stimulus, which had predicted a reward, and then the predicted reward failed to occur because of the animal's mistake or because a mean experimenter had blocked the reward tube. So, we were dealing with a violation of reward prediction, not simply reward absence.

A response that reflects violation of reward prediction indicates a relationship to unpredictability, which could explain all dopamine responses. Reward unpredictability was surprise 4, the biggest and most exotic of all surprises (Mirenowicz & Schultz 1994). And it explained all dopamine reward responses: the activation by unpredicted reward with learning, food touch and free delivery, the absence of response to predicted rewards, and the depression with omission of predicted reward. These results pretty much nailed unpredictability as a condition for the dopamine reward signal. The question was then, what function would a neuronal response have that reflects reward unpredictability? What behavioral or neuronal process might involve reward unpredictability?

From Reward Unpredictability to RPE

It was time to formalize these results to integrate them into existing theories of reward and decision-making. The solution came during my short sabbatical with Tony Dickinson in 1993 in Cambridge. We had sharp discussions about reward function every day in the coffee room of the Psychology Department, which were occasionally interrupted by a low-flying crow that contributed an occasional croak and more solid stuff. I asked Tony what reward function might involve reward unpredictability. Tony had just written his book on animal learning theory and was the right person at the right time (Dickinson 1980). He opened John Pearce's undergraduate textbook on animal learning (Pearce 1987), which stated on half a page the requirement of reinforcer unpredictability for associative conditioning as the basis of Rescorla and Wagner's reinforcement learning theory (1972). Amazing: This was exactly the way dopamine neurons code rewards! These gentlemen must have had a secret look at dopamine signals. What we labeled unpredictability is often referred to as prediction error and constitutes the basis of Rescorla and Wagner's associability term.

From my student days, I was familiar with control theory and the role of errors in correct performance. The only neuronal error signal known at the time was related to movements and involved the climbing fibers of the cerebellum (Gilbert & Thach 1977), as hypothesized by David Marr (1969). But error signals with reward? Why would you get a reward for having committed an error, as Roy Wise asked wisely? I was surely worried about generalizing the error concept to reward learning and its potential neuronal signals. As an experimentalist, I am closer to data than others and feel a responsibility to not mislead the field by engaging in overly enthusiastic interpretations. Others have called my attitude as struggling and not understanding my own data, but that is a small price to pay for being cautious; fancy interpretations can be falsified on the slightest occasion.

An error can be defined more broadly than just a behavioral mistake; an error is the discrepancy between what happens and what was predicted, like a positive or negative surprise, irrespective of requiring an action. This is like when your pub runs out of your favorite ale: You don't care about what exactly went wrong or who misjudged the barrel or forgot to refill it. You just sense the difference against the prediction and are disappointed. With such a definition, the use of the error concept by accepted reinforcement theories would allow us to understand the dopamine reward signal by using fundamental error concepts, and possibly contribute a neuronal foundation to animal learning theory.

While we had done well with intuition and ad hoc definitions, better scrutiny of a dopamine RPE signal would require more formal tests. Tony suggested the blocking paradigm of Kamin that tests the requirement of prediction errors in learning (Kamin 1969). We presented an initial stimulus that fully predicts a reward and added another stimulus without changing the reward; the reward did not change with the added stimulus and hence would fail to elicit a prediction error. Then we tested whether the added stimulus would be learned. Indeed, the monkey's licking demonstrated learning of the initial stimulus but not of the added stimulus. In parallel, dopamine neurons responded to the initial stimulus (that had been learned) but not to the added stimulus (that had not been learned). And the dopamine neurons were excited by a reward that was delivered unexpectedly after the added, nonpredictive stimulus (Waelti et al. 2001). Thus, dopamine responses followed the prediction error assumptions of the blocking paradigm: An RPE is necessary for learning, and the dopamine neurons signal that prediction error. Subsequently, Tricia Janak's group demonstrated behavioral learning when dopamine neurons were optogenetically stimulated at the moment at which a fully predicted reward elicited no response (Steinberg et al. 2013). Thus, the blocking paradigm nails the notion that dopamine responses code an RPE, and that their response drives learning.

While these dopamine responses conform with the Rescorla-Wagner learning model, our relentless playing around revealed another interesting

property that put the icing on the cake. The animals tested by Tomas Ljungberg had been trained in our spatial delayed response task, such that an initial stimulus instructs the animal to touch a left or right lever when a movement-triggering stimulus occurs and that movement-triggering stimulus then predicts the reward. In fact, both stimuli predict reward. Monkeys learn such a task backward: first the movement-triggering stimulus, and then the earlier instruction stimulus. With learning, the dopamine response occurs initially to the reward, then gets lost and transfers to the triggering stimulus, and then gets lost again and transfers to the instruction stimulus, which is now the earliest reward predictor (Schultz et al. 1993). Thus, dopamine responses transfer from reward to reward-predicting stimuli and then to the earliest reward predictor. All these responses depend on unpredictability, not only of the reward but also of the reward predictors. RPEs occur not only with rewards but also with reward predictors: Each stimulus elicits a prediction error relative to the prediction from the immediately preceding stimulus.

Reward-predicting stimuli elicit approach behavior and learning, like primary rewards, and thus constitute higher-order rewards. In fact, primary rewards are often difficult to identify and to distinguish from higher-order rewards. A juice drop at the mouth may be formally called a primary reward, but it predicts the liquid acting on sensory receptors in the mouth and reducing the salt concentration in the blood, both of which may be genuine, real "primary" rewards (of course, the juice taste on the tongue itself may also be a genuine reward). Thus, the liquid drop at the mouth may be primarily a reward predictor or higher-order reward, and the prediction error it elicits is a higher-order RPE. Hence, dopamine neurons detect prediction errors for both primary and higher-order rewards relative to the reward prediction at that moment.

Our finding about reward unpredictability seemed to relate well to RPE as the crucial variable of the Rescorla-Wagner learning model (Mirenowicz & Schultz 1994). But our additional observation that dopamine neurons signal both primary and higher-order RPEs matched a more advanced and efficient reinforcement model. In 1996, Read Montague, Peter Dayan and Terry Sejnowski reported the similarity between dopamine responses and temporal difference (TD) learning (Montague et al. 1996). TD learning was developed in 1981 by Rich Sutton and Andy Barto at Amherst. Their mentor Harry Klopf had been intrigued by the Bellman equation that describes how reward can be maximized by optimizing the pursuit of immediate reward together with all future predicted rewards. Learning through prediction errors at both primary and higher-order rewards can be more accurate as compared with considering only the final reward. Specifically, the single RPE of the Rescorla-Wagner model (reward minus prediction) is replaced in TD learning by (the current reward together with the temporally discounted sum of all future rewards) minus (the current prediction).

Thus, our dopamine responses in the delay task complied with the more efficient TD reinforcement learning formalism by coding RPEs with both primary and higher-order rewards.

Signals like the dopamine prediction error response serve as effective teaching signals; artificial models using dopamine-like TD signals, together with other machine learning algorithms, acquire performance of backgammon and Go games at world-champion level (Tesauro 1994; Silver et al. 2017). The fact that these implementations differ vastly from the brain's architecture emphasizes the importance of neuronal signals that may dominate over any particular architecture implementing the reinforcement algorithm (Purves 2019). The architecture in biological brains seems to represent an implementation that has emerged from evolution but that may not be crucial for reinforcement learning.

Simple Schemes That Convey Complicated Concepts

I was planning to attend the 1995 German Neurobiology Meeting in Göttingen where I would meet Rich Sutton. Rich had been the first graduate student of Andy Barto and developed TD learning together with him, and TD learning has since then become a mainstay in machine learning and artificial intelligence. I thought he might be interested to hear about a corresponding neuronal mechanism. I was frequently visiting my mother in Germany, which required long train rides from Switzerland with many hours to fill. I took essentials from our studies and played around with a graph until I got it the way I wanted. Rich loved it and told me that I was holding some genuine truth in my hands. I used the graph later for a review in the *Journal of Neurophysiology* to which Peter Strick had invited me (Schultz 1998). I also had the graph at hand when *Science* invited Read Montague, Peter Dayan, and me to write a review about the similarity between neuronal responses to RPEs, which had never been heard of in neuroscience, and artificial algorithms that were designed from scratch based mainly on insight and which were highly efficient for learning (Schultz et al. 1997). It goes like this:

1. An unpredicted reward elicits a dopamine activation, like a free reward, or the touch of a food morsel without a preceding stimulus, or a drop of liquid before and during conditioning.
2. Fully predicted rewards at exactly the predicted time elicit no dopamine response, as shown in established tasks or at the end of learning.
3. Omission of predicted reward elicits a dopamine depression, as shown with behavioral mistakes or the withholding of predicted reward.
4. Dopamine neurons respond also to unexpected reward-predicting stimuli (which are higher-order rewards). While a neuronal response to a reward-predicting stimulus may not in itself reflect a prediction error and might be compatible with Rescorla-Wagner learning, the disappearance

of that response when the stimulus is predicted by another, earlier stimulus indicates higher-order RPE coding compatible with the TD formalism.

Multiple Dopamine Response Components

The dopamine RPE response and its correspondence to sophisticated artificial reinforcement models differed from all traditional hypotheses of dopamine function. While the data were accepted by the community amazingly rapidly, dissenting voices were heard. There were indications of attention and punishment activating dopamine neurons that needed to be taken care of. True, we had ourselves seen dopamine activations to novel or known neutral control stimuli, and there were reports about responses to tail pinch, but we put them aside because we lacked solid interpretations. The solution came after a few extra experiments of our own (Mirenowicz & Schultz 1996; Kobayashi & Schultz 2014) and a couple of insightful studies by Masamichi Sakagami at Tamagawa University near Tokyo (Nomoto et al. 2010) and by Christopher Fiorillo at Stanford after he had left our lab (Fiorillo et al. 2013). Looking at these data together, we realized that dopamine neurons do what many neurons in other cortical and subcortical brain regions do: They show an unselective early activation to all kinds of stimuli before the animal and its neurons identify and evaluate the stimulus (Schultz 2016). The early response to unrewarded stimuli grows with the attention the stimulus generates, which would also explain the activation by aversive stimuli. The early response increases also when the stimulus resembles a rewarded stimulus or simply when rewards occur in the same environment. A few tens to a couple hundred milliseconds later, the true dopamine value signal emerges that is not explained by attention or punishment. Only recent experiments in rodents have revealed an anatomically separate subgroup of dopamine neurons that are activated by specific liquids that are also aversive, but without coding aversive prediction errors. These responses are in themselves interesting but leave the RPE response as the main ("canonical") phasic dopamine signal (Menegas et al. 2018). It might be that the early dopamine response component serves to alert the animal about a possible reward in the environment. The animal can prepare a rapid reaction but cancel the action if the object turns out a few hundred milliseconds later to be nothing of interest. The capacity to react quickly, before the second component identifies the object's nature, would provide an advantage over competitors in the long run of evolutionary selection.

From Stereotypy to Free Will

As soon as we saw the first strong phasic dopamine responses after having found the much smaller and rather boring movement-related changes,

we were stunned by the observation that these phasic responses occurred only with rewards and reward-predicting stimuli (when unpredicted) and with no other stimuli. Also, almost all responses were excitatory, whereas depressions of activity occurred almost only with reward omission. Thus, as reward-predicting stimuli are higher-order rewards and are treated similarly to primary rewards by TD models, the strong phasic dopamine responses (distinct from the smaller movement-related changes) were stereotyped for the kind of event that elicited them (reward) and for the kind of response (excitation with unpredicted reward, depression with omitted reward). This simplicity was a huge contrast to the vast response differences between neurons and task events in other brain structures that we had read about and that we saw in the neighbouring nondopamine neurons of substantia nigra pars reticulata while searching for dopamine neurons (Schultz 1986a). Thus, before going any further, we needed to make sure we were not missing anything. Thus, we recorded neurons in the striatum and SMA in the same monkeys and tasks and found time and again responses to all kinds of events and a mix of fast and slow excitations and inhibitions. The contrast to dopamine responses could hardly have been more impressive.

Our recording in striatum and SMA allowed us to catch up with the question of Free Will that remained from my time in Eccles' lab in Buffalo. Thus, we set out to test spontaneous, internally driven, "freely willed" movements. We recorded not only in the SMA but also in the striatum, which receives dopamine axons and is an important hub in cortico-basal ganglia loops with the SMA. At the time, SMA and striatum were both relatively unexplored, and any experiment would be exciting and result in novel data. The SMA receives strong transsynaptic inputs from the basal ganglia, and the striatum receives strong monosynaptic input from the frontal cortex, including SMA. Fuster (1973) and Kubota and Niki (1973) described prefrontal activity lasting several seconds before stimulus-driven movements. We wondered whether such activity might also exist when the animal itself decides to make a movement rather than in reaction to an imperative stimulus. Ranulfo trained monkeys patiently to self-initiate arm movements into our covered food box in a reproducible manner. We then recorded from about 20 arm, shoulder and back muscles to monitor, and control for, hidden muscle activity.

Our caution and self-scrutinizing paid off double. First, neurons in SMA and striatum showed large responses variations, both between neurons and between task events (Schultz et al. 1993); no comparison to the rather stereotyped dopamine responses. Second, subsets of SMA and striatal neurons increased their activity gradually for one to three seconds before self-initiated movements; the increased activity terminated abruptly as soon as the animal started the movement (Schultz & Romo 1992; Romo & Schultz 1992). Many of these neurons failed to show such sustained activity during the preparation of movements triggered by external stimuli and thus

indeed may have been involved in the internal generation of movement. I was very proud of this internally generated activity and wanted to follow it up later. We were touching the neuronal basis of intention-in-action (Searle 1983) on the road to Free Will. But the discovery of the dopamine RPE signal took all my attention, and we never went back to these ideas, which was quite a shame actually.

Given the huge diversity of cortical and striatal neurophysiology signals that others and we have seen, the question remains in what respect the much simpler and more stereotyped dopamine responses are homogeneous. People looking at the molecular, cellular, release and connection characteristics of dopamine neurons are emphasizing their heterogeneity, so the word homogeneous requires specification. Many dopamine responses are not homogeneous, including the slower movement relationships and the specific aversive non-RPE responses. But 'homogeneous' would apply to the events that trigger the phasic, two-component RPE response, which are both primary and higher-order rewards (Schultz 1998; Eshel et al. 2016). But while the RPE response is homogeneous in terms of triggering events, it varies quite a bit among individual neurons. For example, we found time and again that dopamine neurons in medial parts of the midbrain (ventral tegmental area, VTA) respond more frequently and strongly to RPEs than more lateral neurons (substantia nigra pars compacta). A recent and more quantitative study made this point clearer by describing optimistic dopamine neurons whose responses scale well with positive RPEs and hardly change with negative RPEs, whereas responses of 'pessimistic' dopamine neurons don't care much about positive RPEs but scale well with negative RPEs (Dabney et al. 2019; Lovet at al. 2020). Only the population activity of these neurons would code RPEs adequately.

Thus, we had seriously worried about missing something important when we had noticed the homogeneity of the reward triggering events for dopamine neurons. But addressing that worry resulted in our foray into the Free Will problem that together with consciousness still constitutes a main challenge for neuroscience.

What Does Dopamine Do?

I often hear this question when walking the corridors of conferences. Many years ago, a prominent Parkinson researcher asked me, rather incredulously, how a dopamine reward signal could account for Parkinsonian movement deficits. He obviously assumed that a dopamine signal should correspond to Parkinsonian deficits, and the RPE signal hardly did so.

There is a wide spectrum of motor and cognitive impairments in diseases like Parkinsonism and schizophrenia and in experimental pharmacological and lesion tests on humans and animals. Although interference with dopamine neurotransmission induces reward learning deficits in humans and

animals (Zweifel et al. 2009), many deficits cannot be easily explained by phasic RPE signaling, including the Parkinsonian movement deficits. This is the issue: Although it is reasonable to infer a natural function from the deficits that occur when the system is compromised, the RPE signal simply won't explain the many motor and cognitive deficits. I have only one general, and not necessarily satisfying, answer for these deficits that comes back to the notion of "rain on the brain": Dopamine is a chemical that is released in a slowly changing or even tonic manner and serves as a fuel that enables other neurons to do their job; a true neuromodulator without a sophisticated function on its own. Of course, this tonic function is only one of the many dopamine functions (Schultz 2007).

But the confusion about dopamine function goes deeper and persists even when only considering phasic dopamine changes. I have been talking for more than 30 years about dopamine RPE responses and have sidelined the far less interesting slower and lower dopamine changes with behavioral activation, but others are now rediscovering these earlier reported changes. As electrophysiologists, we study subsecond neuronal signals, and an initial sharp and short reward response is not the same thing as a slower and weaker change during behavioral activation. But the two dopamine changes might have a common function, like motivation, that combines reward detection with the behavioral activation that is necessary to obtain the reward. The two dopamine signals may even have a common RPE function (Kim et al. 2020), which might suit the lower time resolution of complex rodent behavior and some new optical recording methods.

Thus, to answer the question: dopamine does have multiple behavioral relationships at different time scales (Schultz 2007): (1) the fast, two-component signal coding RPEs, the phasic changes with behavioral activation and movements (that might overlap with the RPE function); and (2) the slow or even tonic enabling function inferred from compromised dopamine neurotransmission like Parkinson's. In any event, dopamine neurons contradict the widespread assumption that one brain system has one function.

Leaving Fribourg

I will always be grateful for two assets this Swiss city provided me with: the possibility to discover the dopamine RPE signal in a well-organized environment with good grant support, and the safety, structuredness, and coziness of a pretty city between mountains and lakes that allowed us to bring up a family with three children in a joyful manner, including our daughter Carolina who was born in 1980. But I had become infected with, or spoiled by, a certain academic atmosphere as a postdoc in Creutzfeldt's, Eccles', and Ungerstedt's laboratories. I encountered such an atmosphere in Cambridge during the 1993 stay with Tony Dickinson and Nick Mackintosh, and I was longing to have it again. I expressed my interest to Tony and some of the

people I had met in Cambridge, including Trevor Robbins and Barry Everitt. At one point, Martin Johnson as head of the Anatomy Department heard about my interest and suggested that I look into a research professorship from the Wellcome Trust that would bring my laboratories to Cambridge. I applied to the Wellcome Trust and wrote a Home Office animal license application. When both were granted, my move to Cambridge was sealed for a start in October 2001.

I admit that a number of people were stunned by my easy movement from one country to another. Some people even thought I wanted to avoid Germany, but I had not the slightest such intention. It was a job and academic offer I could not refuse, nor get anywhere else at the time. I don't take national allegiance too narrowly. If anything, I am European and would be comfortable to live in any western European country (as well as in the United States). Hence, moving to England for me was more like moving from one U.S. state to another, which Americans are accustomed to doing. The disrespect for local boundaries was reinforced by the excellent travel connections within Europe by Eurostar and the budget airlines that brought travel costs down and made visits with our adult children in Germany, Switzerland, and Britain frequent and routine events. One of my sons living in Geneva even took a job at King's College London and travels there every week for term lectures and meetings, sometimes on the same day going back and forth. It is so nice and easy to benefit from the sophisticated culture of Europe!

Cambridge Work

Neuroeconomics: A New Research Direction

At one of Terry Sejnowski's summer conferences in the 1990s in Woods Hole, I was not only charmed by this wonderful Cape Cod harbor village, but I was also impressed by a talk from Paul Glimcher who suggested investigating reward processing with formalisms of economic theory. He emphasized particularly the subjective nature of reward value, which in economics is formalized in various versions of utility theory. I had encountered utility in small lectures on game theory in Heidelberg, with many informal discussions with the lecturer. So, Paul's presentation rang a bell to an initiated but otherwise naïve mind.

Applying economic theory to neuronal reward studies is not trivial: the amazing similarity between the dopamine RPE signal and the basic TD algorithms provides a biological foundation for efficient learning machines. As the algorithms are superb and should be further developed and exploited, why bother with biological details such as reward subjectivity, as some of my machine learning colleagues argue? Sure, the Bellman equation as a basis of TD learning already takes subjective reward value implicitly into account

when describing reward maximization, so everything should be settled. But biological organisms are not silicon machines, and we need to know whether reward neurons indeed code subjective value as assumed by the Bellman equation. Thus, before going any further in the parallelism between dopamine responses and TD learning, we needed to know whether the dopamine signal represents reward value in a subjective but formal metric. Furthermore, utility coding might also explain how dopamine neurons deal with risk, which is crucial as rewards are, by default, uncertain. And how would the dopamine signal be involved in choice? This is where economic theory comes in. So, I started building my next experimental designs on the basic concepts of economic decision theory. The move to Cambridge offered the possibility to incorporate this new direction into the next 20 years of work even before neuroeconomics became a field.

My entry into neuroeconomics was aided by almost yearly visits to Caltech, which began with initial seminar invitations from John Allman around the year 2004 and were hosted by Colin Camerer and Ralph Adolphs. So, during most of my Cambridge time, Gerda and I spent about two months per year in Pasadena. This was not only sunny California but also provided easy personal access to world leaders in neuroscience and economics. I gave lectures and seminars, participated in a large NIH center grant, visited also with Richard Andersen and John O'Doherty, and learned economics from Colin Camerer, Charlie Plott, David Grether, Antonio Rangel, and Federico Echenique, which amounted almost to a private economics PhD. I was also impressed by Caltech's attitude toward monkey neurophysiology. As a result, Richard Andersen's monkey neurophysiology flourished, and its extension to human brain-machine interface (Aflalo et al. 2015) attracted a large donation from Tianqiao and Crissy Chen for a huge new neuroscience institute (with a Zen Garden in the basement).

The Dopamine Utility Prediction Error Signal

The economic value of a reward depends on its amount and the probability with which it occurs. A bigger reward has more value than a smaller one and a more frequent reward has more value than a less frequent one, provided satiety is ruled out. To understand whether dopamine neurons might code reward value thus defined, Christopher Fiorillo and Philippe Tobler tested rewards of different amounts occurring with various probabilities. Sure enough, most dopamine neurons increased their activity monotonically with both parameters (Fiorillo et al. 2003; Tobler et al. 2005). We then used the reinforcement learning framework to follow the intuitive, frequentist approach in probability theory that derives probability from the frequency of past events. Indeed, the frequency of past rewards translates into a neuronal code for reward probability (Lak et al. 2016). Thus, dopamine neurons code reward amount and probability as the two most basic

economic value parameters. Furthermore, dopamine neurons are sensitive to both liquid and food rewards (and to higher-order visual, somatosensory, and auditory rewards) and code their value in a common currency (Lak et al. 2014), rather than being interested only in particular rewards. These value properties made dopamine reward signals amenable to a neuroeconomics approach.

To address subjective dopamine reward value coding in a simple manner, we used a traditional test called temporal discounting: reward delays decrease the subjective value of rewards without changing their physical properties. When monkeys choose the same reward at different delays, they prefer the early reward; their choices indicate an exponential value drop (Kobayashi & Schultz 2008), exactly as theory would have it. Sure enough, the dopamine responses to delay-predicting stimuli decrease in a similar manner, indicating sensitivity to subjective reward value. The evidence of subjectivity was restricted to temporal discounting, however, and a more general assumption of subjective value coding would require at least one other situation. Risk presented such a possibility.

Risk concerns both loss and uncertainty and comes in at least two forms. When trusting a surgeon with your life, you probably prefer one with a lower failure rate to one with a higher failure rate, everything else being equal. Here risk emphasizes loss while also dealing with uncertainty; it is identical to negative (probabilistic) value, and thus basically the opposite to (probabilistic) reward. By contrast, economic risk emphasizes uncertainty while also dealing with loss. When shopping for a car, you probably prefer one that works more or less well, and thus is only little risky, to one that works sometimes really well and on other days really poorly and thus is quite risky. You never quite know whether the riskier car works, and you incur a big loss when it doesn't. Thus, economic risk varies with the difference between outcomes, with larger difference being riskier. It is captured by the higher statistical "moments" of probability distributions, primarily variance, skewness, and kurtosis. We implemented the simplest and unequivocally defined economic risk, the so-called mean-preserving spread in which either one or the other of two equally probable rewards occurs; both gambles have the same mean reward amount but different variance (Rothschild & Stiglitz 1970). Such abstract academic definitions may have real-world bearing: Stiglitz later became the chief economist of the World Bank and won the Nobel Prize. Monkeys often prefer the riskier to the less risky gamble despite their identical mean as long as the reward distribution is not skewed. Dopamine responses to cues predicting gambles are stronger with the preferred riskier gamble and reflect its higher value (Stauffer et al. 2014). Thus, as with temporal discounting, dopamine neurons code subjective reward value with risk, rather than with physical value.

The probability and subjective value coding of dopamine neurons provided enough arguments for trying utility functions that represent

subjective value in a mathematically coherent way. Bernoulli claimed 280 years ago (1738) that agents value rewards nonlinearly, based on the observation that extra rewards add gradually less subjective value (diminishing marginal utility, represented by concave utility). Such functions have the property of all mathematical functions and can predict events that were not used for their construction, like choices. Von Neumann and Morgenstern's (VNM) utility axioms (1944) conceptualize choice under risk and allow to estimate utility functions with cardinal numeric properties (immune to change in mean and variance), which is mathematically appropriate for neuronal response functions that are intrinsically cardinal.

We estimated utility functions iteratively from psychophysically estimated certainty equivalents in choices between a variable safe reward and a set gamble with two equiprobable rewards. I designed the test during one boring flight to the United States and inadvertently reinvented the fractile, chaining procedure described by Caraco (1980); so we were in business. We fitted splines to the certainty equivalents and obtained an S-shaped utility function that was convex with small rewards (0.1–0.5 milliliters, indicating risk-seeking), linear with larger rewards (risk neutrality), and slightly concave with larger rewards (1.0–1.3 milliliters, indicating risk avoidance). We also estimated risky and riskless utility functions from various direct parametric fits. The utility functions were similar between the different models but varied between individual monkeys. Although VNM utility is a first step toward a fundamental, well-defined decision variable, there are limits to its general applicability, which are addressed in Kahneman and Tversky's prospect theory (1979) by adding probability weighting, reference dependency, and gain-loss slope differences. Now we had an experimentally viable method for assessing subjective value in a general mathematically defined way, which could be used for testing utility coding in dopamine neurons.

We did three tests: First, increasing amounts of unpredicted reward (eliciting a positive RPE) induced a nonlinear dopamine response increase that resembled nicely the curved utility function. That was encouraging but not a very controlled result. Second, in gambles with two outcomes, the larger reward elicited a positive RPE, and the smaller reward elicited a negative RPE. The RPEs were identical in gambles with identical spread, even if the gambles had different means. When the gambles were positioned at different parts of the nonlinear utility function, the identical RPEs became varying utility prediction errors (UPE): The same RPE became a small UPE where the utility slope was flat, and a big UPE where the slope was steep. Indeed, the dopamine response to the large gamble reward was small when the gamble was placed on the flat part of the utility function and large when the gamble was placed on the steep part (the typical low background activity allows only a small difficult-to-quantify depression induced by the small gamble reward). That result basically nailed utility coding by

dopamine neurons and was confirmed by the third out-of-sample test. Being a mathematical function, utility predicts choice: A convex utility function predicts preference for a riskier compared to a less risky gamble, even with the same physical mean amount of the two gambles. The monkeys know the theory and follow that prediction, and their dopamine response as well: the neuronal response is stronger to the stimulus predicting the preferred riskier gamble compared to another stimulus predicting the less risky gamble. These three results suggest that dopamine neurons code economic utility and thus follow the necessary assumptions of a biologically relevant reward signal. Such dopamine signals could indeed be useful for updating decision variable coding in postsynaptic neurons. Now, Bernoulli's suggestion from 280 years ago had a firm biological basis: Utility exists in the hardware of the brain and can be measured, which is much better than inferring it only from observable choice.

Then the real question arises: Would such a dopamine signal have behavioral consequences? The answer is yes. Artificial electric and optogenetic dopamine excitation and inhibition direct animals toward reward and away from nonrewards, respectively (Olds & Milner 1954; Corbett & Wise 1980; Tsai et al. 2009). Bill Stauffer in our laboratory confirmed and dissected the effects in monkeys. Optogenetic dopamine stimulation enhances dopamine responses to reward-predicting stimuli compatible with reinforcement learning and directs behavioral choice toward the stimulus associated with the stimulation (Stauffer et al. 2016). Thus, dopamine excitation follows the definition of reward function, generating learning and approach behavior. Dopamine inhibition has the opposite effect, generating unlearning and avoidance. As there are no real liquid or food rewards in these laboratory stimulation experiments, what the animals seek is a dopamine excitation, rather than a better real liquid or food reward that does not exist. Likewise, animals avoid dopamine inhibition, rather than a worse real reward that does not exist in the lab. Thus, when animals visibly seek and avoid real rewards outside the lab, the rewards are only the means to obtain the desired dopamine excitation and to avoid the hated dopamine inhibition. Or, when I encounter a better-than-predicted reward, it generates a positive RPE and dopamine excitation; when I go after that reward, I do it because I want to experience that dopamine excitation. Likewise for worse-than-predicted reward: I hate dopamine inhibitions and therefore will avoid the worse reward. Thus, when we go to the pub, we only seem to want drinks and food and see friends, but what we really seek is the dopamine excitation elicited by the prediction and experience of drinks, food, and friends.

And there is one more point (Schultz 2021): When we succeed in getting a dopamine excitation, the excitation increases the reward prediction in the next trial through reinforcement learning. But the increased prediction requires better reward for eliciting the same dopamine prediction error excitation the next time; to obtain that same dopamine excitation, we now

need to seek an even better reward. When we obtain that better reward and the desired dopamine excitation it elicits, it increases the prediction again, and the iteration continues and drives us toward better and better rewards. Such a reward maximization process is beneficial for the individual decision maker, but it also makes us restless as we become dissatisfied with what we have; and it drives us continuously to ever better rewards. Little dopamine devils in the brain!

Reward Neurons Galore

The Striatum

Our recordings in nondopamine brain structures that served as controls for the rather homogeneous dopamine reward signal brought us to the striatum. We quickly realized a wider engagement beyond its traditional function in movement. Using a go/no-go task in which a particular stimulus required the animal to withhold a movement, we saw activity in many striatal neurons during the no-go period before the reward. It looked like we had found a movement inhibition signal, which had never been seen before. But then we used control go trials in which the reward was delayed by a similar duration between the movement and the reward and found a similar activity as in no-go trials. So, no movement inhibition activity, but maybe reward expectation? We played with the duration of the waiting period before the reward, and the striatal neurons continued to show the persistent activity before the reward, even with an 80-second delay, way beyond the sampling duration of our computer data collection program. The activity only went down when the animal lost patience and moved his arm a bit and then resumed for the rest of the 80 seconds when he settled down again. This reward expectation activity was just one of many forms of reward, stimulus, and movement processing at various task events in the striatum (Apicella et al. 1992; Hollerman et al. 1998). By contrast, neurons in the ventral striatum, also called nucleus accumbens, showed less diversity and more relationship to reward (Apicella et al. 1991). Thus, striatum neurons were at the crossroads between movement and reward; this was maybe the region in which the reward signals of dopamine neurons met the local reward and movement signals required for economic choice.

The recordings in the striatum allowed for a true extension of our reward work, beyond comparison with dopamine neurons. We were particularly interested in social comparisons of reward. When Raymundo Báez-Mendoza joined the lab from Mexico thanks to a Max-Planck master program, he set out to test cooperative choices of two monkeys in the popular prisoner's dilemma. To provide well-controlled test conditions, we set up cameras by which each monkey watched the well-separated partner monkey. With all that technology, we slowly realized that this artificial lab situation was

nowhere near the natural world of monkeys and that we might not trust the data we were to collect. We also contacted Chris Harris from our economics department, who suggested we use a dictator game as a scaled-down social test with less complexity. One monkey would simply decide how much of available reward he would get and how much the other monkey would get. A good social neuronal reward signal should make a distinction between which of the two animals gets the reward. We confirmed that our monkeys in the lab distinguished between their own and the conspecific's reward and knew who acted. We found that most striatal neurons coded the monkey's own reward rather than the reward for the other monkey. Importantly, the responses depended on which monkey performed the action that led to the own reward; different striatal neurons coded the reward for the recorded monkey when his own action or the other monkey's action led to that reward, thus distinguishing the actor whose movement led to that reward (Báez-Mendoza et al. 2013). The actor-specific responses disappeared in some neurons when the other monkey was replaced by a computer, demonstrating specificity for biological agents. Thus, in one of the first social neurophysiology reward studies on monkeys, we had found a reward signal in striatal neurons that identified the social agent responsible for the monkey's own reward.

Upstairs, the Orbitofrontal Cortex Serves Economic Choice

The frontal cortex is another major brain structure whose activity we wanted to compare with the dopamine reward signal, and the orbitofrontal cortex (OFC) was at that time the neurophysiologically least known frontal area. Edmund Rolls in Oxford had done pioneering reward studies (Thorpe et al. 1983), but there was more to be done when Leon Tremblay joined the lab from Canada via Paris. Starting with the task that had demonstrated sustained reward expectation signals in the striatum, we found similar reward expectation activity in OFC neurons, as well as responses to reward-predicting stimuli and rewards, but the signals were less frequent and less strong than in the striatum (Tremblay & Schultz 2000). With Leon's usual imagination, he tried out something that originated more from a gut feeling than a theoretical concept. He wondered whether adaptation might occur in OFC neurons. He was guided by ideas from Horace Barlow's and Simon Laughlin's earlier work on retina adaptation that makes processing within the neuron's limited coding range more efficient. He first established the monkey's choice preference between two rewards. Then he asked whether the same OFC neuron might show different responses to the same reward when he changed only the alternative reward. Indeed, the neuronal response followed the animal's preference: When the animal preferred a piece of apple to a piece of cereal, the neuron responded more to apple, but when the animal preferred a raisin to the same piece of apple, the neuron

lost most of its response to the apple and instead responded well to the raisin (Tremblay & Schultz 1999). Although this neuronal change seemed compatible with adaptation to the current reward distribution (which changed between the two experimental steps), the responses also followed the definition of preference as the probability of choosing a reward from a given set of rewards. Formal economic preference coding would bring OFC signals into the domain of economic choice variables, which became an active research field after Camillo Padoa-Schioppa and John Assad's ingenious experiments on common currency reward value coding in OFC neurons (Padoa-Schioppa & Assad 2006). But the adaptation issue stuck: When Shunsuke Kobayashi joined us from Tokyo, he thought that neuronal adaptation should not only occur relative to the mean value but also relative to the width (statistical variance) of the reward distribution. Indeed, some OFC neurons show steeper reward-response slopes with narrower distributions, and thus better reward discrimination; the response slopes were flatter with wider distributions, indicating better use of the neuron's limited coding range (Kobayashi et al. 2010). Thus, besides the issue of reward preferences, OFC neurons adapt to the first two statistical moments of reward probability distributions as indication of efficient processing.

As the adapting OFC responses demonstrated sensitivity to reward statistics, we remembered variance as a definition of economic risk. We had seen a slow signal in dopamine neurons that increased with variance, but because of its slow time course, we thought it might serve the dopamine teaching function rather than economic decisions (Fiorillo et al. 2003). Now we wondered whether OFC neurons might carry an explicit risk signal whose latency and duration would be short enough to serve economic decisions and change monotonically with variance. We used Rothschild and Stiglitz's mean-preserving spread and asked whether some OFC neurons show stronger responses to stimuli indicating higher risk, as defined by a larger spread between the top and bottom rewards of binary gambles. Indeed, a population of OFC neurons showed stronger responses with larger risk, but without coding reward value itself (O'Neill & Schultz 2010). Apparently, distinct OFC neurons coded reward value and risk, both of which are important for choice under risk.

The work on nondopamine reward structures encouraged us to investigate reward functions beyond reinforcement learning. There is the old question of how rewards that typically have several dimensions, and thus are vectors, could be coded as single-dimensional (scalar) neuronal signals. An apple has taste and amount, but how could the lower or higher activity of a reward neuron represent the combined value from the two components? The issue might be approached by looking at revealed preference theory, which describes how an integrated subjective value results from different reward components. For best experimental control, we would test choices between two "bundles" that each have the same two distinct rewards, like

apple juice and grape juice, which we could set experimentally to specific amounts. When options with different compositions are chosen equally frequently, they are assumed to have equal subjective value. To participate in the choice between such options, individual neurons should show similar responses to these options despite their different composition, and more preferred bundles should elicit stronger OFC responses even if one bundle component was smaller than in a nonpreferred option. This question stimulated Alex Pastor-Bernier, who had joined us from Montreal and set out with great enthusiasm to test bundles with two juice components on monkey OFC neurons. To get familiar with the concepts and develop a feasible behavioral task, Alex first did an extensive behavioral study in which Charlie Plott at Caltech provided encouragement and the necessary economic background (Pastor-Bernier et al. 2017). I had met Charlie many years ago as a founder of experimental economics (now more fashionably called behavioral economics) during one of my early visits to Caltech and had benefited from his vast practical knowledge and lucid explanations. Based on the well-tested behavior, Alex went on to record from OFC neurons and identified neuronal signals that integrated the values from all bundle components combined (Pastor-Bernier et al. 2019). The neurons showed similar responses to bundles that were similarly preferred by the monkey (choice indifference), while showing larger responses to more preferred bundles (confirming the neurons' sensitivity). The neuronal responses remained strongest with the preferred bundle even when a third bundle was added to the option set, thus complying with Arrow's weak axiom of revealed preference and the independence of irrelevant alternatives (IIA) as a requirement for utility maximizing choice. Interestingly, humans often violate the IIA. This contrast does not necessarily indicate that monkeys are more rational than humans but may reflect the fact that monkeys have experienced tens of thousands of such tests and thus don't get fooled, whereas humans are tested only a few times and get it wrong more easily. As a solid confirmation, a support vector machine classifier correctly predicted behavioral choices from neuronal responses. With this result, the single-dimensional OFC reward signal seems to be involved in choices of typically multidimensional rewards.

The bundle design provided us also with a chance for testing selective reward satiety. Earlier, nonbundle tests had failed because satiety for one reward destroys the monkey's appetite for all rewards. By contrast, the finely controlled reward adjustments afforded by the bundle design preserves the animal's motivation. Indeed, the choice indifference curves show graphically the value reduction of the sated bundle reward in a beautiful, intuitively understandable pattern, and the OFC responses follow the differential change (Pastor-Bernier et al. 2021). Thus, the concepts of economic theory had allowed us for the first time in neurophysiology to identify reward-specific satiety in a controlled manner.

An Open Mine: The Amygdala

While the reward signals in striatum neurons provided a stark contrast to the dopamine reward responses, our experience with OFC neurons went well beyond comparison with dopamine responses and showed us signals for an economic decision machine. But there was one more brain structure in which we suspected major reward signals. As judged by current reviews, the amygdala is primarily involved in aversive processing, because fear is lost after lesions and neurons respond well to fear-inducing stimuli. But we were not convinced. When Istvan Hernadi joined the lab in 2002, he irreverentially designed, together with Ken Tsutsui, a sophisticated reward and economic choice schedule in which the monkey chose between saving a reward with interest for future consumption and immediately consuming the already saved reward. Sure enough, amygdala neurons responded well to reward and reward-predicting stimuli, and their activity transitioned from reward value to coding the actual choice. Amygdala activity could even predict whether the monkey would save more or cash in the saved reward. In the meantime, Dan Salzman convincingly demonstrated amygdala reward responses in a straightforward Pavlovian task (Paton et al. 2006), which discouraged our immediate publication. Then Fabian Grabenhorst joined the lab and wanted to apply his data analysis skills from his Oxford PhD with Edmund Rolls. His analysis demonstrated an amygdala involvement in multiple aspects of sophisticated reward processing and economic choice, thus advancing Dan Salzman's amygdala reward signals toward economic choice (Grabenhorst et al. 2012; Hernadi et al. 2015; Grabenhorst et al. 2016).

After Istvan had left for his native Hungary, Maria Bermudez came from Spain and tested more basic theoretical concepts of reward processing, such as reward contingency. A stimulus is only learned as a reward predictor when the reward occurs more frequently in the presence of the stimulus compared with its absence. Simple stimulus-reward pairing does not explain learning (although it is often claimed so); even a stimulus that is paired with a reward does not become conditioned as a reward predictor when the reward occurs as frequently without the stimulus as with the stimulus, because such a stimulus does not contain distinct reward information. The crucial function of reinforcer contingency was unknown to Pavlov and was first demonstrated with aversive stimuli by Robert Rescorla, but it had never been shown before with neurons. Indeed, Maria's amygdala responses to reward-predicting stimuli disappeared when the reward occurred as frequently without the stimulus as with the stimulus, thus demonstrating a neuronal correlate for crucial contingency (Bermudez & Schultz 2010). Further experiments by Maria demonstrated adaptive coding of reward amounts and sensitivity to reward timing as typical features of reward neurons, which encouraged further consideration of the amygdala as a key reward structure.

But our curiosity did not stop here. Lesion studies of the amygdala have indicated a social function. By the time Fabian was done publishing Istvan's data, he had accumulated a full data set from his new experiments. We had designed a task in which the monkey learns the value of probabilistic rewards from repeated choice and observes a partner monkey sitting opposite (Grabenhorst et al. 2019). Based on his own experience, the monkey learns the reward-predicting stimuli of the other monkey and might understand and simulate his choices. Besides coding their own reward, amygdala neurons learned the value of the reward for the other monkey. A support vector machine classifier developed by Fabian and Arek Stasiak showed that the monkey's own amygdala responses predicted the partner monkey's choices. Thus, while the sophisticated reward signals in the amygdala extended well beyond its traditional role in fear, amygdala neurons are even engaged in social reward processing.

A Look Back

Would I become a scientist again? Definitely, if one could make it a paying profession. It is never guaranteed that the stuff one imagines and designs works. It often looks more like wonder. Yes, there is a gut feeling of what may work and what may not work, but gut feelings can be wrong. Distractions are there, and they can compromise scientific productivity. I believe in the notion that 90 percent of our activities are governed by unconscious processing. You are focusing on a text, and when you wake up in the morning you suddenly have the exact formulation on which you were laboring. Your brain did that for you. But if you flood your brain with tons of unnecessary things, the brain will faithfully advance on these unnecessary things, and the interesting stuff gets left behind. So, your first thoughts in the morning tell you what your mind works on. Even better, you can preload your brain with your problem before sleep, and the brain works it out and tells you the solution in the morning. Thus, your brain's night job does what you are being paid to do during the day and which brings in the big money that your dopamine neurons seek.

Would I become a neuroscientist again? That is more difficult to answer: what would I want to investigate? Many basic processes are known by now, published in hundreds of thousands of papers. There is of course one big and most important question that has stubbornly escaped understanding and might have the importance of Watson and Crick's double helix: consciousness. It would be fantastic to contribute, but one needs to generate ingenious experimental plans and lots of concepts. The other direction is human neuroscience. The new methods for investigating the human brain in action are fantastic and have the added advantage of addressing clinical questions, which is very moral and justified when it comes to the responsible use of research money, often provided by the taxpayer. And

of course, there will always be a new method, like optogenetics currently, which is very exciting. But then, what do you want to discover with the new method that is not already at least partly known? Would this mean filling important gaps, which can be satisfying and attract big grants, or can one, maybe inadvertently, make a surprising discovery? These are the reasons why so many people are interested in new methods that open new research avenues.

Why did my work work? Well, factors were surely luck, persistence, sweat, diligence, stubbornness, and irreverence toward established career goals. I started with a temporary and inferior faculty position in Fribourg, but it became quickly permanent and allowed me to apply for rather easy, relatively small grants with good advice (notably from Dr. Winkler at the Swiss National Science Foundation). This unambitious position was no way to make a career, but I could advance as slowly as necessary to understand what I was doing and work gradually toward a good grasp of the data rather than following publication pressure.

What could have gone wrong? Everything. There is no way to predict a surprising discovery. But my head of department was impressed when a departmental colleague working on blood vessels told him that by experimenting with dopamine neurons, I might be doing something of potential importance. But that was not the point: what mattered was that I could do what I felt was interesting. I did not focus on immediately obvious research questions, and I even had difficulties understanding them; I rather felt that we scientists should generate new problems, in the sense of identifying a new paradigm worth the effort, and then go for it when it seemed to work.

Did people understand what I was doing? It is difficult to say, but initially not many, and I felt sometimes outside of mainstream discussions at conferences. While some people did not sign on, others appreciated the cautious and tedious progress and encouraged me. Discovery at any level does not come for free and is never certain. There is always frustration when trying to understand what one is doing and seeing. The struggle is not always apparent when looking at the result, in particular, when it finally, after lots of effort, becomes simple and straightforward: everything is clear, so why did it take so long to figure it out? It took us 20 years to unravel a particularly telling reward signal, and another 20 years to characterize its biological usefulness, but at the beginning even the term of reward needed a better definition, which is not the case when working on the visual system, for example. People gave us credit for such risky undertaking. In addition, there are the difficulties of experiments with capricious monkeys, supervised by capricious humans, when the animal won't perform the task, and I still don't know whether this is the right test. These experiences teach me something, and I feel the progress—tedious, slow, and never sure, but ultimately a lot of fun and very rewarding.

Selected Bibliography

Aflalo T, Kellis S, Klaes C, Lee B, Shi Y, Pejsa K, Shanfield K, Hayes-Jackson S, Aisen M, Heck C, Liu C, Andersen RA. Decoding motor imagery from the posterior parietal cortex of a tetraplegic human. *Science* 348: 906–909, 2015.

Apicella P, Ljungberg T, Scarnati E, Schultz W. Responses to reward in monkey dorsal and ventral striatum. *Exp Brain Res* 85: 491–500, 1991.

Apicella P, Scarnati E, Ljungberg T, Schultz W. Neuronal activity in monkey striatum related to the expectation of predictable environmental events. *J Neurophysiol* 68: 945–960, 1992.

Báez-Mendoza R, Harris C, Schultz W. Activity of striatal neurons reflects social action and own reward. *Proc Natl Acad Sci USA* 110: 16634–16639, 2013.

Berke JD. What does dopamine mean? *Nat Neurosci* 21: 787–793, 2018.

Bermudez MA, Schultz W. Responses of amygdala neurons to positive reward predicting stimuli depend on background reward (contingency) rather than stimulus-reward pairing (contiguity). *J Neurophysiol* 103: 1158–1170, 2010.

Bernoulli D. Specimen theoriae novae de mensura sortis. *Comentarii Academiae Scientiarum Imperialis Petropolitanae (Papers Imp Acad Sci St. Petersburg)* 5: 175–192, 1738. (Translated as: Exposition of a new theory on the measurement of risk.) *Econometrica* 22: 23–36, 1954.

Bindra D. Neuropsychological interpretation of the effects of drive and incentive-motivation on general activity and instrumental behavior. *Psychol Rev* 75: 1–22, 1968.

Caraco T, Martindale S, Whitham TS. An empirical demonstration of risk-sensitive foraging preferences. *Anim Behav* 28: 820–830, 1980.

Corbett D, Wise RA. Intracranial self-stimulation in relation to the ascending dopaminergic systems of the midbrain: A moveable microelectrode study. *Brain Res* 185: 1–15, 1980.

Dabney W, Kurth-Nelson Z, Uchida N, Starkweather CK, Hassabis D, Munos R, Botvinick M. A distributional code for value in dopamine-based reinforcement learning. *Nature* 577: 671–675, 2019.

DeLong MR, Crutcher MD, Georgopoulos AP. Relations between movement and single cell discharge in the substantia nigra of the behaving monkey. *J Neurosci* 3: 1599–1606, 1983.

Dickinson A. *Contemporary Animal Learning Theory*. Cambridge, UK: Cambridge University Press, 1980.

Engelhard B, Finkelstein J, Cox J, Fleming W, Jang HJ, Ornelas S, Koay SA, Thiberge SY, Daw ND, Tank DW, Witten IB. Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. *Nature* 570: 509–513, 2019.

Eshel N, Tian J, Bukwich M, Naoshige Uchida N. Dopamine neurons share common response function for reward prediction error. *Nat Neurosci* 19: 479–486, 2016.

Fiorillo CD, Song MR, Yun SR. Multiphasic temporal dynamics in responses of midbrain dopamine neurons to appetitive and aversive stimuli. *J Neurosci* 33: 4710–4725, 2013.

Fiorillo CD, Tobler PN, Schultz W. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299: 1898–1902, 2003.

Fuster JM. Unit activity of prefrontal cortex during delayed-response performance: neuronal correlates of transient memory. *J Neurophysiol* 36: 61–78, 1973.

Gilbert PFC, Thach WT. Purkinje cell activity during motor learning. *Brain Res* 128: 309–328, 1977.

Grabenhorst F, Báez-Mendoza R, Genest W, Deco G, Schultz W. Primate amygdala neurons simulate decision processes of social partners. *Cell* 177: 986–998, 2019.

Grabenhorst F, Hernadi I, Schultz W. Prediction of economic choice by primate amygdala neurons. *Proc Natl Acad Sci USA* 109: 18950–18955, 2012.

Grabenhorst, F., Hernadi, I., Schultz, W. Primate amygdala neurons evaluate the progress of self-defined economic choice sequences. *eLife* 5: e18731, 2016.

Griffiths LA. Detection and identification of the polyphenoloxidase substrate of the banana. *Nature* 184: 58–59, 1959.

Guyenet PG, Aghajanian GK. Antidromic identification of dopaminergic and other output neurons of the rat substantia nigra. *Brain Res* 150: 69–84, 1978.

Hernadi I, Grabenhorst F, Schultz W. Planning activity for internally generated reward goals in monkey amygdala neurons. *Nat Neurosci* 18: 461–469, 2015.

Hollerman JR, Schultz W. Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat Neurosci* 1: 304–309, 1998.

Hollerman JR, Tremblay L, Schultz W. Influence of reward expectation on behavior-related neuronal activity in primate striatum. *J Neurophysiol* 80: 947–963, 1998.

Kamin LJ. Selective association and conditioning. In: *Fundamental Issues in Instrumental Learning* (eds. Mackintosh NJ, Honig WK). Nova Scotia, Canada: Dalhousie University Press, pp. 42–64, 1969.

Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. *Econometrica* 47: 263–291, 1979.

Kim HR, Malik AN, Mikhael JG, Bech P, Tsutsui-Kimura I, Sun F, Zhang Y, Li Y, Watabe-Uchida M, Gershman SJ, Uchida N. A unified framework for dopamine signals across timescales. *Cell* 183: 1600–1616, 2020.

Kobayashi S, Pinto de Carvalho O, Schultz W. Adaptation of reward sensitivity in orbitofrontal neurons. *J Neurosci* 30: 534–544, 2010.

Kobayashi S, Schultz W. Influence of reward delays on responses of dopamine neurons. *J Neurosci* 28: 7837–7846: 2008.

Kobayashi S, Schultz W. Reward contexts extend dopamine signals to unrewarded stimuli. *Curr Biol* 24: 56–62, 2014.

Kubota K, Niki H. Prefrontal cortical unit activity and delayed alternation performance in monkeys. *J Neurophysiol* 34: 337–347, 1971.

Lak A, Stauffer WR, Schultz W. Dopamine prediction error responses integrate subjective value from different reward dimensions. *Proc Natl Acad Sci USA* 111: 2343–2348, 2014.

Lak A, Stauffer WR, Schultz W. Dopamine neurons learn relative chosen value from probabilistic rewards. *eLife* 5: e18044, 2016.

Libet B, Gleason CA, Wright EW, Pearl DK. Time of conscious intention to act in relation to onset of cerebral activities (readiness-potential): The unconscious initiation of a freely voluntary act. *Brain* 106: 623–642, 1983.

Ljungberg T, Apicella P, Schultz W. Responses of monkey midbrain dopamine neurons during delayed alternation performance. *Brain Res* 586: 337–341, 1991

Ljungberg T, Apicella P, Schultz W. Responses of monkey dopamine neurons during learning of behavioral reactions. *J Neurophysiol* 67: 145–163, 1992.

Lowet AS, Zheng Q, Matias S, Drugowitsch J, Uchida N. Distributional reinforcement learning in the brain. *TINS* 43: 980–997, 2020.

Marr D. A theory of cerebellar cortex. *J Physiol* (London) 202: 437–470, 1969.

Menegas W, Akiti K, Amo R, Uchida N, Watabe-Uchida N. Dopamine neurons projecting to the posterior striatum reinforce avoidance of threatening stimuli. *Nat Neurosci* 21: 1421–1430, 2018.

Mirenowicz J, Schultz W. Importance of unpredictability for reward responses in primate dopamine neurons. *J Neurophysiol* 72: 1024–1027, 1994.

Mirenowicz J, Schultz W. Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature* 379: 449–451, 1996.

Montague PR, Dayan P, Sejnowski TJ. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci* 16: 1936–1947, 1996.

Musall S, Kaufman MT, Juavinett AL, Gluf S, Churchland AK. Single-trial neural dynamics are dominated by richly varied movements. *Nat Neurosci* 22: 1677–1686, 2019.

Nomoto K, Schultz W, Watanabe T, Sakagami M. Temporally extended dopamine responses to perceptually demanding reward-predictive stimuli. *J Neurosci* 30:10692–10702, 2010.

Olds J, Milner P. Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *J Comp Physiol Psychol* 47: 419–427, 1954.

O'Neill M, Schultz W. Coding of reward risk distinct from reward value by orbitofrontal neurons. *Neuron* 68: 789–800, 2010.

Padoa-Schioppa C, Assad JA. Neurons in the orbitofrontal cortex encode economic value. *Nature* 441: 223–226, 2006.

Pastor-Bernier A, Plott CR, Schultz W. Monkeys choose as if maximizing utility compatible with basic principles of revealed preference theory. *Proc Natl Acad Sci USA* 114: E1766–E1775, 2017.

Pastor-Bernier A, Stasiak A, Schultz W. Orbitofrontal signals for two-component choice options comply with indifference curves of revealed preference theory. *Nat Comm* 10: 4885, 2019.

Pastor-Bernier A, Stasiak A, Schultz W. Reward-specific satiety affects subjective value signals in orbitofrontal cortex during multi-component economic choice. *Proc Natl Acad Sci USA* 118: e2022650118, 2021.

Paton JJ, Belova MA, Morrison SE, Salzman CD. The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* 439: 865–870, 2006.

Pearce JM. An *Introduction to Animal Cognition*. Lawrence Erlbaum, Hove and London, 1987.

Purves D. What does AI's success playing complex board games tell brain scientists? *Proc Natl Acad Sci USA* 116: 14785–14787, 2019.

Rescorla RA, Wagner AR. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: *Classical Conditioning II: Current Research and Theory* (edited by Black AH, Prokasy WF). New York: Appleton Century Crofts, pp. 64–99, 1972.

Romo R, Schultz W. Dopamine neurons of the monkey midbrain: contingencies of responses to active touch during self-initiated arm movements. *J Neurophysiol* 63: 592–606, 1990.

Romo R, Schultz W. Role of primate basal ganglia and frontal cortex in the internal generation of movements: III. Neuronal activity in the supplementary motor area. *Exp Brain Res* 91: 396–407, 1992.

Rothschild M, Stiglitz JE. Increasing risk: I. A definition. *J Econ Theory* 2: 225–243, 1970.

Schultz W. Activity of pars reticulata neurons of monkey substantia nigra in relation to motor, sensory and complex events. *J Neurophysiol* 55: 660–677, 1986a.

Schultz W. Responses of midbrain dopamine neurons to behavioral trigger stimuli in the monkey. *J Neurophysiol* 56: 1439–1462, 1986b.

Schultz W. Predictive reward signal of dopamine neurons. *J Neurophysiol* 80: 1–27, 1998.

Schultz W. Multiple dopamine functions at different time courses. *Ann Rev Neurosci* 30: 259–288, 2007.

Schultz W. Dopamine reward prediction error signalling: a two-component response. *Nat Rev Neurosci* 17: 183–195, 2016.

Schultz W. A dopamine mechanism for utility maximization. *SSRN* 3927732, 2021.

Schultz W, Ruffieux A, Aebischer P. The activity of pars compacta neurons of the monkey substantia nigra in relation to motor activation. *Exp Brain Res* 51: 377–387, 1983.

Schultz W, Apicella P, Ljungberg T. Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J Neurosci* 13: 900–913, 1993.

Schultz W, Dayan P, Montague RR. A neural substrate of prediction and reward. *Science* 275: 1593–1599, 1997.

Schultz W, Romo R. Dopamine neurons of the monkey midbrain: contingencies of responses to stimuli eliciting immediate behavioral reactions. *J Neurophysiol* 63: 607–624, 1990.

Schultz W, Romo R. Role of primate basal ganglia and frontal cortex in the internal generation of movements: I. Preparatory activity in the anterior striatum. *Exp Brain Res* 91: 363–384, 1992.

Searle JR. Intentionality. Cambridge, UK: Cambridge University Press 1983.

Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, Guez A, Hubert T, Baker L, Lai M, Bolton A, Chen Y, Lillicrap T, Hui F, Sifre L, van den Driessche G, Graepel T, Hassabis D. Mastering the game of Go without human knowledge. *Nature* 550: 354–359, 2017.

Stauffer WR, Lak A, Schultz W. Dopamine reward prediction error responses reflect marginal utility. *Curr Biol* 24: 2491–2500, 2014.

Stauffer WR, Lak A, Yang A, Borel M, Paulsen O, Boyden E, Schultz W. Dopamine neuron-specific optogenetic stimulation in Rhesus macaques. *Cell* 166: 1564–1571, 2016.

Steinberg EE, Keiflin R, Boivin JR, Witten IB, Deisseroth K, Janak PH. A causal link between prediction errors, dopamine neurons and learning. *Nat Neurosci* 16: 966–973, 2013.

Stuber GD, Wightman, RM, Carelli RM. Extinction of cocaine self-administration reveals functionally and temporally distinct dopaminergic signals in the nucleus accumbens. *Neuron* 46: 661–669, 2005.

Sutton RS, Barto AG. *Reinforcement Learning*. Cambridge, MA: MIT Press, 1998.

Tesauro G. TD-Gammon, a self-teaching backgammon program, achieves master-level play. *Neural Comp* 6: 215–219, 1994.

Thorpe SJ, Rolls ET, Maddison S. The orbitofrontal cortex: neuronal activity in the behaving monkey. *Exp Brain Res* 49: 93–115, 1983.

Tobler PN, Fiorillo CD, Schultz W. Adaptive coding of reward value by dopamine neurons. *Science* 307: 1642–1645, 2005.

Tremblay L, Schultz, W. Relative reward preference in primate orbitofrontal cortex. *Nature* 398: 704–708, 1999.

Tremblay L, Schultz W. Reward-related neuronal activity during go-nogo task performance in primate orbitofrontal cortex. *J Neurophysiol* 83: 1864–1876, 2000.

Tsai H-C, Zhang F, Adamantidis A, Stuber GD, Bonci A, de Lecea L, Deisseroth K. Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science* 324: 1080–1084, 2009.

von Neumann J, Morgenstern O. *The Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press, 1944.

Waelti P, Dickinson A, Schultz W. Dopamine responses comply with basic assumptions of formal learning theory. *Nature* 412: 43–48, 2001.

Zweifel LS, Parker JG, Lobb CJ, Rainwater A, Wall VZ, Fadok JP, Darvas M, Kim MJ, Mizumori SJ, Paladini CA, Philipps PEM, Palmiter R. Disruption of NMDAR-dependent burst firing by dopamine neurons provides selective assessment of phasic dopamine-dependent behavior. *Proc Natl Acad Sci USA* 106: 7281–7288, 2009.