Nineteen Ways of Looking at Consciousness

Patrick House

Begin Reading

Table of Contents

About the Author

Copyright Page

To Claire Roscow (1988–2010),
Dr. Jonathan Leong (1983–2019),
and Cam Christie (1982–2020);
all, whose neurons turned in early

The movements of animals may be compared with those of automatic puppets, which are set going on the occasion of a tiny movement; the levers are released and strike the twisted strings against one another.

—Aristotle, "On the Motion of Animals"

Introduction

If I were asked to create, from scratch and under duress, a universal mechanism for passing consciousness from parent to child, I would probably come up with something a bit like grafting a plant. Each parent would donate a small piece of their brain and place it on some sort of growth medium, maybe some agar, or some flour with sugar and yeast, and the child would sort of just expand, like those water-absorbent foam dinosaur toys, into its final shape around the pieces of parental brain until it, too, was conscious. How else could it possibly work?

Instead, something much more remarkable happens in nature. An entirely new creature can grow into a fully conscious version of itself, and the entire process occurs, as if by fiat, anytime a certain kind of single cell with the right mix of nucleotide sugars is kept alive for long enough. Which means that consciousness is not something passed on or recycled—like single molecules of water, which are retained as they move about the earth as ice, water, or dew—from one living creature to the next.

Instead, consciousness can be grown from "scratch" with only a few well-timed molecular parts and plans laid out. It is not drawn from a recycled tap of special kinds of cells or dredged from a vein of free will. No, the darn thing just grows. From its own rules. All by itself. And we have no idea how or why.

Despite this, most of the time, except to the philosophers and neuroscientists, there is no "problem" of consciousness. It seems to work just fine, almost effortlessly, which is a large part of its rarely questioned charm. One can have a full and vivid life without giving a second thought to the makings of the first. And yet, even though everybody knows what consciousness feels like, it remains central to the greatest unsolved scientific and existential mysteries we will probably ever face: What is it? And how did it get here?

This book is a collection of possible mechanisms, histories, observations, data, and theories of consciousness told nineteen different ways, as translations of a few moments described in a one-page scientific paper in *Nature,* published in 1998, titled "Electric Current Stimulates Laughter." The idea is an homage to a short book of poetry and criticism, *Nineteen Ways of Looking at Wang Wei,* which takes the poem "Deer Park" by Wang Wei and analyzes nineteen different translations of it in the centuries since it was originally written.

Generally, though not in every case, each of the nineteen "ways" in this book promotes the idea that the brain evolved for, is itself, and will always be dedicated to movement. The best reason to start there is the simple, physical, and verifiable fact that a human brain's only outputs are the muscles it connects to, whether the small ones that dart the eyes back and forth or the lumbering thigh muscles that kick-start a walk. No matter what goes into a brain, only movement ever comes out.

And though I believe that this is the correct starting point for understanding consciousness—that it can be fully understood from the brain's parts and their goals—this starting point is by no

means universally agreed upon. There are as many possible lenses to translate the details through as there are ways to hold a lens. Here, every chapter points toward an explanation for a few moments during the neurosurgery described in the *Nature* article, when a still-awake patient had her brain carefully prodded by an electrode capable of producing blasts of electricity and altering both her behavior and, arguably, her consciousness. Throughout this book, I will call the anonymous patient Anna. (Only her initials, A.K., are given in the original article.)

As a whole, the arguments and ideas in this book represent some of the most popular or plausible theories of how consciousness arises, works, feels, or degrades with use or error. Each chapter does not align neatly with any one theory or thinker, and some theories find their way into multiple chapters. In part, this is necessary because the many theories of consciousness today are as varied as the scientific fields they draw from, a funny quirk of human reason. For those studying the origins of life on the seafloor, for example, consciousness may be the natural interactions of pH, proton pumps, and evolved metabolic efficiency over time. But another theory, from someone who happens to know psychology and decision-making well, may be explained by lots of org-chart-style boxes and arrows. To those scientists who study attention, perhaps consciousness is the mathematical equivalent of a spotlight; to those who study eye movements, it might be the grand effort to cancel out self-generated movement; to those who study flies, it could be as probabilistic as a population of midges' decision to go left or right or loop back on itself; to those who study songbirds, it might be the cognitive consequence of the rewiring of language-learning loops; and so on. There are, in other words, almost as many theories as there are thinkers out there.

I can be clear about my biases, too. I spent the better part of a decade doing laboratory research on a tiny parasite that infects mouse brains and maybe makes the infected host mouse prefer the smell of cat urine just a little bit more than they had before. Thus, I am a mind-control parasite guy. I think in host-parasite terms and relations. I think of free will as something like a river's current with a natural direction and kind of physics to it, which also can, under the right conditions, be redirected. However, as with Wang Wei's original poem, which is lost to history, I believe there is still a way to inch closer to an unknowable truth, guided by the sense that one is on the right track.

For this reason, some chapters are explanations of consciousness as a single allegory, comparing the evolution of consciousness with the evolution of the game of pinball (chapter 2) or even to a small town (chapter 10). Other chapters base the explanation on the latest scientific research into how a brain controls a body's movement (chapters 3 and 17) or on how consciousness arises from waves of electrical activity (chapter 4), language learning (chapter 12), causal power (chapter 14), or efficient compression of information (chapter 11).

Some chapters seek explanations in the physics of temperature and prediction (chapter 8) or in the quantum realm (chapter 15). Others propose that the best clues are found in the minds of learning robots (chapter 18) or in the mind of a cat with thousands of toes, should she lose some of them (chapter 16).

Others consider why removing parts of the brain changes some aspects of self but not all of them (chapter 13) or how and why a conscious brain is often telling itself fictions (chapter 1). Some consider whether the brain is more like a simulation (chapter 6) or a radio broadcast (chapter 9) with networks in the brain the receiving antennae. Others take issue with the very idea of theorizing about consciousness and defining its terms (chapters 7 and 19) or wonder what answer humanity would give if forced (chapter 5).

Each of the nineteen ways described in this book is a patchwork or tapestry of scientific findings and arguments and a collection of compelling or interesting conclusions organized

around a few themes or arguments. For the sake of clarity, and readability, each translation is self-sufficient, and requires no background knowledge of the philosophical or scientific work on which the chapter is based. Names, credit, and caveats have been excluded from the main text and included in footnotes, which give context, references, and sources as appropriate.

The twentieth and final chapter ties all of them together.

★ ★ ★

The hardest part about studying consciousness is that it hides by making the outside world into a version of its wishes, trapping its host in a kind of virtual, fragile snow globe we call subjective experience. We are stuck studying it from the inside, and any attempt at explanation is always going to be an attempt at translating kinds of experience.

In *Nineteen Ways of Looking at Wang Wei,* the book's editor directly confronts the problems inherent to translation. I have a similar hope for this book, where contradictions, ambiguities, and debates around relative priority of theories or thinkers are encouraged, rather than dismissed, as they often are in academic discussion around consciousness. Who, for example, has authority to speak about consciousness? Do scientists poking around mouse brains really know more about human brains than therapists or manicurists, who have spent tens of thousands of hours probing why someone had a bad day? Does an electrode really tell you more than a question? (If so, by what measure?) Is there one answer for how all consciousness works across all of life, or one answer per consciousness, per moment? When different explanations are possible, which aspects of which theories get collapsed or lost and why?

In answering these and related questions, this book offers complex arguments and thought experiments that highlight the differences and similarities among many of the major modern theories of consciousness. The hope is that, through repetition and variation from different angles and points of view, a clearer sense of each and, ultimately, of the original—that is, the subjective feeling of what it is like to be conscious—can emerge.

Why another book on consciousness? Why now? When I was training for my Ph.D., at Stanford, I was once asked by a local vacuum repairman if I could help explain to his brother, a drug addict, what addiction truly was down there in the weeds of his brain. When I responded that I could not, that I am just a laboratory researcher and that, besides, we do not know what addiction is because we do not know what brains are and that addiction is social and complicated and context-dependent and its mechanisms mostly unknown, he responded ruefully, but with a kind of haughty resignation, that he works with vacuums because he can fix them.

That same year, a brilliant friend of mine, whom I had met at a neuroscience retreat, died by suicide, at age twenty-two. I was told by another of her friends, before I had even asked, that she had always been "very sad." Neuroscience is a frustrating field to be in. We have mostly failed, as there is little point to celebrate an understanding of small pieces of something if you cannot prevent the failures of its whole. We know very little about the things we wish to know much more about, and the strain on mental, subjective well-being today is, from the global numbers, only getting worse.

Sometimes, I imagine that modern neuroscientists share both the wonders and frustrations of an astronomer in, say, the Babylonian times; the simultaneous wonders and frustrations of looking up at the sky every night in awe at what was and was not known about the motions of the points of light up there so impossibly high in the sky. At the time, Babylonian astronomers knew where stars would be next but not why; they knew that some stars, which were actually planets, would

appear to move backward in the sky but, again, not why. Their guesses, and they had a lot of guesses, were all wrong.

Similarly, today, across the many scientific disciplines that study the brain and wish to know its secrets, we know where activity occurs in the brain and can maybe even predict where it might occur if you show someone, say, a face—in the "face area" of the inferior temporal cortex—but we do not know why. I have heard it said that the difference between physics and biology is the difference between Galileo dropping both a bowling ball and a pigeon off the top of the Tower of Pisa. But really the difference is that physics can land an autonomous robot on Mars with pinpoint accuracy, study the origins of the Big Bang from under a mountain in Italy, and split a literal atom in half to unleash the cosmic hell of a thousand suns, while we over in neuroscience cannot even tell you what being "very sad" is. That gap in knowledge frustrates the search for consciousness. And the longer we go without understanding it, the more people we will lose for inexplicable reasons.

My preferred translation of Wang Wei's poem "Deer Park" is the following literal, character-by-character attempt at translating the twenty ideograms directly from classical Chinese.

Empty	mountain(s) hill(s)	(negative)	to see	person people
But	to hear	person people	words conversation	sound to echo
To return	bright(ness) shadow(s)	to enter	deep	forest
To return Again	to shine to reflect	green blue black	moss lichen	above on (top of)

What I appreciate most about this translation is that many of the individual ideograms have multiple or ambiguous meanings that the translation leaves unresolved. The middle character of the last line can be translated as "green," "blue," or "black"; the fourth character of the last line as either "moss" or "lichen"; the second character on the first line as "mountain" or "mountains" or "hills"; the second character on the third line as "bright" or "brightness" or their opposite, "shadow."

The character-by-character translation does not settle on which version of a concept or word is "correct." In my years of studying the brain, I have found no more compelling analogy to the difficulty of understanding the human mind than this single page of poetry. It has been said that the uniquenesses of humans' consciousness is our ability to hold more than one contradictory idea in our head at the same time, a feat that this translation embraces.

What does it mean to translate consciousness when some people see images in their mind's eye but others, called aphantasics, "see" nothing? What does it mean that there are twenty amino acids that make up all of life on Earth and that, depending on the context (like each ideogram), can take on separate roles? What does it mean that some people have more kinds of light-responsive cells in their retina, which may allow them to discriminate more colors than others, or that some of those same people see dresses as white-and-gold striped and others see the same dress as blue-and-black striped? Or that some people remember things and can replay them just as they lived

them while others see memories replay as if they were broadcast on a television one hundred feet away?

What does it mean for the ability to talk about the human mind, as if it is only one thing, that one person (James Joyce) described the inside of his mind as a "grocer's assistant" while another (Albert Einstein) claimed to be able to visually imagine the speed of light as a teenager? Or that when the first description of the "stream of consciousness" was used, to describe a work by the novelist Dorothy Richardson, she strongly objected to the phrase, saying that consciousness to her, instead, "sits stiller than a tree"? Which of all these versions of consciousness is "correct"?

Nobody is wrong, but in a way everyone is, because their version is theirs alone. So, too, must we start at the simplest, most literal translation of consciousness to have any hope of grasping it. The character-by-character translation. The one and only thing we know for certain for every one of us. That there is something that it is like to be us. That's it. Everything else is unknowable excess. One of the beauties, and frustrations, of talking about consciousness is that everyone can know only their slice of the world. We have many tools, like language, gesture, and theory of mind, to try to jump into others' heads, but at the end of the day we can only ever scratch the surface of what really goes on inside.

★ ★ ★

Some practical notes on reading this book. The original text of "Electrical Current Stimulates Laughter" can be read in the appendix (here). It is mostly readable to nonscientists. However, this book is meant to be read with no knowledge of the study in question, and it is explained in pieces, sometimes just as a hint, in every chapter. The basic story is that a neurosurgeon, using small, carefully placed blasts of electricity to the brain, was able to cause the patient, Anna, to laugh. Alone, this is not surprising. We have long known that electricity powers our muscles to act, and laughter is just a series of rapid, coordinated muscle movements. What was so surprising was that Anna said afterward that she also felt the subjective sensations of joy and mirth alongside the laughter and that she, when asked why she laughed, gave different and implausible answers each time.

This book is based on the idea that, like an unsplit atom, Anna's story contains within it hidden multitudes and that any attempt to explain consciousness should be able to fully explain the events of that day, both inside and outside Anna's head. Thus, each chapter title can be preceded by the words "Consciousness is..." These titles do not always represent my opinion, nor is each statement inarguable fact, and all are best read with the presumption that the point of view is being argued by someone who believes strongly in that chapter's claims or conclusions.

A clear example of this approach is chapter 7, "The Median Price of a Thrift-Store Bin of Evolutionary Hacks Russian-Dolled into a Watery, Salty Piñata We Call a Head." The chapter starts with a quote from Sydney Brenner, Nobel laureate in genetics and close friend to Francis Crick, that "consciousness" is not the singular problem we think it is and will be remembered in time as a confusing misdirection, like the search for "luminiferous aether" in the late nineteenth century. This was something he told me, in personal communication. I spent days interviewing Brenner, who has since passed away, about his ideas on the brain, in his hotel suite at the Shangri-La, in Singapore. I listened; we argued; I came back the next day.

Brenner, hooked up to oxygen and bound to a wheelchair, had spent each morning combing through genomic databases for the web of life, trawling it for similarity to human genes in hopes of finding one last secret; I, alas, had mostly spent my mornings wondering if it was ironic or unironic that the cuttlefish we ate for dinner the prior night had tasted like nothing. And so, that chapter,

its title, and conclusions are my best efforts to summarize his idea that brains are an evolutionary happenstance full of errors and that the lessons of history apply to genomics and neuroscience all the same. It is his translation, in a sense, but in my words. I'm just the ghostwriter, as everyone translating "Deer Park" was for Wang Wei, for millennia hence.

Similarly, other chapters do the same for a variety of thinkers and theories, some of whom I have met or interviewed for this book and some of whom I have followed and read from afar. (There are many theories and thinkers that this book owes a great debt to—sources, notes, references, and further reading can be found in the Notes section at the end of the book.)

Any repetition or small variation in repetition is on purpose. Of the fifteen translations of "Deer Park" that are in prose, the description of "moss" or "lichen" is translated thirteen times as "green," once as "blue," and twice as no color at all. What a reader is left with if they developed all the translations, as a kind of photograph, would be a single image, multiply exposed, with a combination of mostly green, a little bit of blue, and a little bit of nothing. I believe that every book on consciousness is confronting a similar translation effort, even if naively.

Ideally, any explanatory gaps along the way ("Wait, how did we get from microtubules to the subjective feeling of joy, exactly?") also exist in nature. We are pretelescope. Pre–Isaac Newton. Some of the biggest and best brain-reading devices in existence have not been the clear keyhole into the hallowed rooms of the mind that we had hoped for decades ago. Instead, today's tools are more like the lens grinders that laid the foundations for Galileo to, one day, make glass smooth enough to curve and capture just enough light for him to turn a telescope upward and outward.

★ ★ ★

One last note. In any act of selection, people can get sensitive about rankings or inclusion. If you know of or adhere to a particular theory of consciousness that does not get mention in this book, and take umbrage, I defer to a line from the editor of *Nineteen Ways of Looking at Wang Wei,* which both addresses and immediately closes the door on a similar problem, but for poems: "I have presented only those definitions that are possible for this text. There are others."

Whatever Wang Wei or evolution intended for their respective creations no longer matters. All we have left are versions. I have presented only those theories that I believe give the best clues to explaining the mysteries, horrors, and awe-inspiring details of consciousness and its multiple explanations. (There are others.) At the very least, the various and sometimes contradictory theories are essential background to the gathering of stable and accurate observations about brains and minds, which will persist as data for future scientists or philosophers to figure out what it all means. There is a precedent. For many centuries, observations about magnetism and electricity were tallied and noted, but it was not until a theory came along relating the two that the cataloged unexplainable phenomena were suddenly understood as the symptoms of fully explainable interactions.

Perhaps the same will happen, one day, with consciousness.

1

Relative to the Observer Who Is Also a Liar

In the mid-1990s, Anna, age sixteen, had brain surgery to alleviate suffering from epilepsy. Because the brain does not have pain receptors,[1] she was allowed to remain awake the whole time, with only a tiny amount of localized anesthetic to numb her scalp. During the procedure, the doctors and attendants asked Anna a series of questions meant to keep her talking and, as they probed her brain with electricity and micron-thin blades, hoped that she would not stop.

Though the language parts of her brain were roughly in the same places as they are in all other human brains, the brain moves with each pulse of blood, and every brain, like every coastline, has its own slightly different contours. If Anna had stopped speaking at a certain spot of stimulation—because the electrical current could activate cells relevant to thought and speech—the surgery team knew they were in an area important to language and thus one to be avoided by the surgeon's scalpel.

Oddly, they did not ask Anna while her brain was being probed with electricity to write poetry and stop when the poetry became bad. They did not ask her to intuit a response to a fictional domestic dispute and stop when her response became improbable or immoral. They did not ask her how far in front of her eyes her visual imagination extended and stop when the distance became too uncomfortably far or the description too Daliesque or she suddenly lost perspective. Instead, they asked her to do a variety of seemingly mundane tasks, including name objects, read, count, and flex her hands and toes.

Many imagine the human brain as a series of lit-up wires connected together like telephone poles, strung inside a snow globe made of bone, each neuron brightening like a candle when it has something big to say or wants its owner to notice something. It is not. There is no light, for one. The brain is messy and venous and dense and soaking wet, all the time, and is about as heavy as a hardback copy of *Infinite Jest*.[2] It is not designed, perfected, or neat. It is a thrift-store bin of evolutionary hacks Russian-dolled into a watery, salty piñata we call a head.

If the surgeons had poked Anna's brain with the tips of their fingers, which they would never do, but surely has been done, her brain would have had the give of a very soft Brie cheese. The surgeons could thread a small, loose wire straight through it, but never would because along those tracks could be memory, identity, and bits and pieces of the girl's sense of her teenage self, which is to say the accumulation of her preferences. Hanging like a furled sail off Theseus's ship,[3] a surprisingly tensile outer shell called the dura mater (Latin, and Freudian, for "tough mother") would be visible near the girl's head, as seen by those in the room, though not by Anna, who just wants her seizures to stop.

The surgeons eventually found a spot on her brain while operating that, when stimulated, caused her to laugh, a discomfiting sound in any operating theater. More technically, the surgeons used their fancy electric wand to produce an electrical current that, because the brain also uses electricity to communicate many of its own messages, caused certain neurons in her cortex to send a signal mostly indistinguishable from the natural one to parts of her muscles to coordinate action of these muscles, and it was these contractions that bounced air between them to produce a sound perceived by those others in the room as "laughter."

Strangely, when asked the source of her laughter, Anna gave a different answer each time. The answer was dependent on her immediate surroundings, and often involved an aspect of a picture she was viewing or a person near her ("the horse is funny," "you guys are just so funny...standing around"), even though the correct answer, involving the surgeon's electrode, eluded her. Instead, she confabulated the reasons behind the laughter and mirth because the brain abhors a story vacuum and because the mammalian brain is a pattern-recognizing monster, a briny sac full of trillions of coincidence detectors that are only useful if there are connections between things. Even a wrong pattern, a guess, is at least a pattern to learn against.

Though she did not receive general anesthesia, those before and after her in the same operating room did, and a full description of her awake brain must also explain one of the most remarkable things about consciousness. It can be silenced by anesthesia, in part or in whole, only to recover fully again in a few seconds, minutes, hours, or days. That there is no one kind of anesthesia to turn the consciousness dial to zero means that, although all conscious brains may be alike, the conscious brain has many, and sensitive, failure modes. Consciousness, like Tolstoy's unhappy family, has only one way of adding up to a whole, but many ways of falling apart: xenon gas, propofol, isoflurane, cocaine, nitrous oxide, barbiturates, benzodiazepines, and ketamine, each with different chemical profiles and causes, can all silence consciousness.

Some anesthetics, like propofol, which is sometimes called the "milk of anesthesia" because it is white and oily and repels water, can cause bizarre effects. People who were crying before anesthesia came out of it, hours later, again (still?) crying. These effects are less like a pause button for consciousness and more like a needle lifted off a spinning record.[4] When waking up from anesthesia, or coma, a state sometimes called post-trauma amnesia, people will often have strange, lewd, or primal behaviors, speech, or urges. Legally and socially, people are not often held responsible for what they do or say during this time, which makes one wonder why we ever are. Consciousness is all and every one of these states equally. Any good theory must fully account for each of them.

That consciousness disappears nightly is another of its quirks.[5] Though nobody quite understands what sleep is, we know what it looks like and that anesthesia does not induce it. Some animals can sleep with only half their brains at a time, allowing basic functions of consciousness to persist so that they don't fall out of the sky or get eaten. Those in the water who don't fear being eaten, like humpback whales, often sleep vertically, often in groups, like the large towers of an aquatic city, for less than ten percent of their day. Sleep concerns are highly specific: birds dream of bird problems, whales of whale problems, dogs of dog problems.

Those that can lucid dream, which is a kind of awakeness within dreaming, or an awareness that one is dreaming, can be trained to move their eyes while in the lucid state, under rapid eye movement, and these movements under the eyelids can be detected by an infrared camera.[6] Interestingly, this means one can make a code, like those in a video game or medieval monastery, that lets one break the subjective fourth wall and communicate with the great sleep researchers in the sky. For example, a person can learn to, if lucid, move their eyes in a certain pattern and then count to ten, after which they move their eyes in that same pattern again, to mark the end of

their test. Remarkably, some people take around ten "objective" seconds to do so, which implies that their subjective, incepted time—the waking dream within the dream—not only has a time keeping device but that it may be the same one we always use.

Where did this conscious, clock-making observer, lucid or otherwise, who can wake up and take in their immediate surroundings—who, in physical terms, the speed of light sticks to, who only in observation can collapse a wave of light into a position and determine whether Schrödinger's cat is dead or alive—come from?

From the oceans, of course. The most useful thing that land offers that the oceans do not is the large distance at which things can be detected. Visually, swimming in water is like driving in a milky fog, which reduces the range of even the best mammalian eyes. A small bacteria can sense, crudely, in a small shape around itself—we can call this "sight" if it can respond to a light source, or "smell" if it can detect an unwanted chemical toxin nearby—and the total of its sensory range, the full addition of all its input, stretching through and combined across all its sense, is called its sensorium. The experience of any underwater creature paying attention is an experience of underwater objects or other creatures popping into frame at such high speed that an underwater sensorium needs timing closer to reflexes than contemplation. Even in the clearest water, light scatters and degrades over only a few meters, which means there isn't much need for a brain to come up with long-term planning, because what would be the use?

Thus there isn't much need for clock making beyond intervals of a few seconds, which means that there is no need for the brain to whir up an emergency motor-response plan for the shark cresting over the horizon of the Adriatic shelf, because there is no way to sense the shark cresting over the horizon of the Adriatic shelf. This usefully constrains the metabolism needed to keep track of the far outer radii of the outside environment and means any need to plan movements is limited to the timing of events with a small, near-reflexive range.

Hiding in the center of these plans is the observer, the conscious creature, who is just an accumulation of movement preferences and plans trapped inside a sensorium, keeping track of what it thinks the objects around it are and what it might otherwise do with itself.

The great move of life onto land from the milky oceans changed the range of timing that the newly landed needed to care about. The expanded range of being able to see farther through the crisp air, and with new eyes to boot, meant that prey vis-à-vis their predators had to plan to move themselves across alternately sparse and cluttered landscapes in order to find and not be food. To plan, one needed a sense of time, in order for there to be something unto which the plans unfurled; for there to be a sense of time, there had to be a timekeeper. Thus the timing that mattered most for sensing, predicting, and planning depended on a creature's sensorium but the exigencies of land, like gravity and tripping, made it suddenly necessary to plan seconds and minutes ahead.

In a kind of spherical symmetry, minutes of forward planning (imagination) required minutes of backward recall (memory) and, like an inflating balloon expanding evenly and temporally on all sides, landed creatures needed to be able to pay attention to the future and in so doing pay attention to the past by the same amount. To know what a lion cresting the African horizon will do, one must be able to keep track of what similar-looking creatures once did after cresting similar-looking hills.

Mammals like dolphins or whales, which crawled back into water after a brief stint as hippos, saddened by how murky it was, used all the vocal tricks learned on land to re-create through echolocation and sonar, as much as possible, the range of visual distance the eyes granted.[7] Because some sound waves travel through parts of the ocean almost as far as light waves travel through air, and because a brain can be thought of as a tool, like any other, to make sense of and

expand an animal's sensorium and to efficiently use the information, it seems clear that aquatic mammals have successfully re-created the benefits of moving onto land.[8]

On land, mammals see to the horizon three miles away, but in the sea, they hear it. On land, cave-dwelling bats and two also-cave-dwelling bird groups evolved echolocation, which is a kind of sonar ability to create sound and understand, from the way it bounces back, what lies ahead. These echolocating land species, most or all of whom dwell in lightless caves, faced a similar visual difficulty in the cave as does life underwater, which proves that the brain, as always, does the best it can with whatever information it is given—sound, light, or touch are just fine, if it is all a brain can get.

At birth, the empty brain knows no stories. Seeing is an experience-based inference performed effortlessly and expertly by the adult human brain. The first time your brain lied to you was the second time you opened your eyes. Teenagers with cataracts, blind since birth, upon opening their eyes after cataract surgery and seeing for the first time experience featureless, depthless, shadowless blobs.[9] They could "see" the same number of photons as an expertly seeing adult but their brain sees nothing in the raw stream of light. Their brains had never seen a coincidence before. They had never walked past a brick wall's corner and noticed the lines of light bend around its edge or watched shadows elongate at dusk or compared, from all angles, sunlight shining through a tree versus the light given off by a tree full of tiny white candles.

All these kinds of stories—the girl's of her laughter, the blind of their first visions, and the newly awakened of their sleep—are how the brain hides its strange workings from its owner. Its apparent effortlessness comes at some cost. After the Italian explorer Marco Polo spotted a rhinoceros, in southeast Asia, while he was searching for what he believed to be a very real and very profitable unicorn, Polo wrote that unicorns are "not at all such as we describe them." His prior knowledge and hearsay about the legendary, valuable unicorn had changed what he saw in those brief moments, because even though most of the physical and behavioral features of the rhinoceros, like its weight, coloration, skin, location, and habits, did not match what Polo knew about the story of unicorns, it did have a single horn. The simple story won. His brain's personal history had changed what he saw. Likewise, the teenagers who saw for the first time experienced their own personal Promethean myth by stealing light and turning it into knowledge—their brains, like all of our brains once, got better over time at telling convincing visual stories as they noticed more events coincide between their movements, their brain's guesses, and the outside world. These stories, however, though they became more convincing and useful, did not necessarily become more true.

2

Like the Rise and Fall of Pinball

There are somewhere between one and 8.5 billion ways a brain can work, which means there are between one and 8.5 billion factorial ways of looking at consciousness.

One way to think about the evolution of a simple brain into a primate brain is like a power station that had to transition, all without ever once shutting down, from coal stoves to steam turbines to electric wires to nuclear to solar to an AI-powered, fuel-agnostic grid. One could easily imagine how dreadful this would be to maintain, with pneumatic tubes sticking out in the wrong places, control panels leading to nowhere, extant-parts lists, corrosive materials, and software incompatibilities.

A colleague once told me he preferred the analogy of a car, unable to turn off its engine, all while upgrading from a Roman chariot to a Tesla.[1] I prefer a third version of the story: pinball, because pinball machines were forced to evolve into both story and storyteller as, once, the brain did too, en route to consciousness.[2] Both are the result of a series of add-ons and user constraints impossible to plan for at the beginning. As such, modern versions of both have legacy strengths and legacy faults.

Like life, the game of pinball is never won but, instead, can be lost less badly at some times than at others. The threshold for what counts for any person as a sufficiently good score—or, as it was once described, the moment during play "when the sun goes down and the stars come out"—is as subjective as the threshold for any one person to live a good life.[3] Metaphors tend to stick to pinball machines like the gum on their undersides because every game, also like life, has so much that feels like chance but isn't and so much that feels like the opposite of chance but also isn't.

The last universal common ancestor, or LUCA, of all modern pinball machines can be traced to 1871, to a British inventor, Montague Redgrave, who was granted U.S. patent 115,357 for "Improvements in Bagatelle."[4] Bagatelle originated in France, in 1777, at a party thrown for Louis XVI at Château de Bagatelle; it can mean "a thing of little importance," "a very easy task," or "a game in which small balls are hit and then allowed to roll down a sloping board on which there are holes, each numbered with the score achieved if a ball goes into it, with pins acting as obstructions." Redgrave, in his improvements to bagatelle, evolved the game by adding a spring plunger, reducing the size of the ball to a marble, and inclining the field into which the ball is thrust.

Early-twentieth-century pinball games that evolved from bagatelle had none of the modern trappings of today's pinball machines, such as flippers, coins, or legs. They sat atop a desk, like large liquor barrels, and one changed the course of the ball as it sloped downward in three distinct ways, as one also does the course of history—slightly, by nudging it; heavily, with all one's weight; or accidentally, while attempting to do other things. To hinder people from simply picking up the

machine and modifying the ball's trajectory, designers added ungainly legs in the early 1930s and made the machines heavier, which only disadvantaged the weaker or less leveraged players.

And so, in 1934, a tilt mechanism was introduced to these still proto-pinball machines that prevented the player from moving the machine more than a set amount. Machines could also, for the first time, plug into electrical outlets, allowing them to produce lights and sounds and compete with the sensorium dazzles of motion pictures and World's Fairs, which were still a thing.

The flipper as we know it today was introduced, in 1947, with the game *Humpty Dumpty,* which had three flippers to a side and which, like the development of multicellularity, the human hand, and the atomic bomb, changed the world overnight. Everything preflipper looked instantly vintage. Flippers turned what had been a game mostly of chance into something that one could be good or bad at, compete at, wager on, fight about, cry over. *Humpty Dumpty* looked, to the American consumer, how the bow and arrow must have looked to an Ice Age man after seven hundred thousand years of variations on the hand ax. The "flipper bumper" introduced a way to control Brownian chaos; it meant, contra Newton, that entropy could be slowed and maybe even reversed. The ball had been exclusively a downward-trending thing, but now it could rise, like America, from the ashes of the worst world war man had seen. Unlike death and taxes, the steel ball's demise was no longer inevitable. Protestant America was introduced, again, to reincarnation.

By the 1950s, there were two major aesthetic and functional styles for the machines, which divided players into those who preferred the symmetric machines and those who preferred asymmetric ones. These early machines were much slower than modern pinball, less concerned with points, and were mostly concerned with the completion of small, tactical errands that required step-by-step precision and nonrandom sequences. When solid-state, muted digital computer boards made their way into pinball cabinets, the designers added back the sounds of the clinks, ratcheting gears, bells, and whistles to satisfy the nostalgia gene for the sounds of a writing-on-the-wall, bygone analog era. Slowly, the game became less about accomplishing a goal and more about the accumulation of arbitrary, outsize points. (Even today, some pinball machines reward points simply for playing, as if rewarding the act of inserting a quarter or pressing START alone, even if a single flipper is never hit.)

As the games became more and more about points and stringing those points into sequences of more points, there was a creative lull in the industry. Everything had been tried at least once and, even more damning, the world outside pinball was becoming a lot more interesting and interactive. Hollywood had just had its golden decade, with *Star Wars, 2001, Jaws, Apocalypse Now,* and *The Godfather.* Video games like *Pac-Man, Missile Command,* and *Frogger* were mainstream, cheap, social, and had mapped the goals of the species—avoiding predators, survival, crossing the road—onto buttons and joysticks.

And so, in 1986, pinball made one of its final, and riskiest, gambits: it became fiction. The game *High Speed* introduced a narrative arc and was an instant hit. As easily as most infants know how to grab and suckle, players immediately understood on an instinctive, motor level that the goals of *High Speed* were to change a stoplight from green to red, run the light, and flee from the police who gave immediate chase. Suddenly, the ball was not just a ball but a sports car. The player no longer saw, in the oily reflection of the glass top, their own face but rather the faces of Bonnie or Clyde. They were, for the first time, not playing *with* the small, metal ball but *as* it.

The entire narrative capacity of the human brain to find story in inanimate objects was suddenly brought to bear with every quarter. In the 1940s, psychologists Fritz Heider and Marianne Simmel made a short animation where simple geometric shapes like triangles, lines, and circles of different sizes would bounce around the screen and occasionally clump, bounce off each other, or follow the other shapes around.[5] When viewers were asked to describe what they saw,

they described creatures in conflict and told the tale of the shapes with genders, villains, emotions, and moral feats of high heroism in classic story and character arcs. Of course, these conclusions are illusory, as illusory as the stories we tell ourselves about *why* we laugh after doing so, even if it's because a surgeon has an electrode in our brains making it all happen. *High Speed* is to the Heider-Simmel illusion what fentanyl is to morphine and became, in its potency, more of a threat to the belief industry than the gambling industry, because what was happening in its players' brains resembled animism more than entertainment.

The flipper bumpers that at first only delayed *Humpty Dumpty*'s inevitable fall were now gas pedals; the player became the driver of a car the size of a pinball cabinet; the clean execution of a series of precise shots was the clean execution of thought, plan, and action, all while on the lam. The physiological RPM of the inner circulatory whorls and loops of hormones and multiball reward, tens of thousands of times more complex than the game itself, started whirring up in both the escaping player and those accomplices who stood watching, cheering, and abetting. A drain was no longer simply a lost quarter and a reset of points. It was the player's survival and, for the first time since Louis XVI threatened beheading unless his courtesans told him why the bagatelle ball was like the enemies of the French Republic, the game became a kind of storytelling guide to chapter and life. Thus the 1980s and early 1990s became an era of near-Cambrian explosiveness and introduced, through narrative and the miniaturization of circuit boards, the peak of pinball creativity, interest, revenue, and popularity.

In the mid-1990s, however, when both war and actual car chases could be televised live and other kinds of entertainment had become cheaper, more social, and took up less physical space, pinball was again in trouble. A game of virtual pinball became one of the bestselling games on early personal computers and yet, through emulation and software alone, the game kept all of the physics of weight, tilt, and acoustic echo that defined the earlier, realer machines. A series of speciation events on the corporate side of things caused revenue to decline and manufacturers to shutter until, in 1997, there was an official industry bottleneck. Only two major manufacturers remained.

In the late 1990s, forced to evolve or die, there was a push to incorporate the technology of fully digital arcade games that could project light onto reflective screens to give illusions like depth, texture, and horizon. One manufacturer decided to hedge, stealing both the idea and execution of lateral gene transfer from bacteria in order to cobble the best parts of video games and add them to pinball. Just as pinball had changed multiple times before in its history—legs, tilt, electricity, narrative—it would be forced, in a natural process of change and progress, to do so once again. A reflective mirror, computer monitor, and projector were placed over the standard pinball machine to create a chimeric centaur of old and new, which meant the display was no longer simply a passive addition to the playfield. It was a spectacular technical feat because the projected overlay, through a combination of sensors and illusion, seemed to interact with the ball as if by holographic magic. There was no longer a difference between the screen and the playing field or the analog and the digital. It was all a unified thing and could be glanced at easily without a break in attention as if through a single, cyclopean aperture. Bagatelle had become unrecognizable to all but the ancient geneticists.

This late-stage tack, however, was too expensive and too late. One company, whose slot machine and hospitality revenue had steadily increased along with the popularity and broader legalization of gambling, switched to making slots alone.

One of the most beguiling things about the human brain itself is that it, too, is an overlay of the history of analog and digital.[6] The chemistry of molecules moving rapidly within and between neurons had to become, at some point in the evolution of life, the movement of electrical current.

Then, bored, because predatory consumers demanded that change be gradual, that feature sets stay across generations, and that the best ideas be stolen and made into a new whole, life added bilateral symmetry, legs, reward circuits with arbitrarily dopaminergic counters, and a vestibular system that, in coordination with tilt circuits in the brain, kept us at an even keel. The analog, single-celled stuff is all still there—much as you can find, if you look hard enough, traces of bagatelle in modern pinball.

And then, at some point, when the competition was too great, and the odds looked dire, we, too, became fiction.

We used paddles and flippers to grasp and swim and carve and cook, and the metal ball became the storytelling self that moved through the world picking up points and victories. If there were time and resources to spare, the goal was to multiply and thus create, from only the insertion of carbon, water, and oxygen into the coin slot, the cyclopean aperture of story projection over the historical tabletop of parlor fancy. We got a consciousness—the projection atop a cabinet full of mechanical and analog bells and whistles that would otherwise do fine enough but are enhanced by the holography—and one that might, in a last-gasp effort at survival and relevance, one day also be fully virtual.

Why did evolution not keep single cells single? For the same reason pinball does not stay bagatelle. Why did pinball not stay mechanical or convert entirely to digital? For the same reason why the signals sent around the outer lanes of our brain are both part chemical and part electrical and why zapping a brain in a particular place, with electricity, can cause feelings of joy and mirth reliant on chemical concentrations: because there was a history to consider, because nobody thought to do otherwise, and because nobody came along and pulled the plug on the mammalian brain, ever.[7]

There is something strange about placing a quarter into a pinball machine and knowing that you will receive nothing physical in return. Those who recoil at the logic and waft of watching sunrise through cigarette smoke in casinos or those who judge others who place quarters into machines for the chance of more quarters but do not similarly judge pinheads are, evolutionarily speaking, irrational. By Darwinian standards, the gamblers at least have a chance, no matter how low, to gain more resources than they initially gave.

What reward is granted to the pinhead except the proxy hormones of victory, delayed defeat, high score, training data, or the chance at a free game? The surest way to never lose at pinball, after all, is to never play.


image
not
available