then I am myself the world

WHAT CONSCIOUSNESS IS AND HOW TO EXPAND IT

CHRISTOF KOCH

BASIC BOOKS

New York

introduction

Within seconds, my entire field of view became engulfed by dark, swirling smoke. The space around me fractured into a thousand hexagons and shattered. The speed with which this happened left no time to regret the situation I had gotten myself into. As I was sucked into a black hole, my last thought was that with the dying of the light, I too would die. And I did.

I ceased to exist in any recognizable way, shape, or form. No more Christof, no more ego, no more self; no memories, dreams, desires, hopes, fears—everything personal was stripped away. Nothing was left but a nonself: this remaining essence wasn't man, woman, child, animal, spirit, or anything else; it didn't want anything, expect anything, think anything, remember anything, dread anything.

But it experienced. Did it ever.

It saw a point of cold white light of unbearable intensity, unable even to conceive of looking away, as there was no "away from." There was no left or right, up or down, front or back, far away or close by. There wasn't a black canvas upon which the light existed, as there was no space. There were no other attributes: no color, no motion, no texture, no sound or silence, no

smell, no body, no pleasure, no pain. What there was, was a timeless universe convulsed to a blazing, icy light. That and a profound feeling of both terror and ecstasy, the awfulness of pure experience lasting indefinitely—for there was no perception of time. The experience wasn't brief or long. It simply was.

The remnant of my shattered mind perceived the sublime, the burning furnace of being.

I took a crucial insight from this experience concerning what is. Science explains the world of matter and energy, atoms and galaxies, using a handful of laws of physics and chemistry, supplemented, when dealing with the organic, by the idea of evolution by natural selection. This stance has served humanity extremely well in explaining the cosmos at large, its denizens and how they came to be. Extending this spectacular home run, science tries to retrofit the "subjective" world of experiences onto this "objective" world. That is, without adding anything else to its worldview, it wants to explicate consciousness as arising out of the mindless actions of a gazillion molecules. It is here, however, that science runs into metaphysical difficulties.

Indeed, this approach has it backward. Primacy goes to consciousness, not to the objective world. For me, in that timeless moment, there was no world, no body, no thing. What remained of my naked mind had left the gravitational field of the self, the force field inside which we live our entire lives. At that point, this experience was my only reality, just as it is now, the only reality I am directly acquainted with. Everything else follows from there, including the realist assumption of the existence of objects, out "there," independent of my experiencing them.

The ancient Greeks called the center of their world, the oracle at Delphi, the *omphalos*, or navel; for Judaism, the omphalos is the Temple Mount; for Christians, it is the nearby Church of the Holy Sepulchre. My omphalos is consciousness, the starting point from which I abduce everything else. Consciousness in the sense of having experiences such as boredom, a full belly, or terror.

As any story should, this book starts at the beginning, with the dawn of consciousness, our first subjective experiences. Did they occur in the womb, during birth, or as an infant? How would we know? Answering these questions has surprising consequences, especially for the bitterly fought war around abortion.

I next survey the dizzying varieties of experiences that make up the feelings of life: not just the phenomenal content of the classical five senses of sight, sound, smell, taste, and touch, but also the bodily senses that mediate information from joints, tendons, muscles, and organs; from pleasure and pain in all of its variants; and from the universe of emotions, of thoughts, and of self.

The reduction of the experience of the "self" during intense physical-mental exertions that demand full concentration is known to athletes, soldiers, or fly-fishers as being *in the flow*, *in the zone*, or *in peak experience*. The complete abolition of the sense of self has been reported by many, throughout history, as experiencing a bright light or a luminous expanse, losing the sense of having a body, being someone with a particular history and agency, and a slowing or even a complete cessation of the passage of time. Such experiences can leave deep contentment and awe or even ecstasy in their wake. When the mental

barriers that define us as individuals fade away, when the gravitational field of the self has lost its dominion over consciousness, the mind merges with the universe itself. The distinction between the individual and the world dissolves. They become one and the same. This feeling of oceanic boundlessness is expressed in the title of this book, *Then I Am Myself the World*, taken from *Tristan and Isolde* by Richard Wagner. This opera is one of the most rapturous works of music and singing composed within the Western canon, expressing a yearning by the eponymous lovers to transcend the mundane realm of experiences and attain a union with the ultimate realm of existence, in death. My first encounter with *Tristan and Isolde*, on a scrappy radio, came as a bracing shock. I was utterly transfixed, walking around in a daze for days.

I also reckon with how the mind interacts with the world. The mind is not the passive recipient of sensory data streaming in from eyes, ears, and other sensors, from which it derives an unambiguous description of what is out there. No, the mind constructs what it takes to be "reality"—seeing this chair, hearing music, feeling guilty—from explicit and implicit assumptions about statistical regularities in the world around and within us. These are called *priors* in the language of Bayesian reasoning, or *expectations* in layman's terms. Some are part of our genetic heritage, while others are learned early in life. These priors are usually inaccessible to conscious introspection.

Since each one of us has a different brain and grew up in a different physical, socioeconomic, political, linguistic, and cultural environment, each mind constructs its own, subtly different version of reality. No one has preferential access to the one

"true," "objective," and "unchanging" reality, although there is enough agreement about what is "out there" that we usually get along. I'm always reminded of this provisional, communal aspect of reality when zipping at high speed past traffic in the opposite lane of a nondivided highway—a small nudge to the steering wheel is all that separates life from death. Moreover, a small chemical nudge to your brain, about one-thousandth of a grain of salt of lysergic acid diethylamide, is all that separates this shared reality from a radically altered one.

This fundamental limitation of what the mind can know has important consequences for how we think about ourselves and how we interpret other people's actions, as laid out in Chapter 3. Indeed, the mind shapes the experience of anxiety, depression, and other mental conditions. But this limitation also comes with great benefits. *Neuroplasticity*, the modern understanding that the brain retains an ability to rewire itself, enables us to actively mold how we interpret and understand ourselves. We are not just helpless victims of fate but are the agents in charge of our own narrative, for better or worse, victorious or defeatist. This forceful shaping of our attitudes to events beyond our control has profound consequences for well-being and sickness.

Chapter 4 surveys philosophers' efforts to understand how the mental relates to the physical. How experience comes into the world has been an abiding mystery since the earliest days of recorded thought. Aristotle warned his readers more than two thousand years ago that "to attain any assured knowledge about the soul is one of the most difficult things in the world." Mind is radically different from the stuff that makes up the brain and everything else. Quantum mechanics and general relativity, the

periodic table of chemical elements, the endless strings of ATGC nucleotides that make up our genes—these appear to describe the physical, not the mental (I write "appear to" as quantum mechanics demonstrates that there are no observer-independent events, opening the door for consciousness to enter, at the ground level of reality). Yet we awaken every day to our subjective world of experiences.

The intellectual position that has garnered the most respect in contemporary Anglo-American philosophy departments is the ever more strident denigration or even outright denial of subjectivity. What is real is people talking obsessively about their experiences and acting on them; there is nothing above and beyond these speech acts and other intended or actual behaviors. The feeling part of consciousness, called *phenomenal consciousness*, is a big illusion. Philosophers in the know dispense with the "awful painfulness of my toothache" in the manner that Ebenezer Scrooge dealt with Christmas: "Bah! Humbug!" Furthermore, free will, our ability to deliberate about an upcoming fork in the road and to decide which path to take, is also thrown under this "illusion" bus. This rejection of the reality of lived experience constitutes a mind-boggling repudiation of what is immediately and indubitably given to us. It is also profoundly antihumanist, depriving us of those attributes that make us different from machines—indeed, equating us with machines.

It's an absurd adjuration, akin to Cotard's delusion, a rare psychiatric disorder in which able-bodied patients, often severely depressed, vehemently insist that some of their limbs are missing, that their bodies are rotting from the inside, or even that they are dead. When confronted with the fact that they are having a

conversation, right now, with their doctor, they do admit that the situation is a bit baffling, but the fact is that they are dead, and that's all there is to it. So it is with some contemporary thinkers who insist, against the evidence of their own senses, that experiences don't exist. Truly astounding—gaslighting all of us into believing that our experiences are fake!

Fortunately, consciousness can't be cancelled forever. The mental, having refused to yield, is returning with a vengeance. Indeed, the wheel is turning back to much more ancient understandings of experience, including *idealism*, the proposition that ultimately even matter and energy are mental manifestations, and *panpsychism*, the school of thought that all creatures, and perhaps even matter itself, are ensouled, that it feels-like-something to be anything, not just a human or even a bat. Modern science is supporting aspects of this remarkable turn of events.

Next, I briefly dive into a fundamental topic that might be surprising: existence, and how to define it, which is inextricable from defining experience. Or so argues *integrated information theory*, a quantitative, causal account of consciousness. Its development over the past twenty years has drawn in neuroscientists, neurologists, physicists, computer engineers, and philosophers as it makes startling, controversial (to some), but testable claims concerning who is conscious, of what, and why. According to the theory, consciousness is unfolded intrinsic causal power, the ability to effect change, a property associated with any system of interacting components, be they neurons or transistors. Consciousness is a structure, not a function, a process, or a computation.

However, the theory's insistence that consciousness must be incorporated into the basic description of what exists, at the

rock-bottom level of reality, has also drawn considerable fire from opponents.

The theory quantifies the amount of consciousness of any system by its integrated information, characterizing the system's irreducibility. The more integrated information a system possesses, the more it is conscious. Systems with a lot of integration, such as the adult human brain, have the freedom to choose; they possess free will.

Chapter 6 brings us to the brain, the physical substrate of experience. A worldwide quest seeks to track down the footprints of consciousness to its lair within the dense jungle of the central nervous system. The British molecular biologist Francis Crick—who codiscovered the helical structure of the molecule of heredity, DNA, in 1953 and deciphered the genetic code—and I worked for many years to identify the neuronal conditions sufficient for any one specific conscious percept. We championed a pragmatic, operational approach to the mind-body problem—the mysterious relationship between the nonmaterial mind and the material body—that has proven to be immensely fruitful. Today, more than thirty years later, I and hundreds of other scientists and clinicians are searching for these neural correlates of consciousness with a variety of tools and instruments in volunteers, patients, and laboratory animals, focusing on the back (posterior) regions of the neocortex, the vast lace of dense neuronal tissue layered and folded, like dough, across the outermost layers of the brain. This quest has not yet found its holy grail, as became apparent with my recent loss of a twenty-five-year wager against philosopher David Chalmers, he of the "hard problem of consciousness," the unfordable gap between the brain and the

mind. But as Chalmers admitted, it is only a question of when, not if, these correlates will be discovered.

Indeed, tracking these footprints helped established a beachhead in the mind-body problem, the construction of a consciousness detector, a first in history. Following a traumatic brain injury, stroke, or heart attack, victims can be severely impaired, unable to speak or otherwise signal their conscious state. Do they still harbor a mind stranded in a damaged body, or are they truly "not there"? Clinicians are testing a device, based on integrated information theory, that zaps the patient's brain with a magnetic pulse, records the resulting electrical reverberations via a net of electrodes on the scalp, and computes the complexity of this electrical pattern to infer whether the patient is conscious, like listening to the quality of the sounds a bell makes when rung. Diagnosing whether a mind is present and predicting the brain's likelihood of recovery gives succor to the patient's family and informs decisions on whether to withdraw life-sustaining therapy.

The two most extensive chapters of *Then I Am Myself the World* are given over to transformational experiences. These include religious, mystical, and near-death experiences that leave in their wake a profoundly changed individual. These extraordinary episodes of altered or expanded consciousness, triggered by backgrounding or even abolishing the sense of self, can lead to an epiphany, an enduring and pervasive change in a person's identity, core beliefs, and values. For those who experience ego dissolution, their view of reality and of life's purpose is permanently altered: they lose the fear of death and gain a detachment from material possessions and an orientation toward the greater good.

Religious, mystical, and near-death experiences are rare and strike out of the blue, perhaps as an act of gratuitous grace, as the Catholic Church would have it. Instead of waiting for such an event to occur serendipitously, some partake of substances to intentionally access otherwise inaccessible realms of experience such as psychic death and being at one with the universe. The first two decades of the twenty-first century witnessed a remarkable renaissance in the use of psychedelics, such as psilocybin, the active ingredient in magic mushrooms, that profoundly alter consciousness.

This psychedelic revival is based on the growing recognition that these powerful medicines, in combination with therapy, can ameliorate or even heal a wide range of psychiatric disorders, such as major depression, posttraumatic stress disorder, or general anxiety disorder. Taken responsibly and under the right conditions, they have enabled people to have highly meaningful, sometimes ghastly, but ultimately life-affirming experiences with remarkably few side effects. These experiences open a window of neuroplasticity, lasting for weeks, during which the brain can change its wiring, letting the mind modify deeply engrained attitudes. Sticky thoughts and prolonged ruminations, the hallmark of depression, low self-esteem, and anxiety—"everyone hates me"; "everything freaks me out"; "I will never find love again"—fade away, and a new, more wholesome attitude and outlook on life asserts itself. Psychedelics can teach us much about the mind and its substrate as well as facilitate human flourishing.

Psychedelic and mystical experiences can help us make peace with the inevitable, the ebbing of the stream of consciousness,

the dusk of experience. Given progress in the clinical arts, how we die has evolved over the past century. A brief chapter describes modern death, how it differs from traditional death, and deals with some unusual classes of events in the final hours of the brain, as it irrevocably shuts down.

Can technology provide us with the means to defer death into the indeterminate future? Can we reconstruct our aging brains in software, rejuvenate our minds in the digital realm by simulating it on a computer, thereby living practically forever? The penultimate chapter will discuss the future of human consciousness. Mind-uploading will only be achievable if computational functionalism, the metaphysical assumption that computations, executed on a computer, are sufficient for consciousness, holds. In this view, consciousness is simply a question of discovering the right algorithm. Under a different metaphysical assumption, consciousness cannot be achieved by mere computation as it is a structure associated with the physics of complex systems. If this is how reality is structured, then uploading a "mind" to a digital computer will end up with a deep fake: all action without what we hold most precious, subjective experience.

What about nonhuman, artificial minds, rivaling or even exceeding ours? This topic is treated last. Sentient machines have been a recurring theme in science fiction. In 2022, this topic burst into public view with the startling claim by a Google software engineer that the company's "large language model" was sentient and had to be considered a person with associated legal rights. The linguistic skills and knowledge of these models and their competitors, most famously ChatGPT and GPT-4 by OpenAI, trained on a vast trove of books and online documents

far beyond what any human can read in a lifetime, are astonishing by the standards of even a few of years ago. They write summaries, emails, jokes, (bad) poetry, computer code, letters of recommendation, and dialogue indistinguishable from human-generated material, including plausible-sounding fabrications. They are evolving at an astounding pace and will transform society in fundamental ways.

These chatbots seemingly constitute living proof of the dominant narrative of liquid modernity: the mind is software that can be as readily embodied within silicon wafers as it is within flesh, echoing a pernicious Cartesian dualism. Smart money in Silicon Valley thinks so, most engineers and many philosophers think so, and popular movies and TV shows reinforce this belief.

Against the grain, integrated information theory radically disagrees with this functionalist view. It argues from first principles that digital computers can (in principle) do everything that humans can do, eventually even faster and better. But they can never be what humans are. Intelligence is computable, but consciousness is not. This is not because the brain possesses any supernatural properties. The critical difference between brains and digital computers is at the hardware level, where the rubber meets the road—that is, where action potentials are relayed to tens of thousands of recipient neurons versus packets of electrons shuttled back and forth among a handful of transistors. As we'll see, the integrated information of digital computers is negligible. And that makes all the difference.

It means that these machines will never be sentient, no matter how intelligent they become. Furthermore, that they will never

possess what we have: the ability to deliberate over an upcoming choice and freely decide.

The brain is the most complex piece of self-organized, active matter in the known universe. By no coincidence, it is also the organ of consciousness. Unlike scientific advances in genomics or astrophysics, progress in understanding the brain and the mind directly relates to who we are, our strengths and infirmities, how we can live a contented life, and whether we partake of some larger, ultimate reality. Humanity is not condemned to walk around forever in an epistemological fog—we can know, and we will know.

Let me be your guide through the latest development on the mind-body frontier. Why should you trust me? By instinct and formal training, I'm a physicist with a minor in philosophy. I have practiced neuroscience for the past forty years. I spent a quarter century as a professor of biology and engineering at the California Institute of Technology in Pasadena. Subsequently, I joined the Allen Institute for Brain Science in Seattle as its chief scientist and, later, its president. I continue to work there, now as an investigator. I'm also the head of the Tiny Blue Dot Foundation in Santa Monica. It supports research into neuroscience-based therapies to help people understand that they live in mental worlds of their own making, whose limitations and biases they can overcome.

My guiding principle is the Royal Society's motto in London: *nullius in verba*, or "take no one's word for it"—in other words, rely on the original data rather than someone else's interpretations. That's why when I hear reports of people experiencing

altered states of consciousness via ecstatic dancing, running, meditation, or psychedelics, I try these techniques myself. By doing so, I can ensure that my understanding of these phenomena is based on direct experience, fortified by science, rather than hearsay.

Some call me a consciousness maven. However, I've always been wary of becoming overly confident and dogmatic in my views. So I strive hard to maintain an attitude of curiosity and humility—what Zen Buddhism refers to as a "beginner's mind." By being open to new ideas and perspectives, I hope to keep learning about the only reality we know, consciousness.