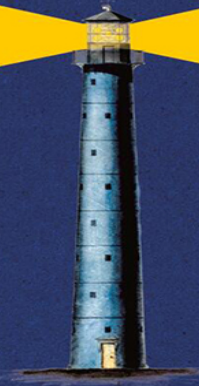


then

WHAT  
CONSCIOUSNESS IS  
AND HOW TO  
EXPAND IT



am

myself

the

CHRISTOF  
KOCH

world

then  
I am  
myself  
the  
world

**WHAT CONSCIOUSNESS IS  
AND HOW TO EXPAND IT**

**CHRISTOF KOCH**

BASIC BOOKS

New York





Copyright © 2024 by Christof Koch

Cover design by Chin-Yee Lai Cover image © SSPL/UIG / Bridgeman Images

Cover copyright © 2024 by Hachette Book Group, Inc.

Hachette Book Group supports the right to free expression and the value of copyright. The purpose of copyright is to encourage writers and artists to produce the creative works that enrich our culture.

The scanning, uploading, and distribution of this book without permission is a theft of the author's intellectual property. If you would like permission to use material from the book (other than for review purposes), please contact [permissions@hbgusa.com](mailto:permissions@hbgusa.com). Thank you for your support of the author's rights.

Basic Books

Hachette Book Group

1290 Avenue of the Americas, New York, NY 10104

[www.basicbooks.com](http://www.basicbooks.com)

First Edition: May 2024

Published by Basic Books, an imprint of Hachette Book Group, Inc. The Basic Books name and logo is a registered trademark of the Hachette Book Group.

The Hachette Speakers Bureau provides a wide range of authors for speaking events. To find out more, go to [hachettespeakersbureau.com](http://hachettespeakersbureau.com) or email [HachetteSpeakers@hbgusa.com](mailto:HachetteSpeakers@hbgusa.com).

Basic books may be purchased in bulk for business, educational, or promotional use. For more information, please contact your local bookseller or the Hachette Book Group Special Markets Department at [special.markets@hbgusa.com](mailto:special.markets@hbgusa.com).

The publisher is not responsible for websites (or their content) that are not owned by the publisher.

Print book interior design by Amy Quinn.

Library of Congress Cataloging-in-Publication Data Names: Koch, Christof, 1956– author.

Title: Then I am myself the world : what consciousness is and how to expand it / Christof Koch.

Description: First edition. | New York : Basic Books, [2024] | Includes bibliographical references and index. | Identifiers: LCCN 2023038544 | ISBN 9781541602809 (hardcover) | ISBN 9781541602816 (ebook) | Subjects: LCSH: Consciousness—Physiological aspects. | Consciousness. | Hallucinogenic drugs.

Classification: LCC QP411 .K629 2024 | DDC 612.8/233—dc23/eng/20231220

LC record available at <https://lccn.loc.gov/2023038544>

ISBNs: 9781541602809 (hardcover), 9781541602816 (ebook)

**E3-20240306-JV-NF-ORI**

# contents

[Cover](#)

[Title Page](#)

[Copyright](#)

[Dedication](#)

[Introduction](#)

[CHAPTER 1 The Beginning of Consciousness](#)

[First Light](#)

[Fetal Consciousness](#)

[To Sleep, per Chance to Dream](#)

[CHAPTER 2 The Varieties of Conscious Experience](#)

[A Panoply of Perceptual Experiences](#)

[A Universe of Emotions](#)

[The Stream of Consciousness](#)

[Loss of Self and Mystical Experiences](#)

[CHAPTER 3 We Each Experience Our Own Reality](#)

[Billions of Bespoke Realities](#)

[Perception Is a Construction of a Description](#)

[Mind over Matter](#)

[CHAPTER 4 Consciousness and the Physical](#)

[What Is the Mental?](#)

[Cartesian Dualism](#)

[Everything Is Physical](#)

[The Computational Mind](#)

[Experience Is Everywhere](#)

[CHAPTER 5 What Truly Exists](#)

[Only What Has Causal Power Exists](#)

[Properties of Any Experience](#)

[Properties of Physical Existence](#)

[The Great Divide of Being](#)

[CHAPTER 6 Consciousness and the Brain](#)

[Tracking the Footprints of Consciousness](#)

[Many Brain Regions Don't Support Consciousness](#)

[Look to the Neocortex](#)

[Stranded Minds in Damaged Brains](#)

[CHAPTER 7 Expanding Consciousness](#)

[Religious, Mystical, and Aesthetic Experiences](#)

[Psychedelic Experiences](#)

[Dying of the Light](#)

[A Common Neurobiological Substrate](#)

[CHAPTER 8 Transforming Lives by Transformative Experiences](#)

[Psychedelics and Their Potential to Change Lives](#)

[Expanding Our Perception Box](#)

[CHAPTER 9 The End of Consciousness](#)

[Appointment in Samarra](#)

[Modern Death](#)

[Anomalous Electrical Surges in the Dying Brain](#)

[CHAPTER 10 The Future of Consciousness](#)

[\*\*Mind Uploading: It's All About the Connectome\*\*](#)

[\*\*On the Difference Between the Real and the Simulated\*\*](#)

[CHAPTER 11 What Computers Can Never Be](#)

[\*\*Imitation Isn't Enough\*\*](#)

[\*\*Only What Truly Exists Can Freely Decide\*\*](#)

[\*Acknowledgments\*](#)

[\*Discover More\*](#)

[\*About the Author\*](#)

[\*Praise for Then I Am Myself the World\*](#)

[\*Notes\*](#)



*Note:* The events that prompted me to write *Then I Am Myself the World* were extraordinary episodes of the sort I had yearned for since youth but had not encountered until recently, in the seventh decade of my life. A word of warning though: the path to these lofty peaks, aided by ancient medicines taken within a communal setting in foreign lands, can be harsh and risky. As medieval cartographers expressed it, *hic sunt dracones*, or “there be dragons.”

This book is intended for informational purposes only. Certain sections of this book discuss psychedelics, including the author’s opinions about such substances and experience with them. However, nothing in this book is intended to, nor does it, constitute medical advice. The reader should consult with their own medical professional regarding all matters of their health and before relying on any research or adopting any idea or suggestion discussed in this book. The manufacture, possession, and distribution of psychedelics is illegal in the United States, and nothing in this book should be taken as encouraging or advocating any activity that could violate federal, state, or local laws. The author and publisher specifically disclaim all responsibility for any liability, loss, or risk, personal or otherwise, that is incurred from the use or application of the contents of this book.

*To all fellow travelers on the river of time who howl, bark, cry, screech, whine, bellow, chirp, shriek, buzz, sing, speak, or are without a voice. For it is only in compassion with all life that we can redeem ourselves.*

**Explore book giveaways, sneak peeks, deals, and more.**

[Tap here to learn more.](#)

**BASIC  
BOOKS**

# 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.

## CHAPTER ONE

---

### the beginning of consciousness

There is no birth of consciousness without pain.

—Carl Gustav Jung

**W**hat was your first subjective experience? Not this morning, but your first experience ever, way back, at the beginning of your life? Was it some dim sense of warmth, of distant muffled sounds, of rocking movements, suspended inside your mother's womb as she walked about? Or was it extreme discomfort, when you were ejected from this dark and confined paradise, squeezed painfully through a narrow tunnel, into a cold world of bright lights, loud sounds, and a desperate urge to breathe? Or did awareness arise by way of tasting your mother's milk, smelling her body's scent, and seeing incomprehensible blobs move about in your visual field as a neonate?

### First Light

Your first moments of awareness booting up mark the beginning, the first babblings of a brook that will eventually turn into the stream of consciousness, the ceaseless flow of musings, reveries, anxieties, regrets, reminiscences, worries, remonstrations, flashbacks, silent speech, and images that form the sound and fury of life itself.

The “stream of consciousness” is a compelling metaphor introduced by the late-nineteenth-century Harvard psychologist William James, father of American psychology and brother of the novelist Henry James. Does this stream originate in the marshlands of the immature fetal brain or in the headwaters of the more developed brain of a newborn infant or even a toddler? Answering this question is challenging given *childhood amnesia*,

the universal observation that adults do not have reliable memories of early childhood, before three or four years of age. Some insist vehemently that they do recall the birth of a sibling or some other notable happening, but they usually confuse an actual autobiographical or “episodic” memory of the event with the knowledge that such an event took place or photos of it. Young children can certainly form memories, but these fade as they age into their teens. Traumatic memories from physical, emotional, or sexual abuse leave deep albeit unconscious traces in the mind. But explicit, pellucid memories of your toddler years are forever lost in time. Whereas Sigmund Freud famously insisted that childhood amnesia was caused by repression of early memories with disturbing sexual content, today’s psychologists attribute this amnesia to the dual absence of language and abstract thought.<sup>1</sup> As those cognitive processes mature, so does your ability to lay down and recall explicit memories, marking the emergence of an autobiographical self.

Yet your lack of remembrance of things past does not mean you were not conscious. Just as you never recall the beginning of a dream (it feels like you’ve been dropped into some situation), it will have been the same with the initial, dim spark of consciousness. It is this moment that marks the beginning of conscious life, the inception of phenomenal existence. This existence for the conscious subject himself or herself constitutes absolute, or intrinsic, existence, a theme I will expand on and frequently return to.<sup>2</sup>

Life begins before consciousness does. You can be alive yet unconscious, an object to others (relative existence) rather than a subject to yourself (absolute existence).

Human procreation, like that of other sexually reproducing animals, begins with the fertilization of a female egg by a male sperm, forming a *zygote*. This single cell contains all the genetic information making up a new life, ground zero for the creation of an individual with a genetic identity unique among all eight billion living people.

From these humble beginnings, the zygote multiplies by dividing into two cells, then four, then eight, and so on, differentiating into the diverse cast of cells and tissues (ectoderm, endoderm, mesoderm, etc.) that constitute a budding embryo, which morphs into a fetus, is born as an infant, and grows into a toddler, then a child, an adolescent, and finally a sexually mature adult who initiates the next cycle of life. Development is

absurdly complex, multiplying by dividing and differentiating, over and over, until a person is formed out of thirty trillion cells—you.

Look back in evolutionary time to unravel this process. This reveals the staggeringly contingent nature of the processes that gave rise to you. You are the endpoint of an unbroken, billion-element chain of organisms, each arising from the preceding generation: your parents, grandparents, great-grandparents, and so on, reaching all the way back to the last universal common ancestor of all life (endearingly known as LUCA), a community of single-cell organisms thriving in a deep-sea vent some four billion years ago.

Now play this movie forward by visualizing every organism along this chain—starting with LUCA, morphing into a membrane-bound cell and then an eukaryote, making the jump to multicellular life, and from there becoming a worm, developing a backbone, turning into a fish, crawling onto land on four limbs, evolving into a small, nocturnal, insect-eating mammal that survived the asteroid crashing into the planet, transitioning into a primate, a great ape, into the last shared ancestor of chimpanzees and humans, into *Australopithecus afarensis*, into *Homo habilis*, interbreeding with *Neanderthals*, until, finally, turning into you, a member of *Homo sapiens*. If twenty-four adjacent frames of this chain of organisms, a hyperorganism existing in space-time, were to be displayed in sequence for one second, the resulting movie would play for over a year, a remarkable record of life on earth! It wouldn't do very well at the box office though, as most of the time little would happen. The story of civilization would be compressed into the last ten seconds of this year-long cinematic record; you, your parents, and your grandparents wouldn't appear until the last three frames, lasting the blink of an eye. This is the incomprehensible nature of deep evolutionary time, a story of breathtaking majesty.<sup>3</sup> As Charles Darwin concluded in his 1859 *On the Origin of Species*, “There is grandeur in this view of life, with its several powers, having been originally breathed into a few forms or into one; and that, whilst this planet has gone cycling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved.”

But I digress.

The point to remember is that you were alive long before you possessed

a developing nervous system, let alone the fancy three-pound brain housing your conscious mind as an adult. Leaving speculations about consciousness in engineered artifacts for later, a living organism is needed to support consciousness. But it is not sufficient. The central nervous system, assuming the organism has one, must be structured in a certain way; it must have a sufficient level of differentiation, complexity, or “something” to support the type of consciousness we are familiar with. What this “something” is will become clearer in a bit.

### Fetal Consciousness

Uncovering the dawn of consciousness isn't just a narcissistic endeavor or a Proustian search for lost memories; it has drastic consequences.

Consider the 2022 landmark decision by the US Supreme Court in *Dobbs v. Jackson Women's Health Organization*. It overruled two long-standing legal precedents concerning abortion (*Roe v. Wade* and *Planned Parenthood v. Casey*), invalidating the long-standing fetal-viability rule, which held that abortion should be legal until the fetus can survive, with proper medical assistance, outside the uterus at around twenty-three to twenty-four weeks' gestational age.<sup>4</sup>

The decision before the court involved lengthy, erudite, and passionately held legal, religious, historical, philosophical, and political considerations on both sides. One justification for overturning the viability rule was the argument that the fetus is conscious as early as fourteen weeks' gestational age and would thereby suffer extreme pain during abortion. This is a scientific-clinical matter that I commented on by joining an amicus brief, filed in the Supreme Court, to support Jackson Women's Health Organization.<sup>5</sup> The evidence for this startling claim was ultrasound imaging of the fetus in utero, demonstrating that the fetus reacts to touches of the mother's belly, to her voice, or even to painful surgical procedures, with facial grimacing and limb movement. These actions imply consciousness and therefore a capacity to suffer.<sup>6</sup>

It is true that the second-trimester fetus has rudimentary behavioral capacities, such as withdrawal from painful stimuli. Yet these are limited and stereotyped reflexes, called *nociceptive responses*, that adults show without any conscious awareness. Both nociceptive reflexes and pain can be

associated with action, including the release of stress hormones and an increase in blood pressure, but only the latter triggers a subjective, aversive conscious experience. Nociceptive responses occur during deep sleep, as when the sleeper withdraws a limb without waking up. Likewise, patients with severe brain injuries that leave them in a near coma can still withdraw their hand if a fingernail is pinched hard. Indeed, tiny fruit fly larvae react to a flame by bending away from the source of the heat. Yet few would advocate for larval sentience.

There is no question that the fetus, like other nascent organisms, is a living entity with regulatory feedback loops supporting stereotyped sensory-motor behaviors that protect it from harm. Yet, for a stimulus to be consciously felt as a painful experience—“ouch, that hurt”—a reflex is not enough. For a person to become conscious of a noxious stimulus, signals from pain receptors in the skin must be relayed, via the spinal cord and the thalamus, to the neocortex, where they set off an alarm perceived as painful. This requires a byzantine, sophisticated network of neocortical cells and their partners in satellite structures, such as the thalamus, closely associated with the neocortex.

The birth of neurons, called *neurogenesis*, starts around the fifth week and is largely completed by the end of the sixteenth week. For the most part, you were born with a full complement of nerve cells.<sup>7</sup> Yet these are immature and will continue to grow, to differentiate, and to extend their processes and tendrils to contact other neurons until well into adulthood. For example, neocortical neurons of a fetus are not properly wired up to receive any peripheral signals until about the thirtieth week. Until this time, the fetus responds to a stimulus such as a heel lance (a quick puncture of the skin to draw blood from the foot) but is unlikely to experience it. Indeed, a preterm infant born at thirty weeks gestational age—not a fetus anymore—will not even wake up following a heel lance.

Based on the way these circuits develop, peripheral pain signals can trigger reflexes but fail to ring the consciousness alarm until well into the third trimester. This implies that a preivable fetus does not experience pain—does not suffer. It still must cross the great Divide of Being separating something that does not experience from someone who does. The former is nothing to itself, while the latter is a subject, albeit still a rudimentary one.

In the final analysis, the supreme court decided *Dobbs v. Jackson*



*Women's Health Organization* not on scientific but on constitutional grounds, returning the authority to regulate abortion to individual states.

### To Sleep, per Chance to Dream

Relevant to consciousness is the discovery by neonatologists that the fetus—floating in its own isolation tank, connected to the placenta that pumps blood, nutrients, and hormones into its growing body and brain, and suffused by sedation-promoting substances—is asleep.<sup>8</sup> By the third trimester, the fetus is usually in one of two states: one state, characterized by fast breathing, an elevated and irregular heartbeat, swallowing, licking, moving the eyes, and isolated facial and body movements, is called *active sleep*, while the other, with slow breathing, a regular heart rate, closed eyes, and almost complete behavioral quiescence, is called *quiet sleep*. Active and quiet sleep morph within the first year after birth into rapid-eye movement, or paradoxical sleep, and deep sleep. The fetus is rarely awake. During its sporadic and brief periods of wakefulness, the eyes are wide open, with high muscle tone and lots of movement.<sup>9</sup>

The extended bouts of fetal sleep raise a fascinating question: If you, as an adult, wake up during certain phases of sleep, in particular during rapid-eye movement sleep in the early morning, you recall vivid, multifaceted experiences, containing the residue of recent events and more remote memories, in particular emotional encounters with family, lovers, friends, and foes. Does a fetus dream while in active sleep? If so, what would a fetus, who is a *tabula rasa*, a blank slate in terms of life memories, dream of?

Longitudinal studies in kindergarten- and preschool-aged children reveal that dreaming develops gradually, tightly linked to the capacity to imagine things visually, to speak, and to other visuospatial cognitive skills. Dreams of four- to five-year-old kids are static, plain, and mundane, with few characters that move or act, hardly any feelings, and sparse memories.<sup>10</sup> Extrapolating backward, what would dreaming be like for a third-trimester fetus, suspended in a warm and dark cave, with its visual cortex, the substrate for visual imagination, still immature? My hunch is that the fetus does not dream in the way you and I dream. But it is difficult to know for certain.

Placing sensitive instruments on the mother's belly to detect the weak magnetic signals produced by the developing brain shows that fetuses older than about thirty-five weeks are sensitive to statistical regularities in a series of sounds directed into the womb. They can discriminate a regular sequence consisting of a quartet of four tones (beep-beep-beep-*boop*, beep-beep-beep-*boop*, beep-beep-beep-*boop*, beep-beep-beep-*boop*) from a slightly deviating sequence (beep-beep-beep-*boop*, beep-beep-beep-*boop*, beep-beep-beep-*boop*, beep-beep-beep-*beep*). In adult patients, this is considered a signature of consciousness.<sup>11</sup>

A third-trimester fetus is unlikely to distinguish itself from the world; it is still egoless. The extent to which it has a primitive bodily awareness, such as pleasant sensations associated with warmth and nourishment through the placenta or painful ones, is at this stage impossible to ascertain. But it cannot be ruled out.

This observation has practical consequences for fetal surgery, a growing specialization in which surgeons carry out open or closed surgery in utero. Until the closing years of the twentieth century, this was done without anesthesia to minimize risks to the fragile and immature fetus. Given the possibility of pain experiences late in the third trimester, this practice has changed.<sup>12</sup>

Things transform abruptly during the dramatic and highly stressful events attending natural, vaginal birth. The fetus wakes up and is forced from the only home it has ever known into an alien world. A powerful surge of noradrenaline from the locus coeruleus deep in the brainstem—more powerful than any noradrenaline released during a skydive or an exposed climb undertaken by the grown adult decades later—and the cessation of sedation when disconnecting from the maternal placenta arouses the newborn. It draws its first breath, open its eyes, cries, and experiences an aerial world that assaults its senses with loud sounds, new smells, and bright lights.<sup>13</sup>

Newborns attend to sounds and sights around them, their gaze attracted to eyes and faces. Their visual acuity is quite low, but the basic thalamocortical circuitry necessary to support simple sensory percepts is in place. Their auditory capacities have been, and will continue to be, honed to their linguistic environment. Exposure to maternal speech in the muffled confines of the womb lets the developing nervous system pick up statistical

regularities distinguishing their mother language (literally) from others. Most impressive is imitation of facial and manual gestures by two- or three-week-old infants—mom wiggles her tongue, and the baby does the same a few seconds later. This requires both the dynamic storage of visual information and the capacity to control the tongue. In adults, such online storage of information is another hallmark of awareness, implying that babies have some measure of sensory-motor consciousness—that they can see, hear, and feel their bodies.<sup>14</sup>

Self-awareness and silent, inner speech develop much later. Just like dreaming, these are complex cognitive processes linked to linguistic processing that take years to mature, with boys usually delayed with respect to girls. If you have raised one or more sons, you will be familiar with the blank look on your teenager’s face when you ask him why he did something particularly stupid. At most you’ll get a shrug and an “I dunno; it seemed like a good idea at the time.”<sup>15</sup>

Understanding of the reasons for one’s actions is extremely limited in childhood. Adults can at least offer some plausible explanations. However, one of Freud’s lasting insights was that adults are often no better at truly understanding the inner sources of their motivation. You are a stranger to your mind.

The actual content that makes up the stream of consciousness grows as the child transforms, day by day, into an adult and encounters romantic and sexual relationships, social media, sports, games, music, movies, literature, alcohol, drugs, art, and work. All give rise to new classes of conscious experiences and add nuances, distinctions, and relations to existing classes of experiences. I turn to this extraordinary variety next.

## CHAPTER TWO

---

### the varieties of conscious experience

Lest you forget what subjective experience is about, let me take you on a tour of the vast universe of human experiences. I explicitly write *human* as each species experiences its own universe sculpted by its specific ways of sensing the world and its cognitive abilities—our experience of the world differs distinctly from that of dogs or bats. I will focus on the way experiences appear or feel, without analyzing the content of these appearances or feelings. This is known as *phenomenology*, a term derived from *phenomenon*, which means “that which appears.” When Eminem sings, “I can’t tell you what it really is, I can only tell you what it feels like,” he takes a phenomenological point of view.

Such a survey is necessary, both as a reminder of how strikingly diverse subjective feelings are and also because modernity has rediscovered states of consciousness that are strange, esoteric, half forgotten, associated with meditating, taking psychedelics, dreaming, and dying. The variety of recognized conscious states keeps on expanding, just like the size of the known universe.

To be conscious is to have experiences. I here distinguish two broad kinds of experiences: *percepts*, also referred to as *sensations*, which can be sensory and concrete or more thought-like and abstract, and *feelings*, which have an emotional character. This distinction between percepts and feelings is a useful one, with various qualifications I’ll come to later. Furthermore, much of adult human consciousness is taken up by reflecting on these immediate experiences, so-called *meta-consciousness*. Life is a stream of such interwoven percepts, thoughts, and feelings, waxing and waning, shifting, moving, metamorphosing, never resting.

The opening pages describe an experience of mine, one whose content is minimal—a bright point of light, terror, and ecstasy. Here is another experience of mine:

Though my calf muscles burn with fatigue, I will my body to continue to run along a trail through the lush forest. Looking far ahead to avoid roots and rocks, I hear the “caw, caw” screech of crows in the canopy above me, but come to an abrupt halt to admire the breathtaking beauty of a beam of the rising sun slipping through the dark forest and striking a moss-covered tree, reminding me of a famous poem.

This moment is but a snippet, one out of the endless string of experiences making up the stream of consciousness. In 1902, William James gave a series of lectures at the University of Edinburgh titled *The Varieties of Religious Experience*, describing and classifying different types of spiritual awakenings, mystical experiences, and religious sensibilities. This chapter will do something similar for all experiences, tout court, mundane or exalted, profane or sacred, sane or lunatic.

### [A Panoply of Perceptual Experiences](#)

This survey starts with the five traditional Aristotelian senses of sight, touch, sound, odor, and taste. Sensory receptors in the eye, skin, ear, nose, and tongue transduce the relevant physical signals—photons, mechanical pressure, sound waves, meaningful molecules—into neuronal signals relayed to the brain proper, where they morph, within a quarter of a second, into the conscious percepts of the face of a certain ex-president, of too-tight sneakers, of the four hammer blows of fate in Beethoven’s Fifth Symphony, and of the invigorating smell and bitter flavor of freshly brewed coffee.<sup>1</sup>

Every sensory modality has its own highly structured perceptual space. Color in most people can be characterized by variations along three dimensions, such as the intensity of green, red, and blue, with their combination characterizing the color you see. This can ultimately be traced back to three distinct pigment molecules in the photoreceptors of the eye, each one responding best to a different range of wavelengths of the incoming light. Color-blind humans, as well as most other mammals, only have two photopigments, while mantis shrimps have eleven or more. It all depends on the particular ecological niche a particular species evolved in.<sup>2</sup>

Taste comes in five basic flavors: sweetness, sourness, bitterness,

saltiness, and umami. The last one wasn't recognized until the work of Kikunae Ikeda in Japan in the early twentieth century. Umami is associated with the savory taste of broths, cooked meats, fish, and soy sauce. This quintuplet of canonical flavors depends on five classes of taste receptors found in taste buds located on the tongue, on the roof of the mouth, and in the lining of the upper throat.

Visual, auditory, and somatosensory percepts are bound to space—when you see, hear, or sense something, you usually see, hear, or feel it at a specific location.

Then there is imagination, the ability of the conscious mind to conjure images, scenes, voices, and music previously encountered or fantasized. Close your eyes and imagine the Statue of Liberty standing guard in New York Harbor: In which arm is she holding the torch? What is she carrying in the other hand? Could it be a laptop? Imagined sights and sounds feel like sights or sounds from the external world, albeit usually fainter, less vivid, with fewer details.

Next up are experiences associated with sensors distributed throughout the body, so-called interoceptive perceptions. These include awareness of the position and angle of your head, limbs, fingers, and torso as you dribble a basketball or type on your phone. The senses of proprioception and balance are underappreciated but are critical to standing, walking, running, or otherwise moving your body and limbs in space without effort or thought. Other bodily sensors mediate the feelings of hot and cold and the diverse sensations associated with individual organs: viscera (when you are hungry or nauseated), heart (when, lying in an isolation tank, you can feel and hear your heartbeat), lungs (when you pant running up a steep hill), bladder (when you feel an urgent need to pee), and bowels. You are usually oblivious to these sensations except if they sound an alarm by becoming uncomfortable or painful, or you focus on them, as during yoga class. And then there is the entire phenomenal space of sexual sensations associated with arousal and orgasm in all its varieties. Unfortunately, these pleasant feelings are usually fleeting; no one, as far as I know, experiences chronic pleasure, but legions suffer from chronic pain.<sup>3</sup>

The boundary between experiencing the outside, seeing a bird in the distance, and the inside, sensing something stuck between your teeth, is fluid, with the sense of touch, mediated by receptors embedded in the skin,

partaking of both outer and inner worlds.

Collectively, these bodily sensors build up the experience of a spatially extended, highly mobile and jointed body, a set of sensations that anchors you in the physical world as much as the visual plenum of the world you see locates you in space.

When your body is injured, inflamed, or otherwise under attack, you feel pain. These experiences have their own subjective signatures depending on the affected organ—a migraine feels different from a toothache, a strained ankle, or a noxious bellyache. Pain has what psychologists call negative valence or affect: it hurts. Unpleasant feelings can range in intensity from the transient and mildly irritating to the long-lasting and searing.

Think of interoceptive percepts as a form of self-monitoring. If things are amiss—you haven't hydrated in a while or are developing a blister—your nervous system generates an alarm, like the gas light in the car's dashboard turning red, notifying you that you're running low on fuel.

Self-awareness, or self-consciousness, is the subjective experience of one's own desires and emotions. Its most prominent feature is a voice inside your head, though not everyone has such a voice. It keeps up a running internal monologue about yourself and the people around you, how they look and act, their motivations, and the environment you find yourself in. The voice ruminates and judges, chews on events in your past, broods about perceived injustices, worries by imagining bad things that might happen, and rehearses the future.

This “I,” with its incessant silent speech, chattering about ten times faster than were you to speak aloud, plays an increasingly dominant role as you grow into adulthood.<sup>4</sup>

The self can become aware of its own experiences, as in “Hmm, my toe is hurting. Maybe I should have bought those shoes in a larger size.” This is *meta-consciousness*: becoming conscious of an experience, a form of self-reflective introspection, consciousness being mirrored by the self. This operation can be applied recursively to itself; that is, you can become conscious that you're conscious of your toe hurting. Meta-consciousness involves selectively attending to an experience, which changes the character of the experience, rendering it more salient. It will typically also involve judging whether this experience is good or bad for the self. *Mindfulness* is the practice of being in the here and now, fully aware of the momentary



experience, but without judging it to be good or bad. It is the act of simply accepting the experience for what it is, decentering the self.

The sense of self is bundled with powerful cognitive abilities. One is the recall of episodic memories, events that involved the self in the recent or more remote past (“I met her at that conference last year”). This is the remembering or autobiographical self. Another is introspection, consulting short-term memory to explain to myself why I’m upset at my coworker. However, as you know firsthand, inferring the true motives for why you said what you said and did what you did is challenging, distorted by wishful thinking and other cognitive biases.

Most importantly, the self wants things: fancy food, a new car, a beautiful sexual partner, a promotion.<sup>5</sup> These desires are expressed in feelings of intentions. Other experiences associated with self include agency (“I made the decision”) and ownership (“it was my finger that pulled the trigger”). Each one of these experiences has a distinct flavor and can be more or less pronounced.

As you become older, you get better at differentiating these experiences and relating them to each other. Your self-awareness grows, and you develop a greater understanding of your emotions, even subtle ones, allowing you to infer the often masked or covert causes of your actions and words. This refinement of your sense of self is a valuable trait that can come with maturity, part of growing wiser with the years.

Whether thoughts, such as “I haven’t seen grandma in a while,” feel-like-something beyond imagining your grandmother’s face remains unclear. Indeed, some psychologists argue that much of thinking is carried out unconsciously; what is consciously accessible are the projections of these thoughts onto the visual, auditory, or linguistic processing machinery in the brain. That is, when you are thinking, “I need to book a ticket to Venice,” accompanied by images of the lagoon of Venice, of an airplane, perhaps of a map of Italy, the cognitive work associated with planning such a trip and turning these plans into action is done away from consciousness’s limelight.<sup>6</sup>

External and internal percepts make up the scaffolding of diurnal and nocturnal life, during both waking and dreaming. Of course, experiences are not constrained to the present. You revisit the past, fondly lingering over the memory of a dinner with a lover or agonizing over a dispute with a



colleague, and imagine the future, planning an office memo or fantasizing about sex. These mental activities, concerning what was, what might have been, or what could be, play the movie of your life backward or forward or insert alternative scenes. In most, such mental time travel is dominated by visuospatial imagery and inner speech: you see and hear things in your head but only rarely smell or taste them or feel a ghostly touch.

### [A Universe of Emotions](#)

The second broad category of conscious experiences, traditionally called feelings, relate to emotions. They range from the basic ones—anger, disgust, fear, happiness, sadness, and surprise—to braided mixtures, such as acceptance, acedia, affection, aimlessness, amusement, anger, angst, anguish, annoyance, anticipation, anxiety, apathy, arousal, and awe (and those are just the ones starting with *a*).

With the world having gone through the Covid-19 pandemic and a war in Ukraine with a nuclear-armed Russia, you are familiar with anxiety or apprehension. It is a brain state that arises from specific circumstances. It goes hand in hand with negative feelings, such as unease and worry; unwelcome behaviors, such as restlessness, irritation, withdrawal, and sleeplessness; troublesome cognitive effects, such as rumination, inability to concentrate, hypercritical self-judgment, and catastrophizing; and a litany of bodily symptoms, including headaches, vague abdominal pain or nausea, heart palpitations, or shortness of breath. What is true for anxiety holds for any emotion. Each one is constituted by a gamut of conscious feelings, behaviors, cognitive modes, and physiological effects.

Emotions are not restricted to self. Empathy, the ability to vicariously feel what others are going through, be they family members, friends, strangers, pets, or even wild animals, is a paradigmatic social emotion, the glue that enables humans to live, relatively peacefully, in large groups. The deliberate development of compassion with the suffering of all creatures is one of the redeeming features of humanity.

It is rare to experience an elementary emotion in isolation, like “pure rage.” Most feelings are composites. Take *saudade*,<sup>7</sup> from the Portuguese word for longing for something irreversibly lost, like the forsaken comfort of a childhood home, suffused with a warm glow and fond memories (the

paradigmatic *et in arcadia ego*). Portugal has an entire music genre known as *fado* that epitomizes *saudade*; it combines sadness, longing, regret, nostalgia, anxiety, and dread.

Emotions and percepts differ in a couple of ways. While percepts are short-lived (you quickly adapt to any sustained stimulus, such as a pungent odor or the rumbling of a car engine and cease to smell or hear it), feelings usually ebb and flow slowly and can persist for long times. At some point, an emotion becomes a mood, like in the Wong Kar-wai landmark movie *In the Mood for Love*.

Percepts, by and large, are not experienced as good or bad. Often an image, a song, or a smell will trigger a memory with powerful positive or negative emotions, but shorn of such associations, percepts lack the affective component that marks emotions: seeing or hearing is experienced as neutral. Emotions, on the other hand, are defined by their valence; these can be either negative, such as fear, or positive, such as romantic love. The associated affect can range from the mild frisson when flirting with the attractive office intern to the overpowering *affaire de coeur* for which you risk your job and marriage. It demands considerable effort to control the impulses unleashed by such strong emotions lest they become destructive. Much of literature, film, opera, and life itself is about the drama that ensues when this self-control fails.

It is one of the tender mercies of modernity, albeit little appreciated, that the experience of acute pain caused by inflammation, infection, and trauma has receded from dominating life in the way it did for previous generations due to the invention of anesthesia, disinfection, and effective, if sometimes addictive, pain medication. What remains are chronic pain and widespread emotional distress, whether in response to external events, such as the death of a loved one, the loss of a job, a traumatic incident, social isolation, discrimination in some form, or due to endogenous causes such as depression, anxiety disorders, or intrusive, unwanted thoughts. Emotional anguish can dominate the lives of sufferers and, in their search for a reprieve, drive them to seek solace in alcohol, drugs, violence, or suicide.

The distinction between a percept and an emotion is fluid and shaped by context. Take pain. A stimulus, say a burning piece of wood unexpectedly falling onto your leg, triggers an immediate withdrawal and a sharp sting. Ouch! That hurt! However, in trying circumstances, say during a

competition or on the battlefield, you can be oblivious to such injuries. For me, this happens when I run trails next to a sudden steep drop, like a cliff, or on a ridge high up in the Sierras, when my mind is utterly focused and at ease. Or, in my younger years, when I climbed big walls, concentrating on tiny indentations to place a toe or wedge a finger, I would remain unaware of abrasions and cuts until past the crux, when the adrenaline had abated.

An extreme example of the dissociation between bodily injury and pain is the self-immolation of the Buddhist monk Thich Quang Duc in 1963 to protest the South Vietnamese regime's campaign to repress Buddhism. What is so singular about this event, captured in haunting photographs that remain among the most readily recognized images of the Vietnam War, is the calm and deliberate nature of his act. As fire consumed his body, Quang Duc remained in a perfectly composed lotus position, without moving a muscle or uttering a sound, until his blackened corpse toppled over, a stunning testament to the supremacy of mind over body.

What about the opposite, pain without bodily injury? The breakup of a long-term relationship leads to feelings of emptiness, regret, anger, hurt, and longing. The heartbreak that the rock group Nazareth wails about in its song "Love Hurts" leaves its echo in the brain's representation for bodily pain that can be detected via a magnetic scanner.<sup>8</sup> A form of societal breakdown, signaled by the surging epidemic of "deaths of despair" from suicide, drug overdose, and alcoholism among rural working-class men and women, is haunting America. Their experience of life is painfully and chronically bleak, driving them to self-destructive behaviors, a leading cause of the shocking decline in life expectancy in the United States.<sup>9</sup>

Thus, while subjective pain and objective bodily damage are usually related, they can come apart. People can have bodily trauma without pain; conversely, there can be emotional trauma sparked by the loss of a relationship without any injury or painful stimulus.

### **The Stream of Consciousness**

With the singular exception of chronic pain and distress, the content of consciousness is usually fleeting. As anyone who has ever tried meditation knows all too well, it is surprisingly difficult to hold the mind still. Experiences are short-lived. Like a hummingbird, the mind continuously

darts about, from thoughts about calling mom to a song on the radio to an unbidden childhood event or an aleatoric scene from a movie, all in defiance of conscious control. It takes years of mental training for your conscious mind to remain in the here and now. This is what makes meditation, fully focusing on nothing but your breath through your nostrils, so very difficult.

This is one of the reasons why William James's metaphor of the stream of consciousness, flowing faster or slower, with emotions, like water, rising and falling, is so powerful. Art has made much use of this phenomenon. A few decades earlier, before James's work, Richard Wagner's heady operas, such as *The Valkyrie* or *Twilight of the Gods*, prefigured this metaphor in continuously evolving soundscapes—so many voices, leitmotifs, and strains intertwined into a single, vast stream encompassing empathy, lust, love, fear, hate, anger, desire,<sup>10</sup> will to power, regret, and compassion—merging and separating, waxing and waning, rising and falling, flowing on, like life itself. Half a century later, Marcel Proust, Virginia Woolf, and James Joyce would perfect the literary equivalent of the stream of consciousness, the interior monologue of the narrator.

As a denizen of the twenty-first century, you only rarely experience the spontaneity of this stream. Whenever you are unoccupied for more than a few seconds, like an addict, you reach for your phone. But even when engaged in a task—working on a spreadsheet, listening to a podcast, doing the dishes, driving—your attention often lapses, and your mind wanders to more pleasant realms or more urgent tasks. In a classical daydreaming experiment, people were given applets for their smartphone that randomly queried them about what they were doing and thinking at that exact moment in time and how they felt. This revealed the remarkable fact that mind wandering occurs almost half of the time and during most activities (except, supposedly, during sex): the less engaging the task, the more the mind wanders.<sup>11</sup>

The stream metaphor is, despite its powerful appeal, misleading in at least three ways. First, there is evidence to suggest that each experienced “now” is a discrete snapshot in time, akin to watching a film, which is essentially a series of stills, with each stationary frame rapidly being replaced by the next one. How long each moment lasts by the clock can vary, subject to attention, arousal, motivation, and so on. This would

explain moments of protracted duration reported upon in the context of accidents, falls, or other life-threatening events—“When I fell, I saw my life flash before me” or “It took him ages to lift the gun and aim at me.”<sup>12</sup> Second, consciousness of the passage of time, slower or faster, can be altogether suspended, as during psychedelic experiences when the passage of time may cease altogether, a topic I pick up in a few pages. Third, the flow of the stream of consciousness, or, perhaps more accurately, the string of conscious moments, like pearls on a necklace, is periodically interrupted by episodes of unconsciousness, when you fall asleep. Trying to catch this transition from wakefulness to sleep, from being to nonbeing, is challenging. You are there one instant and then no more. The novelist Haruki Murakami puts it poetically: “He calmed himself, shut his eyes, and fell asleep. The rear light of consciousness, like the last train of the night, began to fade into the distance, gradually speeding up, growing smaller until it was, finally, sucked into the depth of the night, where it disappeared.”

Conversely, when you awaken from deep sleep, it feels like you came from nowhere into existence. Indeed, when volunteers are roused in a sleep laboratory during episodes of deep sleep and are asked what went through their mind just prior to waking, they typically respond with “nothing”—as well they should, as the conscious mind was extinguished during that time.

A possible confounder is that you may have been conscious but forgotten about it. This can happen following alcohol-induced blackouts, when you later desperately try to recall what you said or did during last night’s pub crawl, or during propofol-induced sedation, which is shallower than surgical-level anesthesia, when you are definitely unconscious. Both reduce arousal and prevent memory encoding without necessarily eliminating awareness. Sedation is beneficial for minor surgical procedures, such as a colonoscopy, as patients can respond purposefully to instructions and breathe by themselves but won’t remember anything. So memory is yet another variable that needs to be controlled for when seeking to isolate consciousness from other processes.<sup>13</sup>

During sleep, usually when the eyes move in a particular jerky manner (giving this phase its name, rapid eye movement, or REM, sleep), the mind roams while the body sleeps—you play, explore, fight, love. The content of

dream consciousness is primarily visuospatial—you see or hear imagined or real scenes, usually mundane ones, populated by people and pets, living or dead. Most dreams depict mundane events; the next day, you tend to remember only the bizarre ones or the ones strongly colored by love, desire, fear, or anxiety. Dreams feel as real as life—the primary distinction between dreaming and waking consciousness is an absence of a sense of self, insight, self-reflection. You aren't surprised that you can fly, walk through walls, or meet long-dead animal companions, lovers, parents, or siblings.<sup>14</sup> You are along for the ride, watching a movie that someone else is directing.<sup>15</sup>

Some sleepers do, on occasion, “wake up” inside their dream, realize that they are dreaming, and take limited control of events, becoming their own “dream director.” Such lucid dreaming is the basis of a compelling multilevel world-building mythology in the movie *Inception* by Christopher Nolan.<sup>16</sup> Some meditative traditions practice something related called dream yoga. Unfortunately, I have never woken up inside a dream and so have no firsthand experience to fall back on. In some sleep research laboratories, lucid dreamers are trained to communicate with observers at the bedside via controlled eye movements (move your eyes three times left and right) while continuing to sleep, exploring this unusual state of consciousness.<sup>17</sup>

Dreams are an extraordinary product of brains, with most of the attributes of everyday subjective experiences. While they last, dreams are as real as life itself. We take dreaming for granted; yet its function remains mysterious.

Recent research discovered a startling facet of life—brief periods during which consciousness is absent while the body carries on its well-rehearsed duties: driving, doing the dishes, or reading a long and uninspiring office memo. To an observer, everything looks normal while the subject is, in fact, zoned out.<sup>18</sup> During these episodes of mind blanking, lasting for seconds to a minute or more, people don't report experiencing anything. Their mind is either devoid of content or offline. Training yourself to be mindful of your inner life, you can catch yourself staring at a point in space, blanking, without perceiving anything. The incidence and duration of mind blanking increases with sleep deprivation and may, indeed, be a manifestation of either episodes of micro sleep or a phenomenon called local sleep, when

only some regions of the brain go offline under accumulated sleep pressure.<sup>19</sup>

So between mind blanking and the previously discussed mind wandering, you spend a large fraction of your waking hours either zoned out or daydreaming!

Less commonly encountered in our always-on culture are states of idleness, languor, boredom, tedium, lassitude, and apathy, in which we let time pass without any ulterior motive. It's the mindset of the deliberate flaneur, who loafs around while enjoying the experience of wasting time. Otis Redding sings evocatively,

*I'm sittin' on the dock of the bay  
Watchin' the tide roll away, ooh  
I'm just sittin' on the dock of the bay  
Wastin' time*

Modernity frowns on stillness, on simply being and watching the world go by, and favors busyness. This comes at a cost to well-being.

### **Loss of Self and Mystical Experiences**

Finally, there is a class of conscious experiences that come with a degree of lucidity uncommon in normal life. By and large, their mere existence was denied by science until the last century, and their study was consigned to the intellectual hinterlands of psychology, at the intersection of spiritualism, the paranormal, and the esoteric. But they are nothing of the kind. They are rare yet genuine subjective phenomena that constitute the outer fringe of what humans can experience. Key to all is loss of self, including the sense of ego and the body.

Consider *flow*, a concept and state studied by the psychologist Mihaly Csikszentmihalyi. It is a mental state in which you are totally engaged with the world while only dimly aware of yourself. It was only when writing this book that I realized that I had been chasing flow and the attendant loss of self my entire life. As a young man, I loved going out dancing to old-fashioned rock 'n' roll for hours on end. My dance partner would put her hands around my neck and jump onto my hips; I would pick her up and



twirl her around my body, onto my shoulder; she would slide between my legs—all in synchrony with the music. We were lost to the world, a symphony of sound and motion, in a trance, without any drugs or alcohol beyond an occasional beer to rehydrate, until the club closed. Later came flow during big wall rock climbing, rowing a shell, biking at all hours of the day and night through dense urban spaces, running along a steep cliff, or building a 3-D structure out of pasta during a management training exercise.<sup>20</sup> I enjoyed, and enjoy, all of them. Others attain a state of flow while racing, fly-fishing, hunting, or fighting. All these activities require a complete focus; all attention is on the task at hand, time slows down to a taut present, and the sense of self fades. That inner nagging voice, your personal critic forever reminding you of your failings, obsessively chatting and commenting on everything, is stilled. Flow conveys a sense of deep contentment. You are in the world and of the world, engaged with your entire being, in the moment, a joyful experience, a selfless state of grace.

Athletes and soldiers seek to lose themselves in flow because being “in the zone” comes with peak performance, devoid of extraneous distraction, with smooth and fluid movements, seamlessly integrating sensing and acting. You train and train, and when the moment for action arrives, you let go and trust your body and its accumulated wisdom.

A more profound loss of self occurs in near-death experiences, sudden religious conversions, mystical experiences, and deep meditation, as well as with ingestion of psychedelic substances (more on these later). This loss of self leaves in its wake a long-lasting residue of thankfulness for being alive and of becoming one with the world.

During these episodes, the stream of consciousness freezes, suspended in place. People often report the presence of something vast, numinous, ineffable. The philosopher Arthur Schopenhauer, alluding to the writings of Indian Vedic literature, used the memorable phrase “piercing the veil of the Maya,” as that is the feeling that these experiences evoke—having come face-to-face with ultimate reality.

Some consider these states a higher form of consciousness. Perhaps. I call them transformative to convey that they differ from everyday experiences in a qualitative manner and not just in the sense of being rare, such as savoring a bottle of 1928 Château Lafite Rothschild that costs thousands of dollars. For the latter is simply an uncommon variant of taste



and smell (perhaps with more “personality” or a longer finish than a typical bottle of Bordeaux), while transformative experiences achieve transcendence, conveying a sense of equanimity, a feeling that everything is as it should be. They transform the life of the experiencer to the extent that the sense of self is extinguished. Experiencing the world with the “I” out of the way—an “I” that always wants something, desires something, fears something—provides the priceless gift of peace of mind. Given their compelling nature, I will return to transformative experiences and their therapeutic and existential benefits later.

The introduction opened with one such experience during which “I” ceased to exist. The remaining lucid and clear mind, having escaped the gravitational bound of the self, was a nonembodied consciousness. The habitual distinction between me and my experience, the apprehender and the apprehended, the subject and the object, the knower and the known, had vanished. Some contemplative Buddhist practices refer to such states without a center as nondual states of consciousness.

This brings up the question of whether an experience must always be about something. Is it possible to be conscious of nothing at all? If so, how would that differ from being unconscious? Can there be an experience not involving seeing, hearing, fearing, wanting? This might bear some resemblance to what long-term practitioners of Buddhist meditation describe as sheer or naked awareness attained during *samadhi*, the complete cessation of all mental content, quieting and stilling consciousness until it is suspended in a luminosity of nothingness. This is a quest common to many spiritual traditions—the mind as an empty mirror, beyond the ever-changing percepts of life, beyond self, thought, hope, desire, and fear.

Mystical experiences do share one trait with other experiences, like tasting a slice of cold, leftover pizza: they are ineffable, having something that can’t be put into words, something impenetrable.<sup>21</sup> The philosopher Thomas Nagel formulates this inexpressibility in a famous essay, “What Is It Like to Be a Bat?”: “An organism has conscious mental states if and only if there is something that it is to be that organism—something it is like for the organism.” This explanation has risen to the level of a definition of consciousness in the eyes of many; it does presuppose a conscious reader who knows what “it feels like” means. Whether a nonconscious entity, such as an advanced artificial intelligence, could ever make sense of any

definition of consciousness, except in the trite sense of regurgitating what other people had written on the topic, remains unanswered and is perhaps unanswerable.<sup>22</sup>

Ending this tour, I want to reemphasize the miraculous existence of any form of consciousness by paraphrasing Ludwig Wittgenstein:

Not how consciousness is, is mystical, but that it is.<sup>23</sup>

That you are intimately acquainted with the way life feels is a brute fact about the world that cries out for an explanation.

But do you and I really feel similarly about life? Do we experience the same world, a common reality, or are we like monads, each one closed upon itself, each experiencing its own reality? That's what we'll go into next.

## CHAPTER THREE

---

### we each experience our own reality

Things are known not according to their natures but according to the nature of the one who is comprehending them.

—Boethius, *The Consolation of Philosophy*

That we each experience our own world was viscerally driven home by #TheDress that exploded on social media and went viral in 2015. Google this internet meme to remind yourself—a washed-out photo of a dress worn at a wedding with horizontal stripes triggered a vigorous dispute among friends and family as to whether it was “white with gold lace” or “blue with black lace.” #TheDress is not one of those bistable illusions that everyone can either see as a rabbit or a duck, an old or a young woman, a crate in one of two orientations, and so on. No, many, like me, insist that the dress is obviously and unambiguously white and gold, while others, equally emphatically, perceive it as blue and black. Ask your family and friends how they experience it; the divergence in what seems like a simple matter is striking.

#TheDress offers a concrete lesson in phenomenology. The frequently asked question “What is the real color of the dress” doesn’t have an “objective” answer, as color is a construct of the mind of the beholder.<sup>1</sup> What the philosopher Immanuel Kant refers to as *das Ding an sich* (i.e., the thing-in-itself) is inaccessible. Science tell us there is a fabric exposed to a light source, the sun, whose photons are reflected by the dress in complex ways depending on the material’s microstructure. Some of these photons make their way into the eye, where they are absorbed by three distinct groups of photopigments. From their outputs the visual brain computes a label, called color, associated with the surface of the dress. It does so based

on a priori assumptions shaped by the visual environment the brain has been exposed to; different brains with different assumptions about the ambient brightness in the scene will come up with different answers. All of this happens quite automatically, effortlessly and unconsciously.<sup>2</sup>

That what we take to be “reality” can differ strikingly from person to person is not widely appreciated. In the context of #TheDress, it makes for a fun icebreaker at a party; yet the malleable nature of reality also has less amusing and darker implications for how we think about ourselves and others in the social and political arenas.

### **Billions of Bespoke Realities**

*Homo sapiens* as a species is defined by its genome. Unlike a digital document that can be flawlessly copied, again and again, without a single character missing or out of place, genomes are subject to random mutations and copying errors as the information is passed from parents to offspring. This comes on top of the genetic lottery as we inherit about half of our dad’s genes and half of our mom’s, but which of our dad’s genes and which of our mom’s genes we inherit is random, like a deck of shuffled cards. We all carry variants of nature’s instruction manual for who we can become—differing in roughly one out of every thousand nucleotides. These “copying errors” are not a bug but a feature, as offspring with diverse genomes are the raw material shaped by the forces of evolution by natural selection. Some may be better adapted to a particular ecological niche, while most are either neutral or do worse. The former will have a higher chance, and the latter a lower one, of surviving and therefore passing on their genome to the next generation.

These genetic differences are superimposed and amplified by the unique conditions we grow up in. Dramatic events—say, a period of malnutrition early in life or even in the previous generation—act directly on genomes through something called *epigenetics*. Abuse and punishment leave their traces in memory by sculpting the thicket of synapses connecting and festooning nerve cells. The brain is like a palimpsest; traumatic memories can be overwritten and effaced but are never truly forgotten. We are shaped by both “nature” and “nurture.” Powerful learning algorithms let us adapt to a vast range of physical, social, and linguistic environments, to mature into

competent adults.

Fundamental to this is *neuroplasticity*—a faddish term for the observation that animals and people can acquire skills, learn a new language, lay down memories, and adapt to the loss of a limb. This malleability of the nervous system is most pronounced in babies and children but persists throughout life, all the way into healthy aging. Like a potter forming pliable clay into a piece of ceramic, your biography shapes you into a distinct person.

Neuroplasticity manifests itself by appropriate changes in the underlying architecture of the central nervous system. In humans, unlike in mice, few new neurons are formed after birth.<sup>3</sup> Rather than adding new cells, the brain continuously adjusts its wiring by up- or down-regulating the heft or synaptic “weight” by which any one neuron extends its influences over the neurons it is connected to. We learn to bike or play the piano; pick up the syntax and semantics of Mandarin or English; absorb cultural attitudes, such as collectivism or individualism; and remember salient events. Indeed, updating synaptic weights is the way deep neural networks, such as the large language models that underpin ChatGPT and its relatives, learn.

The immature brain, like a sponge, effortlessly absorbs spoken and unspoken rules and biases about family members, friends, the privileged in-group, and the less fortunate out-groups living on the “wrong” side of the track. The young, highly malleable brain also picks up, consciously and unconsciously, cues about sexism, racism, and other forms of discrimination. These *priors*, to adopt the language of Bayesian reasoning, are shaped by personal experiences, both beneficial and traumatic, that reach all the way back to early childhood and constrain the way the child, and eventually the mature adult, responds to others based on their looks, actions, speech patterns, and other cultural signifiers. Collectively, these priors determine our view of physical and social reality.

Neuroplasticity extends into healthy aging but declines as we become more rigid, our brains more “crystalline”—as we grow more set in our ways: “You can’t teach an old dog new tricks.” This explains why we soak up a new language by osmosis, without even trying, as a youngster but struggle as adults. Neuroplasticity is blunted by chronic stress or by depression and other mental disorders<sup>4</sup> but can increase following psychedelic experiences.

Brain scientists frequently refer to a mythical creature, the neurotypical. This individual of unspecified sex stands in for the typical member of *Homo sapiens*, born with a full complement of eyes, ears, and other sense organs and a “normal” brain. The last chapter described the experiences of such a neurotypical person. However, in contrast to the “standard kilogram”—a cylinder of platinum-iridium hidden away in a vacuum chamber in Paris that once defined a mass of one kilogram—there is no “standard” human brain. Instead, there is vast genetic and developmental diversity among the eight billion people living on planet Earth, reflected in the astonishing diversity of their brains and their ways of experiencing the world.

Take color: while most of us are trichromats, endowed with three pigments in our cone photoreceptors, some women express four distinct photopigments, experiencing subtle hues forever unavailable to the rest of us.<sup>5</sup> One out of fourteen men are dichromats, commonly, but inaccurately, labeled color-blind. They experience fewer colors than normal, unable to recognize sunburnt skin, distinguish red from yellow traffic lights, or find ripe fruits. Much less common are monochromats, who see shades of grey but no color.

Renowned neurologist Oliver Sacks wrote *The Island of the Colorblind* about Pingelap, a tiny atoll near Micronesia, in which about 5 percent of the population have no cone photoreceptors at all, a rare condition known as achromatopsia. Instead, they must make do with the rod photoreceptors used for night vision, seeing only variations in brightness. In 1775 the atoll was devastated by a typhoon, reducing the population to a bottleneck of about twenty people, one of whom presumably carried the genetic defect for achromatopsia. The achromats on Pingelap fish and swim at night or dusk; in bright light they tend to blink incessantly.

Knut Nordby, a Norwegian vision researcher and achromat himself, noted in his diary, “When I started school, the other children could see something that I could not see. They called it color. They referred to things by names which had no meaning for me. I would not admit to this and tried to use these names, and sometimes I made very strange mistakes.”

Explaining color to achromats is impossible, for colors are more than a semantic label attached to surfaces. It feels-like-something to see the colors of a fluttering flag or the setting sun. Colors, like any other experience,

have what is called a *quale* (plural *qualia*), a unique feeling that makes seeing orange quite different from seeing purple and radically different from smelling garlic or touching a wet towel.

Many people lack binocular depth perception, while others see visual snow superimposed on everything. Some are blind to faces, confusing their spouse in a crowd with a stranger, while super-recognizers never forget a face, even that of a person they may have only seen once, years earlier.<sup>6</sup> The power of imagery runs strong in some (*hyperphantasia*) and is puny in others who can't visualize anything (*aphantasia*).<sup>7</sup> Some primarily think holistically in pictures, others in more abstract patterns, and some in linear strings of words or symbols.

Some always feel cold, while others walk around wearing shorts in winter. Some experience a dish as insipid, while the same food tastes too spicy to others. Some don't have umami taste receptors; others are born with perfect pitch or with the uncanny ability to distinguish a Parkinson's patient from a healthy person by the smell of their oily skin.<sup>8</sup> A few are born without the capacity to feel pain, which proves to be a curse rather than a blessing.<sup>9</sup> Some possess an unusual inner voice, have no inner voice at all, or hear hallucinatory, intruding ones. Some have a highly developed sense of empathy and compassion with the suffering of all beings,<sup>10</sup> while others are psychopaths, deriving pleasure from inflicting pain. Some are optimists, while others are perennial pessimists. Then there are the multiplicities of sexual identities and genders that can profoundly affect how we perceive and judge ourselves and others.

Like an athlete born with exceptional oxygen-binding capability or muscular strength, we may find that our idiosyncratic complement of senses influences our personality, the profession we chose to pursue, and the course of our life. For example, from childhood on, Oliver Sacks was unable to identify people by their faces. To avoid embarrassments, he became withdrawn and socially inept. I would meet Sacks at his home, as he then knew for certain whom he was speaking to. However, his shyness was a product of his face blindness, not a lack of care for people.<sup>11</sup>

What other differences in the ways reality appears lurk out there? The Perception Census, a citizen-science project in the United Kingdom led by cognitive neuroscientist Anil Seth and philosopher Fiona Macpherson, seeks to map the heterogeneity of ways of seeing via interactive tasks



everyone can do on their own computer.<sup>12</sup>

Do those of us with an enhanced sensorium, such as those with exceptional color discrimination, work as artists or in the fashion industry, while those with enhanced taste and smell become chefs and sommeliers?

We each live in our own Perception Box, to adopt an evocative metaphor coined by the writer, creator, and visionary Elizabeth R. Koch, a box whose walls are invisible and unbreakable, as we can only experience what our neural circuitry allows us to experience.<sup>13</sup> These walls become the filter through which we interpret everyone and everything. This ineluctable fact is true of all sentient beings, each adapted to its ecological niche, and, therefore, with its own Perception Box.

Adverse childhood experiences involving physical, emotional, or sexual abuse, maltreatment, and neglect cast a long shadow and influence the extent to which the grown adult feels at home in the world. The more adverse events in childhood, the higher the chances of obesity, generalized anxiety, depression, alcoholism, drug addiction, propensity for violence, and so on. All of this expresses itself in the Perception Box and how wide and expanded or how tight and contracted it is.

Like a young child who closes their eyes and assumes that you can't see them anymore, on the infantile belief that what is true for them is also true for you, we take reality as given and implicitly assume that everyone experiences the same, when in fact few do.

At current count, there are over eight billion distinct human realities on this planet, differing in subtle or often not so subtle ways. We know that people have widely different physical abilities because we can directly observe these differences: some hold a strenuous yoga pose or run a marathon while others puff when they walk up a staircase; some are ripped and robust, while others are skinny and delicate. Unfortunately, the fact that everyone experiences the world differently is not readily discernible. We grow up without realizing that we perceive things differently from others. Why would we? For we have never experienced the world in any other way. Only when reading a news story, watching a documentary, or closely querying friends about some "condition" does the penny drop: "Ah, not everyone experiences the world the way I do."



## Perception Is a Construction of a Description

That we each make such implicit assumptions is congruent with the idea that the brain constructs “reality” based on untold priors about the world. According to Kant, space and time are not inherent properties of the external world but rather fundamental concepts through which our mind perceives and understands the world. He proposed that space and time are a priori intuitions, meaning they are part of the inherent structure of mind, shaping our experience of reality. About a century later, the physicist and physiologist Hermann von Helmholtz argued that perception is a process of unconscious inference. After passage of yet another century, today’s theoreticians view this process, now called *predictive coding* or *predictive processing*, as a probabilistic Bayesian inference in search of the best, meaning the most likely, explanation compatible with all data available to the organism.<sup>14</sup>

Consider a startling visual illusion known as the “Lilac chaser,” which you can find on the web.<sup>15</sup> Twelve blurry pink disks are arranged in a circle, like the numbers on a clock, against a grey background. One of the disks briefly blinks off and on again, before the adjacent disk disappears and reappears, then the next one, and so on. This missing disk, or “hole,” travels continuously around the circle. Yet, when you steadily fixate on the cross at the center of the circle, you see only a single greenish disk, moving along the circle, while the eleven stationary pink disks are gone! Remarkable, you see what’s not there, while not seeing what is there!

The late vision scientist David Marr expressed it succinctly: “Perception is the construction of a description.”<sup>16</sup> This includes not only visual and other sensory percepts but also interoceptive percepts, fears, and other feelings. All these experiences are functionally and lawfully related to past and present events. Consider a veteran with a missing limb who suffers from unrelenting phantom limb pain in the absence of any peripheral tissue pathology. Here, contra naïve realism, which assumes that the external world maps directly, one-to-one, onto the internally perceived one, there is no direct stimulus, but there is an undeniable bad experience.

Collectively, as members of a species, we evolve and develop in similar circumstances and therefore share many aspects of reality with each other (as otherwise we couldn’t even cross a busy street or experience the Lilac

chaser). Yet, ultimately, each nervous system constructs its own description of the world.

### Mind over Matter

Each of us is stuck inside our own reality with shatterproof walls; yet the situation is not hopeless. We can achieve insight into our limitations by reading books, watching movies about diverse experiences, speaking to a therapist, listening to our friends, and educating ourselves about our predicament. Furthermore, we can expand the invisible walls that constitute our Perception Box by interventions and transformative experiences. Reality is malleable. Even if we start out with identical brains—say, as identical twins—circumstances and our choices about what to focus on, what to honor, and what to neglect influence our experience of the world.

Consider the much-maligned placebo effect. A placebo is a harmless medicine or procedure, ranging from swallowing a sugar pill to sham or pretend surgery, without any explicit benefit except for pleasing (the original meaning of the word *placebo*) and calming the patient. Yet these innocuous interventions can have benefits difficult to explain using a conventional mechanistic framework of molecules binding to receptors. The *placebo effect* measures the strength of these beneficial effects. Good doctors have always relied on the healing power of the placebo.

As we learned during the Covid-19 pandemic, it is de rigueur to control for the placebo effect in any double-blind clinical trial—at random, half the recruited volunteers are given a potential vaccine, while the other half are injected with an inactive substance. Neither the medical personnel nor the subjects have any idea who is getting what (that’s what makes it a double-blind trial). The efficacy of the vaccine is measured by how much additional protection it affords above and beyond that seen in the population that received the placebo.

Modern medicine came of age in the late nineteenth and early twentieth centuries, when infectious diseases were rampant, disabling and killing millions: yellow fever, cholera, typhus, tuberculosis, syphilis, and the Spanish flu. This gave rise to the quest for “magic bullets,” a notion introduced by chemist and Nobel laureate Paul Ehrlich, who discovered the first effective compound against syphilis. Antibiotics, insulin, and mRNA

vaccines are magic bullets, homing in on a specific mechanistic cause of a disease in an effective manner. This is not, however, true for the pharmacopeia arrayed against chronic conditions that plague us today: obesity, depression and anxiety disorders, chronic pain, Alzheimer's, and so on. These medicines have minimal specificity and effectiveness.<sup>17</sup>

This is particularly true of mental diseases that are on the rise everywhere in the modern world. People intuitively understand that a physical injury to the brain—a blow to the head, a stroke, a tumor, or a bullet—leaves deficits in its wake, such as loss of speech, defects in memory, and debilitating headaches. Yet mental illness has no simple causes; there is no helpful narrative. Instead, there is pervasive stigmatizing and victim blaming.<sup>18</sup>

Take depression. The accepted treatment is a daily pill of a selective serotonin reuptake inhibitor (SSRI), such as Prozac, taken by an astounding 10 percent of adults in the industrialized West. Yet the evidence from large trials and meta-analyses of groups of trials conclusively shows that SSRIs are only marginally better than placebo. That is, there is a small and persistent benefit when comparing the treated group as a whole to the placebo group; yet it is not at all clear whether this difference is clinically significant at the level of any one patient. Because these drugs have a range of side effects, such as sexual dysfunction and emotional blunting, and consign patients to lifelong dependency on them, this is bad news.<sup>19</sup>

But this is also good news: four-fifths of the drug response is duplicated by the placebo. Patients' expectation that they are receiving an effective drug, prescribed by a highly trained expert, validated by research, and taken by so many others makes them hopeful and lifts their despondency. Their belief has a significantly bigger influence on their symptoms than the actual SSRI molecule!<sup>20</sup> This is remarkable, for it implies that the patient's conscious state, their belief in the drug's effectiveness, can influence their body. More pills work better than fewer, injections are better than pills, and having a person wearing a white coat with a badge reading "Dr. So-and-So" perform the procedure is best.

This highlights the crucial role that expectation plays in shaping the way we learn about the world, about cue and reward. You train your dog to "sit and give paw" by rewarding him with a treat when he does so. He now expects such a reward. You pull an all-nighter to fill out that spreadsheet

and expect words of praise from your boss the next day. If she does compliment you on your work ethic, a circuit deep within your brain rewards you with a dopamine-induced jolt of pleasure, making you more likely to repeat that all-nighter.

Pain is particularly susceptible to the power of suggestion. If you previously obtained pain relief by taking an aspirin, you expect the same benefit when someone hands you a similar-looking pill, even though it may be a placebo. This form of analgesia can be blocked by the opioid antagonist naloxone, implying that your belief recruits your body's own opioid-like substances called endorphins. This is in line with what my mother told me about when she assisted her father, a surgeon, in operations in the hospital's bunker during the Allied bombing campaign in World War II Berlin: when they ran out of morphine, the injured received, unbeknownst to them, a harmless saline injection that nevertheless provided relief.

Placebos don't just work for pain or depression. Belief helps improve motor symptoms in Parkinson's disease and shapes immune response. The placebo effect is everywhere.<sup>21</sup>

Rather than being a troublesome artifact, an annoying nuisance that researchers must deal with, the placebo effect is a powerful mechanism for self-regulation. Think of it as a manifestation of belief: the more the patient believes in a procedure or manipulation, the more likely it is to help. The placebo effect powers practices of alternative medicine and cultural rituals, such as acupuncture or faith healing, that provide therapeutic benefits.

There is no light without shadow. Positive news coverage and stories of miracle cures by a new drug will enhance that drug's power (a challenge for contemporary psychedelic clinical trials), while adverse publicity of serious side effects reduces its effectiveness. Given the legacy of racism within medicine, African Americans are more likely to distrust medical interventions; these will therefore be less effective, a form of systemic inequality. Negative expectations lead to worse outcomes; your belief that something will hurt is more likely to cause you hurt. This is the *nocebo response*.<sup>22</sup>

The placebo and nocebo responses are ill-understood but powerful biopsychosociological phenomena channeling the effects of expectation into the body. Your belief can reach all the way down into your heart, gut, and

other organs believed to be inaccessible to the mind. Just knowing that you are part of a clinical study raises expectations, a sort of Heisenbergian “observer” effect. Furthermore, modern research ethics is quite strict about requiring the disclosure of any manipulations, including giving people a fake pill. That is, you need to tell participants in a clinical trial that they may receive a fake, which makes the placebo effect challenging to investigate in the wild.

No less formidable to study and treat are the mind-body interactions known as psychosomatic or psychogenic disorders.<sup>23</sup> These are shape-shifting debilitating conditions, with heterogeneous and fluctuating symptoms, no specific causes, no effective treatments, no clear organic pathology, and no objective blood marker or brain tests. These include psychogenic spells, which look completely authentic but have none of the electroencephalogram (EEG) hallmarks of genuine seizures; it is extremely difficult to detect psychogenic spells without an EEG, which the patient will refuse. Other psychogenic disorders include Havana syndrome, resignation syndrome, chronic fatigue syndrome, and fibromyalgia (historically known as neurasthenia).<sup>24</sup>

Indubitably, patients suffer; yet the causes of their suffering and the extent to which their conscious or unconscious mind, responding to stress and anxiety, influences their symptoms remain controversial. Patients react with anger at any suggestion that their problems are psychosomatic in nature, for that implies that it is “all in their head” or that they choose to be sick for a nefarious purpose. Instead, they cast about for a physical cause: nerve gas, sonic weapons, vaccination, electric power lines, and so on. It is far easier to believe that those debilitating aches are caused by some unidentified agent than, say, that they are the body’s learned reaction to anxiety, leaving a hypervigilant nervous system in its wake.

The vehement rejection of psychosomatic explanations reflects the widespread acceptance of a brain-is-hardware while the mind-is-software view: if neurologists or psychiatrists can’t find anything wrong with the hardware, well, then it must be the software that is malfunctioning and must be reprogrammed; in plain speak, you are crazy. In reality, the learned interaction between the brain, the mind, and the physical and social environments is vastly more complex than captured by this sort of naïve binary distinction. What is true is that in sickness and in health, conscious

attitude and narrative play a major role. It matters whether you think of yourself as a helpless victim of unknown forces over which you exert no control or as an autonomous agent shaping your destiny in the face of adversity. Some see the proverbial glass as half empty; some see it as half full; some don't see a glass; and some believe it doesn't matter, since the glass can always be refilled.

When I permanently moved from sunny Southern California to the rainy Pacific Northwest, I resolved to never, ever complain about the gloom; I adopted the maxim "There is no bad weather—there is only bad clothing and bad attitude." This has made all the difference, and now, more than a decade later, I can't imagine living anywhere else.

The importance of maintaining an active attitude toward the authorship of your narrative, owning your life, is the prime message of *Man's Search for Meaning* by Jewish psychiatrist Viktor Frankl, imprisoned for three years in Nazi concentration camps in Germany. According to Frankl, a cold, starved, and brutalized inmate can still freely choose how to interpret their terrible circumstance, providing meaning to their suffering. This fighting spirit, what the Greek Stoic philosopher Epictetus called the "last freedom," is what enabled a few to survive this ordeal in contrast to the many who resigned, became apathetic, and "smoked their last cigarette."

The "absurd man" in *The Myth of Sisyphus* by Existentialist philosopher Albert Camus, written, by no coincidence, a few years earlier than Frankl's book, arrives at the same conclusion:

All Sisyphus' silent joy is contained therein. His fate belongs to him. His rock is his thing.... Sisyphus teaches the higher fidelity that negates the gods and raises rocks. He too concludes that all is well. This universe henceforth without a master seems to him neither sterile nor futile. Each atom of that stone, each mineral flake of that night-filled mountain, in itself forms a world. The struggle itself toward the heights is enough to fill a man's heart. One must imagine Sisyphus happy.

This is the power of conscious belief to effect change. Rather than feeling exhaustion in the face of his fate, Sisyphus reframes his situation and

proclaims his autonomy. Your faith may not move mountains, but it can transform your experienced reality. It doesn't take a lot—the word *nowhere*, implying disorientation, can be reframed by cleaving it into two, *now here*, the radical opposite of delocalization!

## CHAPTER FOUR

---

### consciousness and the physical

Having surveyed the vast universe of conscious experiences, including mysterious realms little explored by science, let us turn to how philosophy has dealt with the existence of the subjective. How does consciousness fit into the grand scheme of things? How does the mental relate to the physical world “of shoes, and ships, and sealing wax, of cabbages and kings,” as the Walrus expresses it to the Carpenter in Lewis Carroll’s poem?

Vast forests have been cut down and turned into magisterial and heavily footnoted books by learned authors concerned with the mind. Don’t worry though. I can’t do justice to the collective outpouring of scholars over the past twenty-four hundred years of recorded Western thought, so I won’t try. Instead, I will give a thumbnail sketch of the intellectual wanderings I undertook over the past half century, from Catholic altar boy, confident of possessing a soul that will outlive my death; to physics-trained brain scientist professor, searching for the footprints of consciousness with a belief that all life is ensouled; to an executive organizing hundreds of neuroscientists; to a scholar of unusual conscious states, with a belief that my subjectivity is the only slice of eternity I am entitled to.

I learned that philosophers ask hard questions, uncover unspoken assumptions and unexpected consequences, point out discrepancies, and provide a range of possible answers. Reading philosophy is essential for any serious thinking. But to a scientist used to the notion of progress, the slow and groping ascent toward an understanding of the facts of the matter, and the implicit conviction that we can explain, predict, and manipulate nature better today than yesterday, philosophy can also be deeply frustrating.<sup>1</sup>

Consider the history of philosophy of mind: it illustrates the dizzying varieties of ways scholars throughout the ages have tried to fit the body, the



mind, and the soul into a single, consistent, and overarching explanatory framework. However, there has been no convergence to any broadly acceptable answers and little common ground across different philosophies, except for a decisive turn away from supernatural thinking. Highly polished arguments and counterarguments are exchanged in an eristic and never-ending jousting style that ends in drawn-out sophisticated disagreements but no resolution. There is no broadly accepted consensus that the field is slowly but surely converging toward a true state of affairs. Purely logical, mathematical, or linguistic arguments are, in the absence of empirical verification or falsification, inadequate to break this impasse. For a field of intellectual inquiry posing these questions since antiquity, on the order of a hundred generations of scholars, this is not encouraging.

Fortunately, humanity isn't condemned to ramble forever in an epistemological fog of uncertainty and never-knowing, for we now have science at our back—and that has made all the difference.

### [What Is the Mental?](#)

Hunger is an experience every living creature is familiar with. It ranges in intensity from weak cravings to the hunger pangs of malnourishment or starvation. But what sort of stuff or thing is this feeling? Emanating from an empty stomach, this aversive, unpleasant sensation has no spatial extension—it is neither tall nor wide, big nor small. Indeed, hunger doesn't have any of the usual attributes of the physical—it is neither heavy nor light, stationary nor moving. Hunger is not a form of energy as it doesn't have the attributes associated with energy—that is, an ability to perform work. It doesn't radiate in any wavelength; nor does it generate heat. It obeys no conservation law of the sort that rules the physical—conservation of mass or energy. There isn't a finite amount of hunger you can experience. For as long as you go without food, you feel hungry—in fact, increasingly so.

Because hunger is an experience, you have direct, firsthand acquaintance of this feeling. It does not need to be mediated by something else for you to experience it. You know that you are hungry. You don't have the same certainty about having high blood pressure or being in the early stages of a viral infection.<sup>2</sup> Such knowledge must be inferred by other means, such as a blood pressure cuff or a PCR test.

Notably, this direct and intimate knowledge of the mental is private—only you are aware of your hunger. Others infer that you are hungry—they know you haven't eaten in a while; they listen to you obsessively speak about food; they see you rub your belly and so on—but they don't experience it in the way you do. After all, you could be an actor pretending to be hungry. Conversely, you could be hungry but put on a poker face. The private or first-person attributes of the mental are radically different from the public or third-person attributes of the physical, such as elementary particles, stars, viruses, and brains. Their properties, such as their location, speed, mass, and makeup, are accessible to anyone with the right instruments.

What is true for hunger is true for all experiences. All are defined by their private attributes to which you, and only you, have immediate and direct access.

### Cartesian Dualism

The most intuitive way of thinking about the physical and the mental is that they belong to two fundamentally distinct realms of reality. This is what most of humanity believes. This is what my younger self believed, growing up in a devout Catholic family, attending weekly mass and mumbling prayers in Latin. I fully absorbed the church's teaching of a mortal body housing an immaterial and immortal soul that will be resurrected in the fullness of time, in the *Eschaton*. This is dualism at its finest.

The concept of the spirit or soul—*psyche* in Greek (from which psychology derives its name) and *anima* in Latin (the animating principle)—the repository of a person's true self, was taught by Plato and subsequently transmitted through Augustine of Hippo, who lived during the sunset of the Roman Empire in the fourth and fifth centuries, to Christendom. For Plato, the soul has various faculties and abilities, such as sensing, remembering, reasoning, and deciding. A belief in a soul is a core tenet of the world's monotheistic religions. It persists implicitly in people who think of themselves as nonreligious. The soul is part of our myths, stories, movies, political speeches, and everyday folk psychology.

Consciousness naturally fits into this dual view of reality. It was famously articulated by the reclusive seventeenth-century philosopher,

mathematician, and physicist René Descartes. Although Descartes was born, raised, and educated in France and wrote primarily in French, he lived for much of his adult life in the United Provinces of the Netherlands, where he could freely write on topics considered dangerous in Catholic France.<sup>3</sup>

Descartes's name is associated with the dominant strain of dualism. He was a towering intellectual who helped birth the Age of Enlightenment. At this point, I need to disclose that I attended and graduated from a French high school named in his honor, Lycée Descartes, in Rabat, Morocco. Descartes reified the difference between the physical and the mental by postulating that they make up two fundamentally different magisteria, made from two kinds of substances: *res extensa* (literally, “extended thing”), or physical stuff that has length, width, and a particular location, here or over there; and *res cogitans* (literally, “thinking thing”), or mental stuff, without extension and nowhere located, but with the ability to sense, think, reason, and feel.<sup>4</sup>

Physical stuff is distinct from, and exists independently of, mental stuff, and vice versa. There are stars, billiard balls, and bodies on the one hand,<sup>5</sup> and there are minds on the other. The brain is nothing but matter interacting in complex ways as dictated by the laws of mechanics and chemistry. Informed by his dissections of animals, Descartes speculated that, like the flowing water that powered the moving statues of gods, satyrs, tritons, nymphs, and heroes in the fountains at the French court in Versailles, “animal spirits” flowed through the capillaries, cerebral cavities, and nervous tubules of brains, enabling movement.

But Descartes was at a loss to conceive of mechanisms for intelligence, reasoning, and language. In his day, nobody envisioned yet how the mindless application of meticulously detailed, step-by-step instructions—what today we call an algorithm—can be exploited to play chess, recognize faces, and speak. For this, Descartes appealed to mental stuff. In a living person, the mind is allied with a body. Because the mind is made from different stuff than the body, it exists independently of it and so survives after the body dies.

Descartes had a variety of arguments in favor of these propositions. Here is a flavor of these. In his skeptical mood, Descartes sought ultimate certainty, to find out what was indubitably, without a doubt, true. To this end, Descartes conceived of a “supremely powerful malicious deceiver”

who could fool him about the existence of the world, his body, and everything he saw or felt. He concluded that he couldn't be sure of the shape or size of his body or whether he even had one. Yet he could not doubt that he was experiencing *something*, and therefore he existed. He expressed this in the memorable dictum *Cogito, ergo sum*, or "I think, therefore I am," the most famous deduction in Western thought. Because Descartes knows that his conscious mind exists, but he does not have the same certainty for his body, he concludes that his mind is not the same as his body.<sup>6</sup>

A contemporary version of the malicious deceiver are the sentient machines in the *Matrix* movies. When Neo takes the red pill offered to him by Morpheus, he wakes up to discover that he lives in a gigantic stack of pods and that the life he thought he lived was nothing but a computer simulation. While this prompts him to question reality, he never doubts his own existence; he is, after all, experiencing something. That's pure Descartes.

Half a century after Descartes, German polymath Gottfried Wilhelm Leibniz formulated a compelling argument that the mental does not arise from matter. Leibniz coinvented calculus and binary numbers and built the first digital calculator. His antimaterialist argument is known as the "mill thought experiment":

Moreover, we must confess that the perception, and what depends on it, is inexplicable in terms of mechanical reasons, that is, through shapes and proportions. If we imagine that there is a machine whose structure makes it think, sense, and have perceptions, we could conceive it enlarged, keeping the same proportions, so that we could enter into it, as one enters into a mill. Assuming that, when inspecting its interior, we will only find parts that push one another, and we will never find anything to explain a perception.<sup>7</sup>

Entering the brain in this manner with an electron microscope, today's gold standard in high-resolution imaging, makes neuronal membranes, synapses, and other cellular organelles visible. Peering even deeper, with an atomic force microscope, individual macromolecules come into focus. But

never pain, pleasure, or ennui. Ultimately, the brain is just an amazingly complex set of interlocking dumb mechanisms causally interacting with each other. How can conscious feelings arise from mere mechanisms? They can't, in this view, as the mental is radically different from the physical, and neither can emerge from the other.

Descartes's doctrine liberated natural philosophers, budding proto-scientists in the Age of Enlightenment, to focus on mechanistic studies of the body, birthing the scientific revolution, while leaving the dangerous stuff, the soul, to the theologians. Strong echoes of this dualism survive in contemporary debates around the question of sentient computers, as *res extensa* and *res cogitans* map naturally onto hardware and software. This is a seductive and widespread but also pernicious metaphor I'll return to below.

One of the main objections to Cartesian dualism is the *causal interaction problem*: How does ethereal, thinking stuff impose its will onto concrete, physical stuff? How does the mind direct matter? This concern was articulated in a correspondence between Descartes and the young Princess Elizabeth of Bohemia, exiled in Holland. She posed a straightforward question: "I find from your letter that the senses show me that the soul moves the body, but as for how it does so, the senses tell me nothing about that, any more than the intellect and the imagination do. This leads me to think that the soul has properties that we don't know—which might overturn your doctrine, of which I was persuaded by your excellent arguments in the Meditations, that the soul is not extended."<sup>8</sup>

For something to exert a causal influence on something physical, the former must impart the latter with momentum or with energy. But what momentum and what energy does the mental have? Where do these come from? Their own, mental energy reservoir? Any such transfer must be accounted for when balancing the books in terms of total energy usage. There is no free lunch and no evidence of any such mental energy messing around in the waters of the brain.<sup>9</sup>

Mental causation is the Achilles' heel of Cartesian dualism. If the mental cannot interact with the physical without breaching well-established physical laws, dualism could still hold but would be shorn of its attraction, for the mental would be causally impotent to affect the brain. You think that your desire to raise your hand caused your hand to rise, but that would be

folly, an illusion. The action of the arm is fully and completely accounted for by material events in your brain. Your conscious experience of agency is nothing but a ghost in the machine. Your feelings make no difference to the world. The mind is along for the ride, helpless to do anything but see, hear, and feel. The real actor is the brain. Consciousness serves no purpose; it is an epiphenomenon.

Descartes's assumption that *res cogitans* is immune from decay safeguards the notion of the soul and its eventual union with God in the hereafter. However, a soul leaves difficulties in its slipstream. How can a nonmaterial soul record and store personal memories and the traits that makes you "you," in particular once your bodily substrate has died and decomposed? Where was your soul before you were born? Where will it be when you're dead? In some sort of hyperspace, an antechamber to heaven, where it will eventually be reunited with your body? And with which body? The decrepit one that wasted away in a hospital or the one in the prime of life? The belief in an immortal soul does not withstand scientific scrutiny and falls apart like a house of cards.

Similar problems arise with reincarnation. At a meeting I had with His Holiness the Dalai Lama and Tibetan monks in southern India, the discussion turned to the Buddhist belief in reincarnation and the question of where the mind, with its traits and memories, resides between consecutive incarnations. I raised, one after the other, four fingers on my hand while slowly enunciating the four words that make up what I call the *neuroscientist's dictum*: "No brain, never mind."<sup>10</sup> Consciousness cannot exist in a pure vacuum. It requires a substrate, such as brain cells, electronic circuits, or maybe something more exotic, such as entangled quantum states. But without a substrate, there can be no experience.<sup>11</sup>

### [Everything Is Physical](#)

The dominant attitude among living scholars of consciousness is *physicalism*,<sup>12</sup> a kissing cousin of materialism. It has a venerable pedigree, going back to the philosopher Democritus, a contemporary of Socrates. Little survives of Democritus's work except a few striking statements, such as "By convention there is sweetness, by convention bitterness, by convention color, in reality only atoms and the void."

Physicalism is the metaphysical thesis that at rock bottom everything is reducible to quantities that can be described in an observer-independent manner—that is, both you and I can observe balls rolling on a billiard table, measure their size and velocity, and come to an agreement. In principle, we can do the same for molecules, atoms, electrons, photons, and so on. All facts about the world, including your fierce love for your child or the strenuous feelings associated with a challenging climb, are facts of the physical world, mechanisms going through their motion as dictated by their causal powers.

Physicalism is Cartesian dualism stripped of thinking substance. One substance, rather than two, rules all. In that sense, physicalism is a simpler explanation than dualism. Why postulate *res cogitans* when everything can be explained already by *res extensa*?

To many scholars, physicalism seems the obvious and, indeed, the only stance to take in a postreligious era once God, the soul, and the afterlife have been cast into the waste bin of discarded ideas, together with Zeus, Wotan, the four humors of medicine, and a flat Earth.

But more precisely defining physicalism has been challenging. Which physical laws are meant? Today's physics, thought to be incomplete? A final theory of physics, if such a thing exists and is discovered in the fullness of time? But such a final theory might include the mental if properly construed! Physicalism is a promissory note, claiming that in due time natural laws, not supernatural ones, will explain the whole shebang, everything that is the case, including phenomenal consciousness.

This perspective, which I call *classical physicalism*, is in crisis mode, although that has gone largely unnoticed. It assumes that objects, such as my bicycle, have definite properties that do not depend on an observer interacting with them. You don't need to see my bike to know that it has a definite rest mass.<sup>13</sup> Since Heisenberg, we know that this is not the case for microscopic variables. Furthermore, object properties are thought to only depend on what is happening within a certain spatial neighborhood around my bike. This is called *locality*. But ultraprecise measurements over the last several decades have conclusively demonstrated that this is not the case for microscopic objects. Quantum entangled systems, say a pair of photons, can be arbitrarily far apart yet remain tightly coupled to each other in some mysterious way until one feature, say the polarization angle of one photon,



is measured, in which case, instantaneously, the angle of the other entangled photon also assumes a specific value—even if it is a million miles away. Until measured, the photons remain linked to each other in some strange way, as if by magic, no matter their distance. And what is true for a pair of entangled particles can also be true for millions of them.

That physical reality is observer dependent and nonlocal is established fact—so much so that the 2022 Nobel Prize, one pragmatic measure of what is considered “established beyond a reasonable doubt,” was awarded to the three quantum physicists, Alain Aspect, John Clauser, and Anton Zeilinger, whose experiments with entangled photons established nonlocality.<sup>14</sup> What this means for the relationship between the physical and the mental is acutely debated. Is consciousness a nonalgorithmic process—that is, something that cannot be described or expressed by a set of specific rules or instructions running on a computer? The physicist Roger Penrose forcefully articulated this view in a series of books, particularly *The Emperor’s New Mind*, intimately linking any conscious experience to the reduction, or the collapse, of the quantum mechanical wave function.<sup>15</sup>

The weight of the evidence supports the conventional view that quantum-mechanical effects wash out when dealing with large organisms at room temperatures, hot and wet by the standards of quantum mechanics. But whether this is a valid assumption for all aspects of brains, shaped by hundreds of millions of years of natural selection by evolution, is difficult to conclusively prove and remains an unresolved question. There is no question, though, that having access to quantum resources within or across cells would yield dramatic benefits in terms of efficiency and speed of learning compared to classic resources.<sup>16</sup>

Leaving these major challenges aside, a school of thought called reductive physicalism assumes that every mental state is fully reducible to the physical state of the underlying substrate—every subjective experience has a unique, associated brain state. Experiencing an orgasm is identical to a particular brain state, with some neurons firing and others being silent. There is nothing above and beyond these neural processes. End of story.

According to reductive physicalism, once scientists understand how every conscious state is fully and uniquely reducible to the firing of a vast assembly of neurons (or the vibration of microtubules, the collapse of the wave function, or whatever else), psychology can be replaced by



neuroscience and, ultimately, by physics. The mind is the brain. Phenomenal feelings do not make any causal difference to the underlying physics. There are no autonomous mental phenomena above the causal powers of the underlying brain state.

Indeed, once mental states are accounted for in terms of the physical, they become superfluous, as they don't add any explanatory power. Mental states are a shorthand for associated actual and potential behaviors. The painful feeling of an inflamed tooth is identical to groaning and wincing, rubbing that side of the mouth, avoiding eating on it, increased blood pressure and heart rate, the release of stress hormones, and so on. Once all of that is explained, the hard work is done. Feeling and experiences need never bother us again for they have been marginalized as illusionary or even outright eliminated. In these “deflationary” or “eliminative” views, phenomenal, subjective experience will eventually join the superego, the Oedipus complex, and other constructs as an antiquated way of thinking.<sup>17</sup> Daniel Dennett, the doyen of American philosophers, is quite explicit and outspoken: “The elusive subjective conscious experience—the redness of red, the painfulness of pain—that philosophers call qualia? Sheer illusion.”<sup>18</sup>

In a letter to Dennett following an aborted climbing trip in the Sierra Nevada due to an infected tooth, I expressed my incredulity that my excruciating pain was but an illusion, brought on by some sort of cognitive confusion. This denial of the lived reality of conscious experience is absurd, an extreme form of gaslighting. What would Parmenides, Plato, or Aristotle have thought of an age where such a disavowal is taken as supreme wisdom?<sup>19</sup>

Any -ism that fails to explain the qualitative, felt aspects of pain, despair, distress, despondency, suffering, trauma, loss, and melancholy does not account for the human condition. *Darkness Visible* by William Styron is a searing telling of the author's descent into the hell of severe clinical depression, full of lucid descriptions of bouts of existential meaninglessness and emptiness, “the grey drizzle of horror,” “the despair beyond despair,” the “torpid indifference” to life. It is the torment of these experiences that drives depressives to suicide (Styron himself contemplated it many times) and that needs to be explained rather than being denied. What is valid for depression is equally valid for other mental conditions; treating these by

focusing exclusively on objective behavioral criteria or physiological ones (biomarkers in the lingo) at the expense of patients' experience has limited effects on their well-being. Neuroscience knows enough to be dangerous but not enough to be helpful to patients. Psychiatry and philosophy must return experiences to "mental diseases."<sup>20</sup>

A further conceptual obstacle that stands in the way of unambiguously reducing a mental state to a physical state is the fact that any experience can be instantiated in many ways. Consider, for example, pain. All living creatures experience pain. Adults and children do, as do preterm infants and newborns whose immature brains are not yet fully wired up. If you share life with a dog or cat, you know that they too feel pain. Modern laboratory practices pay a lot of attention to eliminating the discomfort and pain in research animals. There is a vigorous debate in neurobiology concerning the extent to which nonmammalian vertebrates, such as fish, or invertebrates, such as insects, experience pain when they exhibit associated behaviors, such as a centralized orienting response and tending to their wounds.

The nervous systems of all these creatures are different, sometimes even radically so, implying that pain must be instantiated in innumerable distinct ways. And what is true of feeling pain is also true of being scared or seeing. This is the multiple realizability argument against a one-to-one mental-to-physical reduction. The program to map a mental state to a single, unique and quantitative state has hit yet another roadblock.

### [The Computational Mind](#)

In reaction to these challenges, the philosopher Hilary Putnam introduced *machine* or *computational functionalism*, the mythos of our Information Age.<sup>21</sup> Putnam argued that pain is not a set of behavioral dispositions but a functional state that detects actual or potential damage to the body and takes evasive action. Part of pain's job description is to generate a bad, aversive feeling, to signal a sense of urgency to do something "now." Pain is triggered by inputs, including temperature, pressure, joint angle, and other interoceptive sensors distributed throughout the body, and initiates defensive behaviors, such as withdrawing the limb, orienting toward the source of pain, vocalizing, and so on, up to overriding such actions because something of higher priority is going on—say, being chased by a wolf.

A functional state that supports such causal relationships will be associated with pain, no matter whether it is instantiated in the nervous system of an octopus, a dog, a human, or even an alien, as in Ridley Scott's eponymous movie. What matters for a mental state is the role it plays in the organism, including its causal relationship to the environment, sensory input, motor output, and other mental states. The physics of the mechanism, the stuff out of which the system is made and how it is wired together, is an implementation detail.

Putnam further suggested that such functional states can be implemented by a universal Turing machine. Digital computers and smartphones approximate Turing machines; nervous systems can be thought of as Turing machines in which deterministic, all-or-none state transitions are replaced by probabilistic ones. The details don't matter; only the abstract operations are relevant.<sup>22</sup>

Given that the functional state description is independent of the underlying material substrate, if the same causal relationships held in a Tesla as in a dog or a human, the Tesla would feel pain when damaged or tired when its batteries were getting drained and would seek the nearest service station.

In this worldview, you are a Turing machine made flesh. You are a robot unaware of its programming. The mind as software is the reigning doctrine in philosophy and computer science departments. It dominates the tech industry and is at the core of contemporary arguments that an advanced AI will be sentient. It is a mainstay of Hollywood movies: Rachael in the cult science fiction movie *Blade Runner*, Ava in the dark psychodrama *Ex Machina*, David in *Prometheus*, not to mention characters in *Westworld*, *Dark Mirror*, and countless other shows and movies.

Yet the reduction of consciousness to a type of function or a class of computations leaves me cold. Why would one type of computation—say, computing the distance between me and other objects in my immediate visual surroundings—be unconscious, while a different one—say, evaluating the state of my body—would go together with subjectivity? Where would these feelings come from? What are they? How do they emerge out of the physical? This calls for some new level of explanation, above and beyond the theory of computation.

Indeed, some philosophers argue for the existence of an unbridgeable

*explanatory gap* between the physical and the phenomenal. You might have seen a famous *New Yorker* cartoon by Sidney Harris in which two physicists work out a calculation on the left and right sides of a blackboard. In the middle of the board, one of them has written, “Then a miracle occurs...,” to which the other responds, “I think you should be more explicit here in step two.” This is the explanatory gap: Even if you accept that any experience depends on a substrate, such as a brain,<sup>23</sup> why should any physical state feel-like-something? A vacuum cleaner doesn’t feel anything, so why should a brain? What is it about certain states that singles them out as “feeling-like-something”? There is a non sequitur between the physical state and the experience. This implies that physicalism as currently conceived is, at the least, incomplete.

This brings me to zombies! Introduced by the philosopher Robert Kirk and refined and popularized by David Chalmers, zombies are imaginary creatures indistinguishable from you and me except that they have no feelings. They have no minds. Unlike their Hollywood incarnations, philosophical zombies don’t have a predilection for human flesh. To fool us and lull us into a sense of complacency, zombies speak about their feelings. But it’s all a deep fake.

In his PhD thesis, which turned into an unlikely bestseller, *The Conscious Mind*, Chalmers asks whether the existence of such zombies is compatible with physicalism. Can one conceive of a world, indistinguishable from ours in all ways, except without experience? Chalmers answers yes. None of the foundational equations of quantum mechanics or general relativity mention experience; nor does chemistry or molecular biology. No scientific or logical fact contradicts the possibility of zombies. Yet, in our world, experience is a given—a brute fact. Thus consciousness is an additional facet of reality, beyond physicalism.<sup>24</sup>

On one side of the explanatory gap is what Chalmers calls the Easy Problem. Elucidating the function of pain implies tracing, step by step, the causal chain of events between, say, spilling acid over your hand and your reaction, running to the water faucet to wash the caustic substance from the skin. Identifying such mechanistic chains is what scientists do. These may be immensely complicated, but the work is conceptually straightforward. On the far side of the explanatory gap is the Hard Problem, explaining the conscious experience of burning pain. Chalmers argues that it may be

impossible, Hard with a capital *H*, to cross the gap to reach subjectivity. The distinction between the Easy and the Hard Problems of consciousness has generated a huge secondary literature, including a play by Tom Stoppard. We shall see whether it will remain truly unfordable.<sup>25</sup>

### Experience Is Everywhere

Physicalism remains orthodoxy in philosophy and science, despite the profoundly counterintuitive but well-established nonlocality of quantum mechanics.<sup>26</sup> The eagle of physicalism, having successfully defeated its enemy, dualism, screams supremacy, voiced by Dennett, who crowed, “Dualism is not a serious view to contend with, but rather a cliff over which to push one’s opponents.”<sup>27</sup> This triumphalism, however, turned out to be premature and classical physicalism in decline. Much older ways of thinking about the mind are enjoying a remarkable renaissance.<sup>28</sup>

Philosophy of mind has been captured within the forcefield of two poles, the physical and the mental, associated with physicalism on the one hand and idealism on the other. The waxing and waning of the military, political, economic, and cultural dominance of the Anglo-Saxon world in the twentieth and twenty-first centuries are reflected in the rise and fall of analytic philosophy, with its single-minded (*sic*) focus on physicalism. While idealism, which holds that, ultimately, everything is a manifestation of mind, is beyond the pale within traditional analytic philosophy, it is foregrounded by philosopher and computer scientist Bernardo Kastrup, who articulates a full-throated defense of idealism as the most rational approach to understanding reality.<sup>29</sup>

Idealism is not the only resurgent school of thought; older and fainter voices are becoming audible again. Whereas dualism tries to reconcile the physical with the mental but is challenged to explain how they can be so tightly linked, other views hold that the physical and the mental are intimately related at the bedrock level of reality. This ancient teaching, known as panpsychism, assumes that anything physical is made from conscious parts, or forms part of a greater conscious whole. Ultimately, consciousness inheres to the constituent of matter (particles or their associated fields). Panpsychism has, somewhat to my surprise, reestablished a degree of popularity.<sup>30</sup>

A related view holds that the physical and the mental are different aspects of the same Ur-substance. Some of the brightest minds in the West took this position, including the founding figures of psychology, Gustav Fechner, Wilhelm Wundt, and William James, and the logician and philosopher Bertrand Russell.

Russell argues that physics has nothing to say about the intrinsic nature of matter; rather, it describes how bits and pieces of matter relate to each other, say, via Coulomb's law governing the force among electrically charged particles at rest. A physical description at any level of granularity is about causal interactions among organs, cells, organelles, molecules, atoms, nucleons, fields, and so on. Ultimately, the material world is a gigantic matrix of fundamental entities having causal relationships with each other. Physics does not speak about the inner nature of these entities. They remain beyond reach.

Russell contends that the mind *is* the intrinsic nature of the physical. Consciousness is the interior view of the causal structure of brains. It is the way they feel like. And so it is for everything else. Every physical mechanism, such as a brain, has both external, publicly available attributes, as well as internal, private attributes. Panpsychism's metaphysics is elegant as it solves the interaction problem and eliminates the need to explain how the mental "emerged" out of the physical once it became big and complex, like in human brains. There is no need to postulate emergence, as the mental is always allied with the physical, two sides of the same coin.

One implication of panpsychism is that the mental is ubiquitous: not only present in animals but also in plants, protozoa, and bacteria. But it doesn't stop there. According to panpsychism, the mental extends all the way down to the ultimate constituents of matter, although no one has claimed that it feels like a lot to be an elementary particle. Most people find this notion profoundly counterintuitive. Yet it is not inconsistent with the facts.

Panpsychism is bedeviled by a conceptual challenge, known as the *combination problem*. Where are the boundaries of the mental drawn? There is you, and there is I, but there is no merged mind that shares both of our experiences. How does panpsychism explain what keeps us separate? The United States of America does not feel-like-anything, even though there are more than three hundred million conscious American citizens.

Groups can engage in highly synchronized and well-rehearsed rituals, such as a military parade or a corps de ballet, but they don't give up their individual minds to a group mind. Unlike atoms that combine into molecules, organelles, organs, and organisms, experiences do not combine into larger experiences, into a hive mind. How do a hundred billion mindless neurons aggregate to make up a commodious human mind but then aggregate no further into a collective über-mind of all people? What delimits my mind from yours?

Besides claiming that everything has both extrinsic and intrinsic aspects, panpsychism is silent about these questions. Is the mind determined by the physical makeup of its substrate, its constitution, or by the way it is interconnected, its structure? Or perhaps by what it does, its function? Where are the boundaries of mind? Does the Milky Way galaxy, with its several hundred billion stars, have a mind more capacious than the human one constituted from a roughly similar number of cells? Unlikely, but panpsychism won't tell you why.<sup>31</sup>

But today's most exciting theory of consciousness will answer these and other questions, including the burning question of the day: the prospect of machine consciousness.

## CHAPTER FIVE

---

### what truly exists

It is impossible to come to a clear understanding of the nature of the mental without a proper understanding of existence. To do this, we must distinguish the absolute existence of conscious minds from the more derivative, lesser form of existence of objects, such as brains. Let me explain.

When I am in a truly dreamless sleep, there are no experiences; there is nothing; I do not exist for myself. My sleeping body is there, in bed, breathing in and out, for others to observe. But not for me. This is *relative*, or *extrinsic*, *existence*; existence for others. Stars, rocks, cars, and garbage cans exist in this derivative manner, only for others, not for themselves.

When I wake up, groggily groping to turn off the alarm on my phone, I come from nowhere into being. This mind hears a jarring sound and senses a supine body, without yet even being fully aware of where it is or what day it is. But at that point, the conscious mind exists already for itself, intrinsically. It doesn't have to experience anything exalted, mystical, or searing. Just *being* without much conscious content is entirely sufficient for *intrinsic existence*. Being as compared to nonbeing. This is *absolute existence*, the only existence worth having.

Consciousness is the launching pad for everything else—not its physical substrate, the brain. This starting point is what makes integrated information theory (IIT) so different from contemporary theories that start with the brain and then seek to squeeze the juice of consciousness from it using computational functionalism. IIT starts with consciousness, not with the brain.

### Only What Has Causal Power Exists

Formalizing the distinction between intrinsic and extrinsic existence in terms of *causal power* is at the heart of integrated information theory.



Causal power, the ability to be a source of change, is what lies beneath the theory's elaborate and, to many, daunting mathematical carapace. The theory arrives at some rather surprising conclusions: phenomenal experience is widespread, it can be quantified, digital computers as currently conceived only have a miniscule amount of it, and minds like ours have free will.

The theory and its mathematical underpinning constitute a singular intellectual edifice, architected by Giulio Tononi, a brilliant Italian psychiatrist and professor of neuroscience at the University of Wisconsin, Madison. Guided by Tononi's vision and aided over the years by numerous collaborators, including me, IIT has evolved into an empirically testable framework that explains many facts, predicts new phenomena, and can be extrapolated in unforeseen ways. Many are attracted by IIT's stark beauty, its ontology, which is spelled out in considerable detail. Indeed, there is computer code for computing  $\Phi$  and unfolding the causal structures of very simple model circuits.<sup>1</sup> Such a starting point is rare in science and makes the theory unique. However, its uncompromising stance on the centrality of subjectivity for existence, and some of the theory's implications, is also why it has attracted considerable controversy.

The theory's omphalos is consciousness and the way any experience feels like. Again, all other theories start in very different places: a particular type of event, such as the collapse of the quantum mechanical wave function; a particular neural signal or activity, such as high-frequency brain waves; a particular behavior, such as pushing a button when seeing a face; a particular function, such as writing information to a central buffer, a blackboard; a particular computation, such as a global broadcast. These theories then argue that this special event, activity, behavior, function, or computation is closely linked with, or constitutes, consciousness. It is this difference that sets IIT apart from all other theories of consciousness.<sup>2</sup>

Francis Crick and I proposed many years ago that electrical brain activity buzzing about forty times a second is an essential neural signature of consciousness. A principled problem that besets these ideas, including ours, is the question why any event, activity, behavior, function, or computation should give rise to consciousness. Your brain is a piece of furniture of the universe like any other, although more complex than most. Why should some—indeed, why should *any*—molecular or neuronal

commotion within the brain, and only the brain, be associated with subjective experiences of pain or pleasure, of seeing or hearing? No one believes that the oxidation of gasoline inside a combustion engine makes the car feel warm or that electrical charge flowing onto the gate of transistors is associated with tickling. So why should the buzz of some neurons inside your skull do the trick? These are two distinct domains: the material world of mechanisms going through their motions, on the one hand, and the mental world of experiences, on the other. That's Chalmers's Hard Problem, crossing the explanatory gap between objective physical events and subjectivity.

Integrated information theory does not need a miracle to convert the water of brain activity into the wine of consciousness, because it starts with the latter, with intrinsic existence.

The theory assumes realism, the metaphysical stance that people, dogs, tools, and atoms exist independent of my experience. The world exists even when I am not conscious. Some of the minutiae of the world may differ, pace quantum mechanics, when I fall into a deep sleep, but not its existence. To assume realism is, surprisingly enough, a radical act. Take the opposite belief that nothing exists outside my experience. The universe was born with me, and it will die with me. This is *solipsism*, the thesis that you, this book, and all other things are figments of my mind. It is impossible to refute on strictly logical grounds. I don't spend any time on this extreme form of narcissism as it doesn't explain anything about the world. It just turns back everything to my experience. Solipsism is for the ego-infatuated.

While we're on the topic of outlandish beliefs, let me dispense with another one: the *simulation hypothesis*, the argument that the reality we're living in is fake, a hyper-realistic computer simulation running in the next universe up. I can't refute with utter certainty that I'm not living in the Matrix, just as I can't refute, in the words of the ancient Taoist philosopher Chuang Tzu, that "I am really a butterfly dreaming that I am Chuang Tzu." Logical possibility does not imply, however, physical possibility, even though it is intellectually enjoyable to while away the hours in heated discussions. The simulation hypothesis shares much with angelology, the historical debate concerning hierarchies of angels and other supernatural beings among scholastic philosophers of the Middle Ages. Both are equally sterile in explaining this world.<sup>3</sup>

Of course, whether something can be rigorously mathematically proven is of less consequence than one may at first imagine. As the logicians of the last century, in particular the reclusive Kurt Gödel, formally demonstrated beyond doubt, every sufficiently complex axiomatic system contains statements that cannot be proven true or false. It cannot even be shown that such an axiomatic system is self-consistent—that it doesn't contain contradictory statements. To see what I mean, consider the syntactically perfectly valid phrase "This statement is false." If we accept it as a true statement and just read it out, it clearly states that it is false. But that implies that its opposite must be true, which contradicts what it explicitly states. This is a linguistic version of an ouroboros, the snake eating its own tail. If logic can't even deal with a paradox expressible in four words, it should be no surprise that it fails to deal with the complexity of the world.

Besides the reasonable assumption that there are persistent objects outside my experience, integrated information theory presupposes that things exist to the extent they have cause-effect power. If something—such as Santa Claus or the luminiferous aether, a space-filling substance postulated by classical physicists that would let light waves propagate through otherwise empty space—does not make a difference to anything and nothing can affect it, it has no causal power and can therefore be disregarded.<sup>4</sup>

Causal power is not some intangible notion but rather something quite concrete: the extent to which something can be the source of change—say, the fact that those three neurons over there being simultaneously on will cause that neuron over here to go off. It is the ability of the system's recent past to specify its present state (cause power) and the extent to which the present state specifies its immediate future state (effect power). Consider a hammer. I can pick it up and grasp it by its handle. I feel its hefty weight and use it to pound a nail into the wall. In this way, I satisfy myself that this hammer exists as much as anything does: something exists if it can be manipulated, if it can take a difference (cause power), and if this has an observable effect, if it can make a difference (effect power). Manipulations include not only hammering but also toggling switches or neurons on and off, injecting current with an electrode, probing with ultrasound, radio waves, laser light, and so on. Observations include looking with the naked eye or through telescopes, microscopes, magnetic scanners, and other

instruments.<sup>5</sup>

Physical objects—molecules, nerve cells, bodies, planets—have causal power by dint of their associated gravitational and electromagnetic fields,<sup>6</sup> attracting and repelling each other as the case may be. Something that cannot make a difference to anything or be influenced by anything is causally impotent. It can be disregarded from the point of view of existence.

You might object that nonphysical entities, such as the idea of “God” or the “United States of America,” can have causal power. This is indubitably true. However, the causal power of God is channeled through the mind of the devout who believe. It is this belief in a supreme being that causes them to be virtuous or to kill. Likewise, the United States derives its causal powers, such as passing laws, printing money, or starting wars, only to the extent that people consciously assume that it has such powers. Without such a collective belief, all the might of the United States shrivels away (how much causal power does the Roman Republic possess today?).

This is also true of the causal powers of small, greenish notes of paper. They have the remarkable ability to be traded for stuff, real estate, and labor. Their causal power depends on a planetwide conscious belief, the psychology of money, whose strength waxes during a banking crisis and wanes when inflation rears its ugly head.<sup>7</sup>

How does one go about measuring causal power? An operational measure of the power of money is the cost of a basket of standardized goods at any one time and place. How much does one dollar buy? This is effectively how the government measures inflation and makes cost-of-living adjustments. Integrated information theory likewise adopts an operational definition, calling it *cause-effect power*. It is quantified by manipulating all components of the system under investigation and observing the outcome.

Consider a neural circuit. It is fully described by listing all possible synaptic inputs to the circuit and all associated outputs. When this synaptic input is applied, the firing activity of those neurons changes in such-and-such manner. This can be done for individual neurons, turning them on and off and noting the effect, if any, on the overall circuit. Because neurons can be highly nonlinear, next all possible pairs of neurons should be turned on and off and the outcome determined. Then this should be repeated for all triplets and so on. Doing this exhaustively for billions of neurons is, of course, infeasible, but in principle this way of proceeding is quite

unambiguous and doesn't appeal to anything magical.

Due to inherent and unavoidable randomness in nature, toggling tiny synapses on and off does not always lead to the same, reproducible result each time. Thus, a more general approach is to consider conditional probabilities: if ten synaptic inputs are activated, the neuron turns on 75 percent of the time and remains off the other 25 percent. The outcome is tabulated in the system's *transition probability matrix*, a table that has as many rows and columns as the system has synaptic inputs and neuronal outputs. This matrix exhaustively describes what a system does. Indeed, the transition probability matrix defines what the system *is*. The same principle applies to the central processing unit of a computer, with its several billion transistors. Some gates over here are flipped, triggering changes in those transistors over there. Doing this exhaustively defines the computer's transition probability matrix.

Defining existence as causal power traces its origin to Plato<sup>8</sup> and is a near-universal but rarely acknowledged principle for what science means when it stipulates the existence of something, such as a Higgs boson, a virus, or a black hole. Everything that exists in a fundamental sense has causal power, the power to take and to make a difference. All of physics can be expressed in this operational manner, using conditional probabilities. That is what is meant by being physical: having causal powers on others.<sup>9</sup>

### [Properties of Any Experience](#)

With this definition of causal power in place, IIT accounts for subjective experience via five axioms of phenomenal existence. These axioms are indubitably true for any and all human experiences, are consistent with but independent of each other, and are complete; that is, there are no other axioms that hold universally true for all experiences. These phenomenal axioms serve a similar role in IIT as the axioms of Euclidean geometry taught in school. From these axioms, translated into postulates, all the theory's conclusions are derived or extrapolated.

The axioms are about essential properties of experience. Many articles and my last book have dealt with these, so I will be brief.<sup>10</sup>

The first axiom is *intrinsicity*. This means that any experience is subjective, existing for itself, not for others. It exists from the intrinsic

perspective, from within, not from an outsider's perspective.

The second axiom is *information*. Every experience is specific. It feels a particular way to read this book. If it were different, it would not be this experience.

The third axiom is *integration*. It reflects the unitary, undivided nature of every experience. As I write these lines, I see a canopy of leafless trees (it is winter) on a backdrop of Lake Washington, with the wind driving whitecaps on the waves. This experience is not composed of a left field of view next to a right field of view, with the whining sounds of the wind superimposed onto the two views. No, it is a single holistic experience. Of course, I can shift my eyes, or I can shift my attention to the left or to the right or to the sounds. But these are all slightly different experiences, with each one of them, again, being unitary.

Erwin Schrödinger, the Austrian quantum physicist after whom the famous equation and the unfortunate cat are named, expressed it as follows: "Consciousness is never experienced in the plural, only in the singular. Not only has none of us ever experienced more than one consciousness, but there is also no trace of circumstantial evidence of this ever happening anywhere in the world. If I say that there cannot be more than one consciousness in the same mind, this seems a blunt tautology—we are quite unable to imagine the contrary."<sup>11</sup>

The fourth axiom is *exclusion*. It states that every experience is definite. It has the content it has, neither less nor more. It could not be otherwise. Consider that you don't see outside the border of visual space. It is not part of your visual experience. It is neither grey nor black. Rather, it simply does not exist. Or consider experiencing "a vague hunch" when you enter your home, an ill-defined feeling that something is off, without knowing yet quite what. This is a definite experience, excluding a universe of experiences with other content that you could be experiencing.

The fifth axiom is *composition*. Any experience is structured into components. Its components are phenomenal distinctions—the three distinctions of my terrifying experience in the opening pages were a point of icy-blue brightness conjoint with feelings of terror and ecstasy.<sup>12</sup> When looking at Lake Washington, I see a diversity of trees, with their swaying branches superimposed onto the waves and a cloudy sky in the upper half. Each tree is in a particular location in my visual field, to the left or right,



above or below something else that is farther away or closer by, colored in its specific way. Each wave has its own spatial relationship to the surrounding water, to every other wave, and to everything else. A very large number of such relations make up the phenomenal structure of this experience.

These five axioms are indubitably true. I can't conceive of any experience, including altered states, that does not exist intrinsically, that is generic rather than specific, that is multitudinous rather than unitary, that is indefinite rather than definite, and that does not have some content, however minimal.<sup>13</sup>

### Properties of Physical Existence

IIT pairs each of these five phenomenal axioms with a physical postulate. For something to have an experience, its substrate must satisfy five rules, formulated in terms of physical existence and defined operationally as having the power to make a difference (cause power) and to take a difference (effect power). These intrinsic powers are operationally defined: manipulating something and observing the effect of this manipulation. That's what every biologist does every day.

A specific physical mechanism in a particular state—a circuit with these neurons turned on and those turned off—that satisfies all five postulates is the substrate of a particular conscious experience. In the case of the brain, this physical substrate is also known as the *neural correlates of consciousness* (NCC).<sup>14</sup>

However, the theory is not limited to brain-based experiences. Indeed, it is agnostic as to whether the substrate is a nervous system, an extended root system of a tree, electrical currents circulating in an ocean of superfluid helium II on a lone exoplanet wandering between the stars, or silicon circuitry. However, to keep things simple, I'll remain focused on the brain, picking up the theme of machine consciousness in the final chapter.

Intrinsicity means that the substrate of consciousness must have cause-effect power that is intrinsic: it must be able to take and make a difference within itself.<sup>15</sup>

Information means that the substrate of consciousness is in a specific state, with these neurons on and those ones off; it therefore has cause-effect

power that is specific: its subsets must have specific causes and specific effects, not generic ones.

Integration means that the substrate of consciousness must have cause-effect power that is unitary: the substrate must not be reducible to separate subsets that do not exist for themselves. Nor must the distinctions and relations they specify. Their degree of irreducibility is measured by integrated information, a number symbolized by the lowercase Greek letter phi, written as  $\phi$ , and pronounced *fi*. The sum of the  $\phi$ 's of all distinctions and relations is the integrated information of the circuit in this state, symbolized by the uppercase letter phi,  $\Phi$ . This number measures the irreducibility of the substrate. Something with no integrated information does not exist as an integrated entity, as it can be reduced to two or more subsystems without any loss. The more integrated the information, the more irreducible the substrate, the more it exists for itself, the more it is conscious.

Exclusion means that the substrate of consciousness must have cause-effect power that is definite: its cause-effect structure must be specified by a definite set of units at a definite grain.

Some neurons will be part of the substrate, the neural correlates of consciousness, and some will not, even though they are directly or indirectly connected to each other (shades of the Six Degrees of Kevin Bacon parlor game).<sup>16</sup> The theory makes a sharp distinction between neurons that are part of the substrate, identical to the NCC, and those that are not, a distinction that strikes many neuroscientists as unrealistic given the heavily interconnected nature of neural lace. But consider the US-Canadian border: People born just south of the border are Americans, while those born a mile to the north are Canadians. Although a small distance separates these neighbors who look, speak, and behave quite similarly, the border has massive political, societal, legal, and financial implications. There will likewise be invisible borders between the actual neural correlates of consciousness and connected, nearby neurons. These excluded cells provide the unconscious biases and the enabling factors for the actual substrate of consciousness to be in the state it is in.

So when asking whether a particular neighborhood of the brain are NCC, one must consider a staggering number of candidate networks—those that exactly define the target region, that leave out some neurons along the



border or leave out all neurons in some layer, or that partially overlap with the target region but include nearby neurons, and so on.

How do we pick which makes up the true substrate, the true neural correlates of consciousness? The theory says that of all possible candidates, only the one with maximum integrated information,  $\Phi$ , exists for itself. This follows the principle that only that which exists maximally truly exists. None of the other ones exists intrinsically. This dictat is an example of a so-called extremum principle quite common in physics (e.g., the least action principle<sup>17</sup>). It implies that all the other circuits—say, the left half of the neural correlates of consciousness or this plus some connected neurons in the basal ganglia—exist from an extrinsic point of view but not for themselves if their integrated information  $\Phi$  is less than that of the maximum. Only the maximum exists for itself. Only the circuit that has a maximum of integrated information is conscious. None of the partially or fully overlapping subsets, supersets, or paraset are.<sup>18</sup>

The same approach applies to the question of the spatial or temporal grain at which the unfolded causal powers are evaluated. What exists for itself is the spatio-temporal grain that maximizes integrated information. A priori, the spatial grain could be atoms, molecules, proteins, synapses, dendrites and other subcellular compartments, individual neurons, groups of neurons, or whatever. Likewise, the relevant time scale could be microseconds or less, milliseconds, seconds, or more. Among this multitude, which one is it? It is the one that maximizes integrated information!<sup>19</sup>

Finally, the composition axiom implies that the neural correlates of consciousness must have cause-effect power that is structured: the substrate must have subsets that specify causal distinctions bound by relations, yielding a cause-effect structure. All these distinctions and the hyper-astronomical number of possible relations, each with their own causes and their own effects (growing with the power of two to the power of two to the number of neurons considered), need to be unfolded. This form is vast beyond anyone's power of imagination.<sup>20</sup>

In principle, all these causal powers can be derived from the transition probability matrix of the circuit in a particular state, with these neurons on and those off. The transition probability matrix encapsulates how all components of this circuit respond to all possible perturbations of their

states. The causal powers of the entire network can then be unfolded from this transition probability matrix into a maximally irreducible cause-effect structure, called a  $\Phi$ -structure. Nothing else is needed.

The heart of IIT is an explanatory identity: an experience is fully accounted for by its associated  $\Phi$ -structure. The phenomenal properties of an experience—its quality or how it feels—correspond one-to-one to the physical properties of the intrinsic cause-effect structure unfolded from the underlying substrate. Put poetically, an experience is how the substrate feels like from the inside. Note that an experience is not identical to its  $\Phi$  value, a number, nor to its substrate, the NCC, but is completely accounted for by the unfolded cause-effect structure.

Every aspect of any experience maps one-to-one onto substructures within the  $\Phi$ -structure, with nothing left over on either side. All the content of my experience here and now—space, time, colors, conscious thoughts and beliefs, intentions and desires, doubts and convictions, hopes and fears, memories and future expectations—corresponds to aspects of the  $\Phi$ -structure unfolded from the relevant circuits in my brain. This is a simple yet quite radical claim.

All quality is a structure, not a function, a process, or a computation. One implication is that consciousness is nonalgorithmic; it is not (Turing) computable.

### **The Great Divide of Being**

To discover which entities exist intrinsically, an observing neuroscientist would have to measure the irreducibility of all candidate neurons at all possible spatial and temporal sizes, all candidate distinctions and relations, and all candidate entities. In practice, such an attempt to assess maxima of integrated information exhaustively is out of the question. Instead, one must resort to various approximations and simplifying assumptions. This is a problem for the scientist, however, not for the substrate of consciousness. What exists maximally does not need to perform such exhaustive measurements. Just as a bicycle chain held up at its two ends will naturally fold into the configuration that minimizes its potential energy without exhaustively trying out all possible configurations, so will the form that maximizes integrated information of a particular physical substrate. This

form or structure exists, intrinsically, for itself. All other partially overlapping circuits of lower integrated information exist, but only extrinsically, for others, not for themselves.

Between intrinsic and extrinsic existence passes the most fundamental of all divides: the Great Divide of Being. This is the unbridgeable chasm between what exists in an absolute sense, in and of itself—namely conscious, intrinsic entities—and what exists only in a lesser sense, for others.

Once this Great Divide of Being, between existence for itself and existence for others, between a subject and an object, between absolute and relative existence, is recognized as fundamental, we have grasped the difference between the mental and the physical.

When I dream, I exist for myself, although without any insight into my condition, as the “self” is muted during dreaming. When transitioning into a deep, dreamless sleep, I cross the Great Divide of Being. My consciousness ceases to exist. Likewise when I become comatose following a stroke or accident. I am still alive, albeit on life support, but I have lost what is most precious to me. What remains is a reduced form of existence, for my loved ones and the clinical personnel caring for my body.

Think of early Earth, after it had cooled sufficiently to possess stable oceans. At this point, four billion years ago, the planet was lifeless. Assuming that the integrated information of small assemblies of hydrocarbons is close to zero, ontological dust from the point of view of absolute existence, almost nothing existed for itself. Despite the sun shining, Earth’s surface was dark, without the inner light of consciousness. Erwin Schrödinger rhetorically asked whether such a world would be “a play before empty benches, not existing for anybody, thus quite properly speaking not existing.” Yes, without a conscious audience, there is no play. Just stuff happening.

This also answers the classic Philosophy 101 riddle: “If a tree falls in the forest, does it make a sound if no one is around to hear it?” Indeed, there is no sound without a conscious observer to hear. Furthermore, there isn’t even a tree or a forest, as these concepts depend on a conscious subject discerning a tree from other trees, treating them as different from the soil they are planted in or from the air that surrounds them. Nature knows nothing of these distinctions, of trees and forests, but only of formless stuff.

Without a conscious subject, there are only “atoms and the void,” as Democritus stated, or *toho wa-bohu*, as the Hebrew Bible describes the Earth before the act of creation in Genesis.

The first flickering of phenomenal light likely appeared during the Cambrian explosion, 530 million years ago, as multicellular animals and their primitive nervous nets arose and proliferated. Driven by relentless competitive pressure, the nervous systems of some species on the tree of life grew and became more entangled, such that the light of intrinsic existence burned ever brighter. With capacious brains came minds sophisticated enough to reflect upon themselves and to become aware of the terrible and beautiful universe they found themselves in.

Integrated information theory shares some of the intuitions of panpsychism, the school of thought that holds that consciousness is a fundamental aspect of reality and that experience is much more common than assumed. Paraphrasing Hamlet, there are more things in heaven and earth that are conscious than are dreamed of in modern analytic philosophy. If an organism, such as the neural net of a jellyfish, satisfies the above five postulates, it feels-like-something. Its experience will likely include a sense of hunger, pain and primordial fear when attacked by predators, and bodily feelings of undulating in the sea—perhaps not too dissimilar from the experiences of a third-trimester fetus, floating in its tight aqueous environment.

Of course, most things in the world, like a random group of atoms, grains of sand, cells, neurons, people, trees, cars, pieces of furniture, and stars, are not maxima of cause-effect power and therefore do not exist for themselves—are not conscious. So in that sense, IIT is much more conservative in attributing sentience than panpsychism.

Integrated information theory has no requirement that a brain must house only a single substrate of consciousness. Provided they do not causally overlap, there might be one large substrate, a maximum of integrated information, that includes part of Broca’s area and houses the egoic consciousness that can speak about its experiences, peacefully coexisting with another, nonoverlapping region of neocortex in the same brain that is also a local maximum of cause-effect power. This “minor” substrate would also have experiences yet be muted if it didn’t control the brain’s language regions. This could go a long way toward explaining a

variety of phenomena such as driving while being completely engaged by a phone conversation or a radio show, mind wandering, conversion disorders, and so on.<sup>21</sup>

Finally, there is no threshold of integrated information, say forty-two, below which there is no experience. Provided the system has some itsy-bitsy intrinsic causal power, it will feel-like-something.<sup>22</sup>

Congratulations on making it through the conceptually densest chapter in the book. The next three chapters are easier, dealing with the neuronal footprints of consciousness in the brain and with mystical and psychedelic experiences.

## CHAPTER SIX

---

### consciousness and the brain

It took a while for humanity to discover that mind relates intimately to the drab and inert goo inside the skull: during mummification in ancient Egypt, the brain was scooped out and discarded while other internal organs were lovingly preserved; the Old and New Testaments never mention it even once; even Greek philosophers had little to say about the brain, with Aristotle belittling it as nothing but a cardiac cooling unit.<sup>1</sup>

Instead, most cultures throughout history located the soul in the heart, the one organ visibly and palpably alive. When its beating ceases, consciousness is lost within seconds.<sup>2</sup> The early-seventeenth-century recognition that the heart is nothing more than a muscle circulating blood throughout the body helped evict the soul from the chest to the cerebrum. The birth of the brain-centric age dates to the late seventeenth century, with the publication of *Cerebri Anatome* by English doctor Thomas Willis. This book introduced the term *neurology* and included realistic drawings of the brain's convolutions (instead of depicting them like intestines). At the same time, English polymath Robert Hooke discovered the basic unit of biology, *cells*, using cutting-edge technology: a microscope. His *Micrographia*, in which he published his finding, is a landmark in our groping progression toward understanding life.<sup>3</sup>

As far as we know, all life is cellular. A typical human consists of some thirty trillion cells, the bulk of which are red blood cells. If the human body were a society, it would be an autocratic one, for the central nervous system that lords it over the rest of the body is made up of less than 1 percent of all cells.<sup>4</sup> Understanding of the mind must therefore be centered on the cellular nature of life.

In the late nineteenth and early twentieth centuries, Spanish anatomist Santiago Ramón y Cajal revealed brain cells in all their stupendous glory. Just as heart cells differ from liver or skin cells, there are different types of

brain cells, depending on where they are, how they look, and what they do: pyramidal neurons, Purkinje cells, amacrine cells, spinal motor neurons, chandelier neurons, and so on.<sup>5</sup> Ramón y Cajal's breathtakingly beautiful ink-and-pencil drawings of neural circuits adorn museum exhibits, coffee table books, and my left bicep.

As the centuries accumulated, theorizing about the mind shifted from bulk-mechanical to cellular-electrical models. This is where things turn personal.

### [Tracking the Footprints of Consciousness](#)

I obtained my PhD with a doctoral thesis that modeled how synapses positioned within the dendritic tree, the input region of a single neuron, interact with each other in ways that approximate logical operations.<sup>6</sup> After four years at the Massachusetts Institute of Technology (MIT), I joined the California Institute of Technology (Caltech) in Pasadena as a professor of biology and engineering. It was there that I started to seriously think about the relationship between the mental and the physical.

Enter Francis Crick. He was the physical chemist who, together with James Watson, identified the double-helical structure of DNA as the biochemical carrier of genetic information, formulated molecular biology's central dogma, and helped decipher the code of life. By the time I met Crick, he had left England to live in La Jolla, in Southern California, just two hours south of Caltech, and had switched his intellectual focus from molecules to brains.

Both thrilled to have found a kindred spirit willing to debate synapses, neurons, firing rates, and subjective feelings, Crick and I struck up an intense mentor-protégé relationship that endured for sixteen years and led to two dozen jointly written papers and essays. Purposefully disregarding interminable philosophical debates, we articulated an empirical program to study consciousness in people, monkeys, cats, and rodents, based on a rich literature of neuropsychological case studies of blindsight (seeing without being able to consciously see), differences in the conscious capabilities of the left and the right cortical hemispheres, and amnesic patients, all interpreted through a neuronal lens.<sup>7</sup>

We encountered considerable resistance when exploring that-which-

must-not-be-named. Indeed, I was urged by a Caltech mentor to wait until I had either retired or won a Nobel Prize (as had Crick) to pursue a topic best left to philosophers, religious folks, and mystics. I completely ignored his well-meaning advice. How could science claim that everything in the universe fell under its writ when it failed to address the central fact of existence: phenomenal consciousness? Sure, a solution to the mind-body problem had evaded efforts by humanity's sharpest minds, but that was no excuse for defeatism. Our joint efforts ended on Crick's deathbed, with him dictating edits to our last manuscript on the claustrum, a mysterious sheetlike neuronal structure underneath the neocortex, hours before he died, a scientist until the very end.<sup>8</sup>

Our thinking centered on the footprints left by the mind in the delicate nervous lace of the brain, the neural (or neuronal) correlates of consciousness (NCC). We proposed that an essential signature of the NCC is the firing of many neurons in a periodic manner, about forty times a second. Our 40 Hz hypothesis stipulated that whenever a stimulus evokes such rhythmic firing, the brain becomes consciously aware of the associated information. With such a hypothesis in hand, a myriad of questions in the speculative realm became amenable to experimentation.<sup>9</sup>

Take the experience of mild pain. To discover its footprints, volunteers lie inside a magnetic resonance imaging (MRI) scanner while an uncomfortable amount of heat is applied to an arm via a heating device; brain activity is recorded and contrasted to the situation when the thermal device is at a lower setting. How does brain activity change as the intensity of the experienced pain ramps up? Under ideal conditions, rarely achieved in any single experiment, what shows up will be the NCC, the neuronal substrate of the feeling of pain.<sup>10</sup>

Rigorously defined by philosopher David Chalmers as the minimal neuronal mechanisms jointly sufficient for any one specific conscious percept, the NCC provide a foothold into the mind-body conundrum. This operational approach allows experimentalists to proceed in an agnostic and pragmatic manner, without the need to swear allegiance to any one particular philosophical school of thought.<sup>11</sup>

Research can't stop with finding some correlates of consciousness, as the footprints of selective attention, which usually covary with consciousness, can be mistaken for the latter. If your molar is inflamed, you



will pay attention to the tooth that radiates the pain, much more so than if it were silent, not signaling anything out of the ordinary. Moreover, the correlates of consciousness can be conflated with the correlates of pressing a key on a computer keyboard: if every time the subjects feels a strong pain, they press “3,” while for a lesser pain they press “2” or even “1,” the experimentalist has to make sure that the inferred brain activity reflects the intensity of the pain experience rather than the pressing of those keys or remembering the task. Worrying about how to deal with such confounders is what keeps graduate students up at night.<sup>12</sup>

Finding these correlates is but a starting point for the next round of investigations. Does someone with a high tolerance for pain exhibit reduced NCC activity compared to someone who winces at any minor discomfort? Do the NCC last as long as the pain experience? That is, a brief pain should have a brief NCC while a pain that lasts for minutes should likewise be reflected in the NCC extending for minutes. What are the cellular constituents of the NCC? Do specific types of neurons consistently correlate with the pain? What processes lead up to the NCC? What processes follow it? How do the NCC for thermal pain, a cut finger, a toothache, a migraine, or the distress and agony associated with a divorce differ? Are the same neurons activated but located in different brain regions or are different types of neurons activated altogether? What about the difference between acute and chronic pain or between pain and hunger, pleasure, color, or boredom? Most importantly, what are the causal powers of the NCC? That is, do the brain regions in which the NCC are located correspond to known locations whose loss—say, due to stroke—makes the patient unable to experience pain? Conversely, does activating this brain region, via drugs or implanted electrodes, cause pain?<sup>13</sup>

You’ve probably heard the saying “Correlation is not causation.” Given the highly entangled nature of the central nervous system, neuroscientists evoke this mantra daily. This means that two variables can be correlated, even highly so, without one being the cause of the other. Consider the consumption of ice cream and sunburns. Both variables rise and fall predictably with the seasons. They are correlated, but obviously eating ice cream does not cause a sunburn. Rather, they are linked via a confounding variable, the annual change in solar influx that makes people slurp ice cream in summer, when it is hot, without putting on sunscreen. Conspiracy

theories willfully ignore the “correlation is not causation” mantra and causally seek to link coincidental events via a protracted, covert, and highly unlikely chain of events.

In the biomedical sciences, observing two variables and inferring that, because they move up and down together, one causes the other, does not work. Inferring causation based on correlation is inadequate.<sup>14</sup> Instead, causal interventions are needed: perturbing something over here and observing the effect of this manipulation over there. This is what a well-controlled clinical trial is all about—say, intervening by giving a vaccine and observing whether this reduces illness above and beyond what is expected because of the placebo effect. In the software industry, webpages and graphical user interfaces are constantly tweaked to probe which manipulations maximize stickiness and the likelihood of purchase.<sup>15</sup>

Moreover, it is essential that the NCC are linked, in a causal manner, to subjective experience. This can be achieved by manipulating the NCC, say by turning the associated neurons first off and then on again. Ideally, this should cause some aspect of the experience to disappear and then to reappear.

Let me give you a dramatic case study: a fourteen-year-old girl experienced distinct episodes of guilt and distress in social situations.<sup>16</sup> She eventually developed generalized tonic-clonic seizures and was diagnosed with epilepsy. Brain imaging revealed a tumor bordering the subcallosal cingulate gyrus, tucked underneath the prefrontal cortex. The neurosurgeon who treated her, Itzhak Fried, stimulated this region via an implanted electrode, evoking intense guilt, similar to what she experienced during her seizures. When Fried surgically removed the tumor, her seizures and the associated episodes of intense guilt were eliminated. This is a rare clinical example of linking the gain and the loss of a specific conscious experience, guilt, to a particular brain region. Likewise, electrically stimulating sites within the right fusiform gyrus on the bottom of the posterior cortex causes distortions when seeing faces, while loss of this region following a localized stroke leads to face blindness: the patient does not recognize familiar faces, including their own face in a mirror, although there is nothing wrong with their eyes; for the most part, they see perfectly fine.<sup>17</sup>

The ideas Crick and I advocated for in the early 1990s were innovative because they spelled out a systematic, operational program of how to

discover the NCC (not so much for the idea that any experience must have some specific neuronal substrate, for that had already been discussed a century earlier by none other than Sigmund Freud<sup>18</sup>). I was so excited by the promise of this research program that I challenged the philosopher David Chalmers, after a late-night session in a bar in June 1998, to a wager that he immediately accepted. I gave hostage to fortune by betting a case of fine wine that neuroscience would discover the neural correlates of consciousness within twenty-five years, by 2023 (which seemed a lifetime away).<sup>19</sup>

### **Many Brain Regions Don't Support Consciousness**

Numerous physiological processes are necessary for consciousness. Your lungs need to extract oxygen from the air you inhale and deliver it to trillions of red blood cells that the heart pumps through the brain's extensive network of blood vessels to keep its energy-hungry cells operating. When this supply is even briefly interrupted, such as when the carotid arteries are blocked, people "lose" consciousness, as the brain doesn't have any built-in power reserves and starts shutting down.<sup>20</sup> So blood flow enables consciousness but is insufficient for mind—a comatose patient whose heart is beating gives silent testimony to this.

Many parts of the nervous system play little role in consciousness. The spinal cord, a foot-and-a-half-long conduit of nervous tissue running inside the backbone, houses two hundred million nerve cells. If it is severed, motor and sensory functions of legs, arms, or the trunk are lost. The higher up the site of injury, the more extensive the damage. People lack bodily sensation in their paralyzed limbs, are confined to a wheelchair, and lose control of their bowel, bladder, and other autonomic functions. Yet they continue to experience life—they see and hear, love and fear, imagine the future and recall the past.

The spinal cord merges into the two-inch-long brainstem at the base of the skull. It combines the functionality of a communication hub with that of a power plant. Its circuits regulate arousal, sleep, wakefulness, the pulsation of the heart and lungs, and other vital homeostatic functions. Through its narrow confines pass many of the cranial nerves innervating the face and the neck, transmitting incoming sensory and outgoing motor signals.

If the brainstem is damaged or compressed, death frequently follows. Even quite focal destruction can lead to a profound and sustained loss of consciousness, especially if the damage occurs simultaneously on both sides.<sup>21</sup> This is because brainstem neurons suffuse the neocortex with a cocktail of modulatory substances, helping set the stage on which mental life plays out. But brainstem neurons are not the actors. They do not provide the content of consciousness. Patients with an intact brainstem but widespread cortical damage typically show no signs of consciousness.

The NCC must be distinguished from all these *enabling factors*—cardiac, vascular, metabolic, or neuronal. They are necessary, but not sufficient, background conditions for the mind. Think of the battery powering your laptop. Without it, your computer is just an inert slab of machinery. Yet the power source can't stream movies, generate the stunning graphics of *No Man's Sky*, process files, or store photos.

The cerebellum is the “little brain” tucked underneath the neocortex, at the back of the head. It instantiates the automatic, “thoughtless” processes that silently coordinate sensory information streaming in from stretch and position sensors embedded in muscles and joints, as well as the equilibrium organs in the inner ear and the eyes, with the motor commands that go out to the hundreds of muscles in the limbs and trunk. The cerebellum choreographs actions—running while visually tracking someone, driving while speaking, typing on a keyboard, dancing to rock and roll, playing basketball, tennis, and so on. With sufficient training these actions flow effortlessly, a synchronized orchestra of motion.

If parts of the cerebellum are lost to stroke or to the surgeon's knife, the orchestra turns discordant, cacophonous, at least in some of its sections. The patient becomes ataxic, their movements clumsy. They lose the fluid ability to play a musical instrument or speed type on their phone. Yet their subjective experience remains intact; they can be highly articulate, witty, vibrant.<sup>22</sup> A handful of individuals are born without a cerebellum altogether, leading to delayed development and cognitive deficits. Yet they are not zombies; they experience the world in all the usual ways.<sup>23</sup> The brain's most distinct neurons are cerebellar Purkinje cells with their flat, fan-shaped dendritic tree, collectively receiving excitation from tens of billions of cerebellar granule cells. That's four times more than all the neurons in the rest of the brain combined. Yet even this generous neuronal

endowment is insufficient to generate feelings.<sup>24</sup>

Hints for why this might be so are found within the cerebellum's stereotyped, crystalline circuitry, divided into independent modules that number in the hundreds or more. Each one operates in parallel, with nonoverlapping inputs and outputs, and is wired in a feed-forward manner. That is, one set of neurons feeds the next one, which in turn influences a third. There are none of the reverberatory, self-sustaining excitatory loops prevalent in the neocortex, but there is plenty of inhibitory feedback to quench longer responses.<sup>25</sup>

These observations refute the myth that consciousness simply arises from neurons doing their thing. Here are billions of cerebellar cells doing what comes naturally to them, firing action potentials and releasing little squirts of neurotransmitter, yet without any feelings. What matters is not the constitution of brain tissue but the way it is wired, its structure. A cerebellum-like architecture, with its myriad independent circuits, is insufficient for consciousness.

### [Look to the Neocortex](#)

The substrate of the mental is the neocortex and allied satellite structures, sited like a crown on top of the brain, just underneath its protective skull. It is strongly and reciprocally connected to the thalamus, about the size of a quail egg, in the middle of the brain. Making up about 80 percent of the mass of the brain, the neocortex is a defining hallmark of all mammals, a multilayered, extended sheet of dense nervous tissue, the proverbial grey matter. Unfolded, it has the size and width of a fourteen-inch pizza with toppings. Two of these highly folded sheets, the left and right hemispheres, squeeze into the cranium.<sup>26</sup>

A large region in the back of the neocortex, including temporal, parietal, and occipital neighborhoods, is closely linked to subjective experiences of sight, hearing, touch, and sensing of the body and the self. Because of its close association with consciousness, this region, the current best NCC candidate, is termed the posterior hot zone (see [Figure 1](#)).

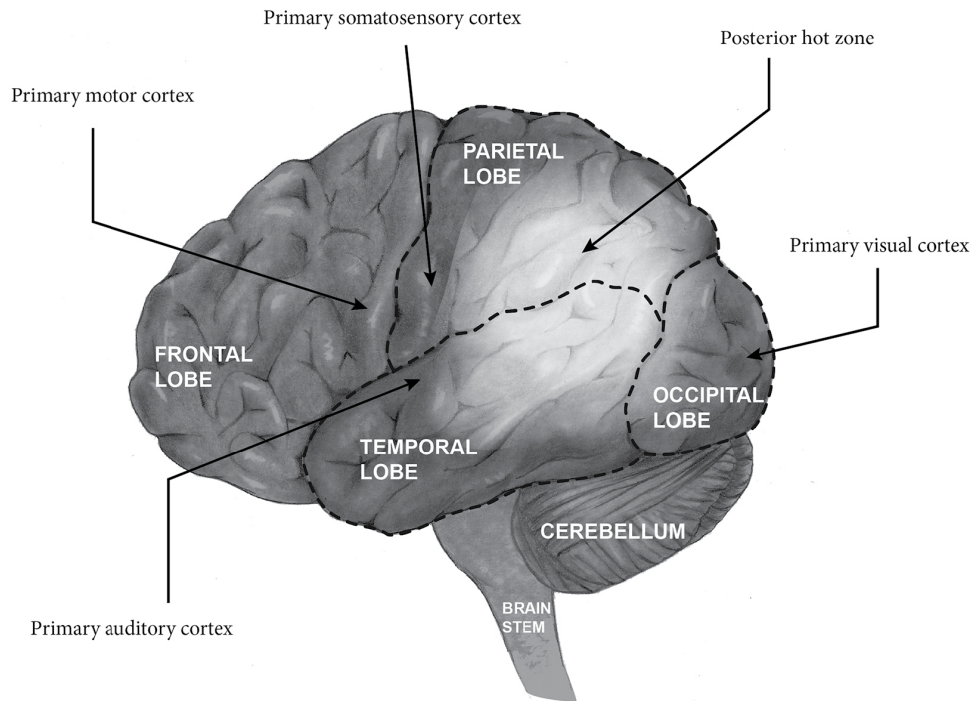


Figure 1. A drawing of the left hemisphere of the human brain.

Neurons in the brainstem, merging downward into the spinal cord, provide critical background conditions for consciousness. The cerebellum, with four times more neurons than the rest of the brain, is incapable of producing feelings. The outermost, highly folded and layered nervous tissue is the neocortex, the jewel in the crown of the nervous system. A set of adjacent regions in its back, called the posterior hot zone, extending into occipital, temporal, parietal, and posterior cingulate areas, is the substrate for consciously seeing, hearing, and feeling the body and the self.

Most of the brain's frontal regions—in particular the prefrontal cortex, the territory in front of the motor strip (that includes the primary motor cortex) that runs perpendicular to the midline separating the two brain hemispheres—are not the substrate for seeing, hearing, touching, willing, or the sense of self.<sup>27</sup> The frontal lobe, greatly expanded in humans compared to great apes, is critically involved in reasoning, planning, speaking, and other cognitive operations closely linked to intelligence but not with consciousness per se.<sup>28</sup>



That the mental hails from cortical neighborhoods in the back is supported by three broad classes of evidence: First, spatially localized destruction of nervous tissue, called *lesions*, whether caused accidentally by a clogged artery, a ruptured blood vessel, or deliberately, by a surgeon to remove a tumor or an epileptic focus, offer clues in the pattern of deficits they leave behind in the mental lives of patients. Second, electrical stimulation of some cortical regions elicits distinct experiences or memories, shades of *The Matrix*. Again, this points to the direct involvement of these regions in distinct types of conscious experiences. Third—and weakest from an evidentiary, inferential point of view—brain-scanning experiments correlate specific experiences to activity in specific neocortical regions.

Vis-à-vis the first source of evidence, an entire class of experiences can be wiped out by the circumscribed loss of tissue in different neighborhoods within the posterior hot zone. Depending on the exact location, affected individuals may be unable to perceive color or moving objects, recognize faces, identify objects by touch, discern familiar voices, or comprehend speech or text. Such specific deficits in the absence of anything wrong with the eyes or ears, called *agnosia* (from the Greek for absence of knowledge), implicate the destroyed regions in conscious color, face, or motion perception and so on.<sup>29</sup>

Remarkably, these individuals often don't realize that something is amiss. If they lose feeling in an arm or a leg, they may not recognize the limb as their own anymore, with unusual consequences, such as alien hand syndrome.<sup>30</sup> They won't notice that they are blind in part of their visual field. A patient with a lesion in his right parietal lobe will deny that he's blind on his left. He might reluctantly admit that he runs into doorways on the left and doesn't find food or utensils there, but the inference that he doesn't see is overpowered by the fact that there is no blank or black region in his field of view. Clinicians refer to such deficits of self-awareness as *anosognosia*.<sup>31</sup>

How can this be? If half of your computer screen were to freeze or turn to static noise, you would know because your representation for visual space is intact. Yet were you to lose this neuronal substrate of seeing space, you wouldn't see anything. That's why there isn't blank or black space outside your field of view, such as behind your back; there's simply

nothing. Your brain didn't develop a substrate for visual space in those regions.<sup>32</sup> Not-knowing is common in Wernicke's aphasia following damage to the left superior temporal gyrus in the back—patients talk fluently but what they say makes little sense. Remarkably, they are not aware of this, happily speaking their mind, as they do not hear themselves.<sup>33</sup>

When prefrontal regions are either surgically removed or fall prey to some calamity, patients continue to see, hear, and sense. Deficits affect the higher mental faculties, resulting in apathy, a lack of curiosity about the world, and a disability in planning, introspecting, reasoning, and regulating emotions.<sup>34</sup> Patients perseverate or consistently make bad choices. This supports my contention (to which I'll return in the final chapter) that intelligence is different from consciousness. Intelligence is about behaving in a flexible, context-dependent manner, predicting and planning for the near and far future, distinct from subjectivity.<sup>35</sup>

Interfering with the neurophysiology of a particular brain region provides a second source of strong, causal evidence about its contribution to mind. This can be caused by aberrant electrical activity of a localized epileptic seizure. Or brain circuits can be willfully disturbed via an electrode as part of a routine clinical workup prior to neurosurgery. Stimulating the posterior hot zone in this manner can spark a litany of sensations and feelings: flashes of lights, geometric shapes, colors and movement, face distortions, auditory hallucinations, feelings of familiarity (*déjà vu*) or unreality, the urge to move a limb, burning pain, or an out-of-body experience. Contrariwise, most regions of the prefrontal, non-eloquent cortex are silent when so stimulated. Patients feel little; if they do, their experiences relate to smells and thoughts.<sup>36</sup>

The induced experiences can be quite bizarre: a high-functioning Silicon Valley executive suffers from partial seizures in a region within the posteromedial cortex (on the inward-facing side of the cortical hemisphere, along the midline) during which his sense of self is distorted, including his perceived location in space; he eavesdrops on his own thoughts, a form of depersonalization. Direct electrical stimulation of this region, or its match on the other side, in this individual and in other patients induces feelings of slipping, falling, floating, being displaced, self-dissociation, and euphoria. It appears to be a site critical to the subjective sense of self, both its



experiential (as in, “I am standing”) and its narrative (as in, “I walked across the airport”) aspects.<sup>37</sup>

The third source of evidence is untold observational studies in which healthy volunteers do some task—say, push a button when they recognize a particular face—while the bulk-tissue activity of their brain is monitored using functional magnetic resonance imaging (fMRI), that tracks changes in blood flow and volume; electroencephalogram (EEG) electrodes picking up electrical signals; and magnetoencephalography (MEG) instruments recording magnetic signals.<sup>38</sup>

Let’s look at the latest and arguably the most powerful study. Called an *adversarial collaboration*, this large international project depended on getting two conflicting theories, and the associated scientists, to agree (in writing) on two experiments whose outcome would clearly favor one or the other rival.<sup>39</sup> Reconciling protagonists who fundamentally disagree on what they mean by consciousness and aligning a dozen labs in different countries on the same exacting protocols is nontrivial, as they say in the trade. The matchup was between integrated information theory (IIT) and global neuronal workspace theory (GNWT), today’s two most prominent theories of consciousness. The latter is a functional, computational account of the mind in which information accesses a global workspace by broadcasting it from the prefrontal cortex, in front of the brain, to the back of the cortex, thereby generating consciousness.<sup>40</sup>

The two contending theories differ in many ways, including their metaphysical foundations and their attitude toward phenomenology: IIT starts with consciousness and infers its substrate from there, while GNWT seeks to distill consciousness out of computations carried out by the brain; IIT emphasizes the rich, subjective nature of perceptual experience, while GNWT stresses that what people report is limited to a handful of items at any one point in time, such as the identity of a face or a thought.

This question—Is the content of consciousness rich or sparse?—is epitomized by Pieter Bruegel’s famous winter landscape *Hunters in the Snow*. Like all of Bruegel’s paintings, it is vibrant with life, resplendent with detail. This realism reflects a common experience when looking around: the world is indeed replete with particulars—textures, colors, blobs and oriented lines, some stationary, some moving, all painted onto a spatial canvas. According to workspace theory, this apparent richness is an illusion

—consciousness is restricted to a few high-level, thought-like items, describing the gist of the painting, akin to the label a museum might place next to the painting: “Winter scene of hunters and their dogs returning empty-handed with a late-medieval village in the background.”<sup>41</sup>

The two theories also differ on where the NCC are located, their timing, and the way the NCC are coordinated across the cortex; these differences were to be experimentally resolved by the adversarial collaboration. Crudely put, workspace theory postulates that activity in the prefrontal cortex is essential, that activity there arises when the subject initially becomes aware of the image (in these experiments, subjects were shown highly visible pictures, flashed on the screen, one by one) and, again, when the image disappears and is replaced by a grey and empty screen (but not in the intervening interval when the image remains clearly visible on the screen). IIT argues that the NCC must be in the posterior hot zone as its topographic connectivity maximizes intrinsic causal power and posits that the NCC persist for as long as the subject is conscious of the image. Both agree that their candidate NCC should achieve a high degree of synchrony between the front and the back or between the posterior hot zone and the early visual cortical areas.<sup>42</sup>

The two experiments were exquisitely designed to test these three predictions for visual perception—that is, for consciously seeing individual faces, objects, numbers, and letters. The experiments combined the three standard instruments of cognitive neuroscience (fMRI, EEG, and MEG) in volunteers and combined that with a fourth method: recording from electrodes implanted underneath the skull in patients monitored for occurrences of epileptic seizures. Given the proximity of these electrodes to the neocortex, this yields high-quality signals, less noisy than EEG electrodes positioned outside the skull and the scalp. All these conjoint methods, together with the intended analysis algorithms, had to be pre-registered. This step is rare; it means that scientists had to decide, ahead of time, which type of mathematical decoding in what brain regions to apply to the data, with what statistical parameters, and so on, instead of making this crucial determination after looking at the data and opting for the type of analysis that confirms one’s biases (this is usually what happens). The collaboration recruited 250 subjects, vastly more than for any comparable experiment; all data was collected by at least two independent laboratories;

and all data will be released to the public so anyone can mine it or come to their own conclusions.<sup>43</sup>

The results of the first experiment were publicly announced at an exuberant event, including an artist rapping about consciousness and zombies, in New York in June 2023: the results strongly favored the NCC as being in the back, the posterior hot zone, with its timing consistent with the duration of the stimulus's visibility.<sup>44</sup> Both of these predictions clearly favor integrated information theory and seriously call into question the necessity of the prefrontal cortex for seeing. However, the third prediction favored workspace theory. Of course, the dataset is massive, with more discoveries to be mined, and the second experiment has not yet been evaluated. To no one's surprise, spokespeople for both theories profess themselves satisfied with the results!<sup>45</sup> The quarter-century wager between me and Chalmers was also up that evening: given these discordant results, I publicly admitted that the community had not converged on a broadly accepted NCC and handed over a case of fine wine to Chalmers. While *Nature* proclaimed, "Philosopher 1, neuroscientist 0," to me it felt like a case of losing the battle but neuroscience winning, as more had been learned about the footprints of consciousness in the past twenty-five years than in all of previous history.<sup>46</sup> However, this public event, widely covered by the mainstream press, a triumph for the scientific method, had an ugly aftermath.<sup>47</sup>

No one denies the import of the front of the brain for action, reasoning, and intelligence. But these capabilities are not the same as consciousness. The latter appears to have a close association with the posterior hot zone in the back. Of course, the underlying atoms of consciousness are not entire regions containing hundreds of millions of cells but far-flung coalitions of highly heterogeneous neurons. It is to discrete neuronal assemblies—perhaps as small as a few hundred neurons, perhaps as vast as tens of millions—that we must look for any one specific experience. Disentangling the vast multiplicity of mechanisms to identify those responsible for any one experience, the true NCC, is a monumental task that will take the rest of the century.

The neocortex is one of a handful of defining features of the class of mammals. It is likely that in all, loss of the neocortex precludes conscious experience.<sup>48</sup> Other vertebrates—fish, amphibians, reptiles, and birds—do

not have a neocortex but possess functionally related anatomical structures that can support experience. As I wrote these lines, a beautiful iridescent-green hummingbird flew against a glass pane on my deck. It was tweeting, its tiny body beating rapidly, obviously in distress; I tried to help it as best as I could, a fellow creature, stunned, in pain, and, probably, afraid. Similarly, invertebrates, such as bees and cephalopods, display complex cognition and capabilities. Yet they have no extended sheetlike neural architecture. Should we deny them sentience because they do not have a neocortex? No! The evolutionary and behavioral evidence is compatible with the thesis that all animals are sentient to a larger or smaller degree. This is also the prediction of integrated information theory. Given the ten-times-greater circuit density of the tiny brain of a bee (compared to the mammalian neocortex), with its approximately one million neurons wired up in stunningly complex patterns, it too will have a cause-effect structure with a nonzero amount of integrated information. It too is likely to experience some degree of contentment, flying in the warm rays of the sun, carrying a load of golden nectar back to its sisters in the hive.<sup>49</sup> Indeed, it may be that every organism on the tree of life feels-like-something, is sentient, although its phenomenal content may take a primitive form unrecognizable to us.<sup>50</sup>

The intimate relationship between the posterior hot zone and experience might appear paradoxical as it is the prefrontal cortex that is most highly developed in humans compared to close evolutionary relatives, such as monkeys and great apes.<sup>51</sup> But if we admit that they too see, hear, dread, and desire, the paradox disappears. What makes humans different is not so much our basic sensorial, experiential repertoire but a powerful language module, flexible intelligence, an ability to self-reflect, and a hypertrophied sense of self-importance.

### **Stranded Minds in Damaged Brains**

This knowledge is not just of interest to geeky neuroscientists but has direct clinical application to detect consciousness in people who can't communicate in any way with the outside world. *Disorders of consciousness* occur in a patient whose brain has been severely injured by blunt force trauma to the head, cardiac arrest, ischemic stroke, encephalitis,

drug overdose, and a litany of other tribulations. Bedridden and on life support, they can't speak, wink, or otherwise signal. Is the injured person like a stranded astronaut, conscious but entombed in an impaired brain, unable to communicate with loved ones at their bedside? Or is truly no one at home?

Clinically, consciousness is assessed using a series of simple bedside tests that require the patient to be able to speak or otherwise respond intelligibly. The examiner goes through a standardized checklist, such as the "Coma Recovery Scale—Revised," asking the patient to track a bright light with their eyes, move a limb, respond to a painful pinch, or say something. If they know their name, where they are, the time of year, and so on, or, when unable to speak, if they can still follow instructions and communicate by blinking their eyes or moving their hand, they are assumed to be conscious, however impaired they may otherwise be. If they repeatedly fail these tests, they are assumed to be unconscious. If this lack of voluntary responses lasts weeks, the patient is considered to be in a vegetative state, nowadays called a behaviorally unresponsive state.

However, if the damage is to the auditory cortex, the patient may be conscious but unable to hear; if motor regions are destroyed, they may be unable to signal. Or they may be cognitively so impaired that they are unable to answer coherently but can still feel. That is, the patient may be conscious but unable to signal that they are present. It is estimated that of the millions of people worldwide with a severe brain injury, up to one in five may be conscious yet unable to communicate with the outside world—like Major Tom in "Space Oddity" by David Bowie, a marooned astronaut whose radio link is down.<sup>52</sup> This is tragic because it would give succor to the patient to know that they were being heard and would certainly lessen the emotional burden for the family to know that their parent, sibling, or child was "there." Furthermore, early recovery of consciousness is good news, predicting long-term functional recovery. But most importantly, when a patient is unresponsive and has a bleak long-term prognosis for recovery, the family and the clinical team caring for them often opt for withdrawal of life-sustaining therapy. Indeed, cessation of life support, usually by discontinuing mechanical ventilation, is the most common cause of death among patients suffering acute disorders of consciousness.<sup>53</sup>

What is needed is a biomarker for consciousness that works in the

intensive care unit or convalescent home. This is a big unmet need, as recording spontaneous electrical activity from a brain so severely impaired that the patient is unable to signal their “presence” does not reliably distinguish consciousness from its absence.

Over the past twenty years, a team led by Marcello Massimini at the University of Milan in Italy has developed a technique that stimulates the neocortex with a magnetic pulse, delivered by a transcranial magnetic stimulation coil to the skull, and records the electrical echoes using EEG electrodes. It’s a bit like knocking a bell with a small hammer and listening to the reverberations of the sound waves. Inspired by the theoretical principles of differentiation (information) and integration discussed in [Chapter 5](#), this method computes the complexity of the EEG response to the magnetic pulse and spits out a number, the perturbational complexity index. If this index is high, the subject is conscious, whereas a value below a certain threshold indicates unconsciousness. Probing the complexity of the brain’s response in this manner works flawlessly in a validation dataset of volunteers and patients whose state of consciousness is known. That is, this test distinguishes the unconscious sleeping or anesthetized brains of volunteers or of patients in a coma from the conscious brains of awake or dreaming volunteers, people under ketamine (a dissociative), and neurological patients who are conscious. Furthermore, this complexity test confirms consciousness in nineteen out of twenty minimally conscious patients.

Hospitals in the United States and Europe are currently evaluating the extent to which this technique can be used as a primitive consciousness detector. Indeed, Massimini, Giulio Tononi, and I cofounded a company, Intrinsic Powers, that seeks to develop a practical device to routinely monitor consciousness in unresponsive patients at the bedside in this manner. The road to approval by the regulatory bodies and acceptance by clinicians is long and winding, but we seek to provide certainty to the caretakers and families of patients who must deal with harsh and tragic circumstances.<sup>54</sup>

Mind-body discussions often have an abstract and erudite quality about them. Yet in the ICU, when the chips are down, determining whether or not life support should be withdrawn is as serious a decision as it gets in life.

Technology, guided by the integrated theory of consciousness, will bring

light to this gloomy discussion concerning impaired consciousness. Let me now describe the opposite: expanded consciousness.

## CHAPTER SEVEN

---

### expanding consciousness

The person who is involved in this perception is no longer an individual, for in such perception the individual has lost himself; he is pure will-less, painless, timeless subject of knowledge.

—Arthur Schopenhauer, *The World as Will and Representation*

**W**e spend most of our lives on autopilot, driven by incessant demands of family, career, money, and other daily concerns. But on occasion, something happens that abruptly resets the seemingly determined path of our lives.

The book opened with my experience during the year of the pandemic—entering a solipsistic state of awe and terror, outside time, without a sense of self, of body, or of the world—that left lasting traces. Throughout the wide arc of history, many have experienced similar states that bring about enduring and pervasive revisions in basic beliefs concerning the nature of reality. Deeply rooted habits are abandoned, and attitudes change in ways best described as spiritual: loss of fear of death, detachment from material possessions, and an orientation toward the greater good. Common to many transformative experiences is the dissolution of the self, including loss of ego and the body it is chained to.

The self is a malleable set of mental processes that mediate thoughts, inner speech, memories of the past, and plans for the future. It is the “I” that experiences and that bears autobiographical memories, strengths and weaknesses, likes and dislikes. The experiencing and the remembering selves are an enormous complex bundle of overlapping cognitive modules. These develop during childhood and do not fully mature until middle age.



Their neural basis includes the posterior and the anterior cingulate cortices (structures along the midline of the two hemispheres) and the medial prefrontal cortex.

This “I” consciously thinks, perceives, senses, and interacts with the world. It is your perennial companion, essential for self-reflection and for pursuing long-term goals—say, going to professional school for an advanced degree. Yet the mental chatter and negative thoughts that accompany the self can become overwhelming at times. Ruminating and obsessing about perceived slights, catastrophizing, brooding, and so on can drown out other thoughts and make life miserable. You can, for a limited time, become self-less, during states of flow when you are fully immersed in playing soccer, climbing, coding, or doing mathematics; when you are in the here and now, content. Yet, if you lose this focus on the task at hand, the self will quickly reassert itself, grounding you.

The experience of self is as real as any other conscious experience, such as pain or pleasure. What is illusory, as emphasized by Buddhism, is the idea of a permanent and fixed essence that constitutes the “true self,” the “real me.” The elusive authenticity that politicians on the campaign trail seek to project is a myth, a shape-shifting illusion.<sup>1</sup>

Transformative experiences let the individual escape the gravitational field of the self to float weightlessly above planet ego. All those clamorous thoughts, darting hither and yon, the “monkey mind,” are gone, replaced by a heard silence. The life of the subject who underwent such a profound encounter is marked by a clear distinction between “before” and the transformed “afterward.”<sup>2</sup>

There are three categories of transformative experiences depending on their proximal causes: a first class that includes religious, mystical, and aesthetic experiences; a second class of experiences induced by psychedelics; and a third associated with near-death experiences. I subsume all under the general heading of transformative experiences as this best describes their impact, is neutral with respect to its ontological implications, and reflects my belief that they are all woven from the same skein and share a core of common psychological and neurobiological traits.<sup>3</sup>

### **Religious, Mystical, and Aesthetic Experiences**

The lives of saints and sages from many religions are filled with foundational episodes of encountering the living God or otherwise directly sensing the presence of something all-encompassing, all-powerful, numinous. Consider Blaise Pascal, the seventeenth-century French philosopher, mathematician, and physicist, pioneer of probability calculus, and coinventor of the mechanical calculator. After Pascal's death, a handwritten parchment was discovered sewed inside his coat, documenting his divine experience on the night of November 23, 1654. The description starts, "Fire. The God of Abraham, the God of Isaac, the God of Jacob. Not of the philosophers and intellectuals. Certitude, certitude, feeling, joy, peace."

People return from experiences like Pascal's sobbing or radiant, babbling about "everything being love," about "being bathed in the cosmic light of harmony and unity," and becoming one with the universe. These and similar sentiments may sound trite and don't make much sense to those who haven't experienced them. But... but they have the power to radically reshape the life of the experiencer. In some, they prompt a religious awakening. The classic one is the Pauline conversion, described in the Acts of the Apostles: Saul's transformation on the road to Damascus from persecutor of Christians into their greatest advocate, St. Paul.<sup>4</sup> Others shed their old ways, abandoning chasing the almighty dollar to work for communal, healing, environmental, or spiritual causes. They become less driven, at peace with themselves and the world.

Little wonder such experiences appear miraculous. Changing deeply engrained habits is arduous and well-nigh impossible—think of your struggles to lose weight, drink less, quit smoking, work out more, or stop obsessively checking your phone. It takes discipline and protracted willful efforts over the years. Yet a single experience has the power to sweep all difficulties out of the way.

Religious experiences strike with the unpredictability and force of lightning. They occur but once in a lifetime, if at all. Some historical personalities—Hildegard of Bingen, Julian of Norwich, St. Teresa of Avila, Joan of Arc, Meister Eckhart, Rumi, Khalil Gibran—have had multiple encounters, perhaps triggered by temporal lobe seizures.<sup>5</sup>

William James's *The Varieties of Religious Experience* contains a detailed account of them. James remarks that experiencers bring back a

state of knowing, what he calls a noetic quality: “Although so similar to states of feeling, mystical states seem to those who experience them to be also states of knowledge. They are states of insight into depths of truth unplumbed by the discursive intellect. They are illuminations, revelations, full of significance and importance, all inarticulate though they remain; and as a rule they carry with them a curious sense of authority for after-time.”

Not all such experiences are explicitly religious in character; some are more spiritual, without a revelation of a deity. The *overview effect* is a self-transcendent experience frequently reported by astronauts and cosmonauts during spaceflight when viewing the staggering beauty of the cloud-covered blue orb of Earth, floating in the immense vastness of empty space, directly apprehending the planet’s perceived fragility and the fundamental kinship of all living things. Astronaut Edgar Mitchell described his experience on the Apollo 14 flight as an “overwhelming sense of oneness and connectedness... accompanied by an ecstasy... an epiphany.”<sup>6</sup>

Primatologist Jane Goodall recounts her experience in the forest of Gombe after an exhausting six-week trip as follows:

Lost in awe at the beauty around me, I must have slipped into a state of heightened awareness. It is hard—impossible really—to put into words the moment of truth that suddenly came upon me then. Even the mystics are unable to describe their brief flashes of spiritual ecstasy. It seemed to me, as I struggled afterward to recall the experience, the self was utterly absent: I and the chimpanzees, the earth and trees and air, seemed to merge, to become one with the spirit power of life itself. The air was filled with a feathered symphony, the evensong of birds.... That afternoon, it had been as though an unseen hand had drawn back a curtain and, for the briefest moment, I had seen through such a window.<sup>7</sup>

People believe they have been granted a vision of the world beyond the confines of Plato’s cave, accessing unmediated reality. “For the first time in my life I knew exactly—because, thanks to the power, I was doing it—what it means to love one’s neighbor as oneself” (the poet Wystan Hugh Auden), “the direct, total awareness, from the inside, so to say, of love as the

primary and fundamental cosmic fact” (the writer Aldous Huxley), “that in this I had attained a union with the primal being or the godhead” (the philosopher Martin Buber). To an outsider, this is incomprehensible.

Some seek the exalted in nature. While most hikers, climbers, and campers flee from snow, thunder, lightning, and storms, two sorts of people head into dramatic weather in the mountains: those tasked with rescuing folks and those seeking to test their mettle and encounter the sublime. I have always experienced the attraction of the latter!

Many feel awe and a sense of wonder when losing themselves in music, the most powerful of all art forms, be it chamber music or heavy metal. Aesthetic experiences take us away from our mundane concerns, into the realm of the infinite. For philosopher Arthur Schopenhauer, the perennial duality between the subject and the object, the perceiver and the perceived, the knower and the known, ceases in aesthetic experiences. “We... devote the whole power of our mind to perception, sink ourselves completely therein, and let our whole consciousness be filled by the calm contemplation of the natural object actually present... We lose ourselves entirely in this object... we forget our individuality, our will... so that it is as though the object alone existed without anyone to perceive it and thus we are no longer able to separate the perceiver from the perception, but the two have become one.”<sup>8</sup>

One obvious way of dealing with these radical and uninterpretable claims is to disregard them as the delirious gibberish of drunken *confabulatores nocturni*, men paid to tell stories during the night, to dismiss their mental states as pathophysiological and retreat to the lab bench to safely study genes, receptors, brains, and other concrete things. This avenue is foreclosed to me given my own experience of terror and ecstasy. I know beyond a doubt that this is what it was. That is my noetic quality. Of course, you don't have to believe me.

It is, however, a matter of historical record that people have been impacted by their extraordinary experiences, discovering the biblical “peace of God that passeth all understanding,” altering their way of life. Thus a more nuanced approach is to accept these reports as authentic and honest descriptions. They teach us that our central nervous systems can enter physical states, with these neurons on and those off, commensurate with experiencing the sidereal and the infernal and everything in between. This is

a gift we should all take advantage of.

What precipitates these experiences? They seem to depend on years of fervent commitment to a worthy cause.<sup>9</sup> But that itself is insufficient; otherwise legions of investment bankers, start-up employees, and other workaholics would routinely gush about the absolute and the ineffable. Some additional, unknown, and rare component is necessary since only a vanishingly small fraction of humanity report such experiences.<sup>10</sup> God, fate, or chance seems to grant to a few a vision of the Whole. The special mindset that produces such states may as well be thought of as an act of unearned grace.

The likelihood of such experiences in those not prone to them endogenously can be enhanced by a dizzying variety of practices developed by different cultures over millennia. These consciousness-modification techniques range widely, from the extreme social isolation and fasting of the early Christian hermits and ascetics living in the desert to facilitate interior silence, to extended meditations in the Buddhist tradition to ascend a hierarchy of higher mental states, to the endless practice of sword fighting or archery to lose the sense of agency and self, to self-flagellation and other forms of mortification, sensory deprivation (such as in a floatation tank), intense physical exertion (such as long-distance running), chanting and ecstatic dancing to induce a trance (such as practiced by the Whirling Dervishes in the Sufi tradition or at a rave), and rapid, protracted breathing (such as holotropic breathwork).<sup>11</sup> And then there are occult substances.

### [Psychedelic Experiences](#)

Many cultures throughout history cured illness, divined the future, or communicated with spirits by drinking, eating, smoking, or inhaling special “medicines,” what we today consider psychoactive drugs. Participants experience a realm beyond the everyday, in which they leave their bodies, converse with gods and demons, or undergo psychic death. Shamanic healing sessions are part of regulated group rituals that include fasting, chanting, meditating, and praying to achieve the right mindset. *Curanderos*, traditional healers or shamans, guide participants, watch over them, and help them make sense of their experiences. Today, these substances are known as *psychedelics*, a word derived from the Greek for “mind-

manifesting,” or as *entheogens*, referring to their power to induce spiritual feelings of presence, wonder, awe, and the divine.<sup>12</sup>

Most psychedelics are naturally occurring substances derived either from plants or from animal glands. In the twentieth century, chemists isolated, purified, and characterized the psychoactive molecules. These can now be manufactured in synthetic forms at scale. They bind to specific sites on the surface of neurons and elsewhere in the body, primarily to those associated with the endogenous neurotransmitter serotonin. They affect the mind in profound ways. Despite their illicit nature, the use of these medicines is flourishing.

Besides magic mushrooms and their primary psychoactive ingredient, psilocybin, the classical serotonergic psychedelics include mescaline, or 3,4,5-trimethoxyphenethylamine; N,N-dimethyltryptamine (DMT), known as the “spirit” molecule and synthesized in the brain;<sup>13</sup> the DMT-containing ayahuasca brew, known as “the leaf and the vine”; 5-methoxy-N,N-dimethyltryptamine (5-MeO-DMT), known as the “God” molecule; and lysergic acid diethylamide (LSD), or “acid” (a synthetic substance). Their acute effects last anywhere between ten minutes and ten hours, depending on the molecule, amount, and delivery method.

At lower doses, psychedelics distort normal waking consciousness without impairing the senses, general intelligence, or motor control. They induce motion trails and render colors more vivid and shimmering. They enhance aesthetic and spiritual feelings of gratefulness and sacredness and turn down, sometimes all the way to zero, the self. That voice in the head that is always cajoling, complaining, criticizing: gone—leaving the mind free to contemplate the beauty of the world. Sheer existence appears more authentic, more real than everyday life, like the *ligne claire* Tintin drawings by Hergé with his fanatical devotion to detail. This hyper-realization of the world feels like coming home and is the opposite of what in psychiatry is known as derealization, when patients complain that the world feels unreal, fake, insipid, lacking in emotional coloring. Yet psychedelics can also elicit chthonic visions; induce temporary states of fear, panic, or confusion; and bring up suppressed emotions.

It was in the Indian summer of my life, decades after I started working on the neural basis of consciousness, that I tripped for the first time. For several hours, my diminished “I” was under an enchanted spell of

indolence, languor, and lethargy, being rather than doing anything. The experience filled me with a profound sense of contentment, wonder, and gratitude. I wrote in my notes,

I find myself in a corner of Eden before the Fall. I stare at a wooden bench and see it in a way I have never seen before—I apprehend the essence of this specific bench in this specific setting, with the green-blue-cyan moss growing on its humid planks, fluorescing, pulsating, palpitating, quivering, coalescing, and disassembling, fusing and breaking up, throbbing with life itself. I'm utterly absorbed by its enigmatic beauty, in awe at its complexity, deciphering the ineffable message it conveys—a thing of mesmerizing beauty. I can't stop grinning, like a child seeing the world for the first time. I viscerally hear a rhythm, the beating heart of the cosmos, animating everything. I eventually detach my attention from the bench and focus on the bubbles in a glass of sparkling water—individual air pockets coalescing and disappearing into the greater whole. I look at my hand and at the faces of those around me—everything is overlain with a patina of serenity.<sup>14</sup>

Psychedelics can induce profound visions with an astonishing verisimilitude that any Hollywood director would envy. These are not hallucinations, as the user understands that what they experience is not “waking” reality; furthermore, these visions occur with closed eyes. A trained observer, a psychologist, describes the power of ayahuasca, a bitter, entheogenic concoction made from shrubs and vines of the Amazon Basin to induce visions:

All things imaginable and non-imaginable can be seen with Ayahuasca. One can see all the moments of one's life, all the people and places that one knows, Nature and the Cosmos in all their manifestations, human history and the different cultures that it has and has not produced, and scenes that lead one above the planet to the far reaches of the cosmos, to the heavens. One can see the inner parts of one's body and the deeper strata of one's soul, one can



encounter the infinite richness of myth and fantasy, meet fairies and dragons, angels and devils, taste the nectars of the Eternal, be washed by the bounty of the Supreme Good, witness the perennial light, encounter the Divine.<sup>15</sup>

At higher doses, ego, agency, memory, body, space, and time are stripped away. The loss of everything that tethers the mind to the world is, quite frankly, terrifying, but leaves consciousness free to enter a state known as *pure experience*.<sup>16</sup>

The Good Friday Experiment, conducted in 1962 in Marsh Chapel at Boston University with twenty graduate divinity students supervised by Timothy Leary and the Harvard Psilocybin Project, vividly demonstrated that in their felt intensity and character, psilocybin-induced experiences share commonalities with religious experiences. Participants reported ego loss, being in an eternal now, feeling a sense of beauty mixed with terror, seeing the world with new eyes, and encountering celestial or demonic beings, including God. In a follow-up study twenty-five years later, most of the original participants characterized their experience as a spiritual high point of their lives.<sup>17</sup>

Traditional psychedelics are associated with the peoples of the New World. The Old may have once known them too. There is credible historical and spotty archaeological evidence that the Eleusinian mysteries in classical Greece involved psychedelics.<sup>18</sup> With the coming of Christianity, such practices and the associated knowledge were violently suppressed. Whenever Europeans, in their colonization of the Americas in the sixteenth and seventeenth centuries, encountered ceremonies involving substances that gave users visions of the supernatural realm, such as the Aztecs consuming “sacred mushrooms,” known to them as the “flesh of gods,” the colonial powers bloodily crushed these rituals. The eradication of the “devil’s work” by Christian missionaries continued through the closing decades of the twentieth century.

As a lapsed Catholic, I find this profoundly tragic as well as ironic, as I was raised to believe that during the sacrament of the Eucharist in Holy Mass, the consecrated bread and wine become the body and blood of Jesus Christ. After eating the host and sipping the wine, I dutifully returned to my



pew and kneeled in quiet contemplation. But not once, not a single time in the more than one thousand masses I served in as an altar boy or attended as an adult parishioner, was I rewarded with a revelation of heaven or hell, a burning bush, or a booming voice from the sky.<sup>19</sup> I know that, per Catholic teaching, this is not the point of partaking of the Eucharist. But I had an implicit expectation that taking part in mass, Sunday after Sunday, year after year, would eventually be rewarded with some sign. Ah well...

I recently participated in several communal Santo Daime ceremonies in Bahia, Brazil.<sup>20</sup> These featured hours and hours of praying, meditating, drumming, chanting, dancing, and drinking ayahuasca in the forest and on the beach, under a star-studded southern sky. Guided by our *padrinho* Paulo Roberto Silva e Souza, the godfather or shaman who leads the ceremony, I experienced something wonderful—Mind at Large. This has led me to question the metaphysical worldview I had grown comfortable with. Perhaps everything is but a manifestation of the mental? In my struggles to make sense of this ontological shock, the wobbling of my model of reality, I came upon the striking Schopenhauer quote at the beginning of this chapter.<sup>21</sup> The word *daime* in the Santo Daime tradition comes from the Portuguese for “give me.” Indeed, this ceremony has truly given me something I have searched for all my life.

Psychonauts, people who explore these altered states, insist that they have accessed a “higher” form of consciousness, that these substances “expand” or “blow” the mind. This seems to be more than just a figure of speech as it is such a commonly expressed sentiment. There are at least three ways to interpret this claim.

First, psychonauts consistently report that their experiences leave them feeling more curious about the world, more connected, more joyful, less conscious of self. Such psychologically defined “open” states can be contrasted with “closed” states, such as anger, distrust, frustration, anxiety, obsession with perceived inadequacies and insults, with inner or outer conflict. It may simply be that people intuitively associate these open states with an expanded mind and closed states with the everyday selfish mind. Psychiatrist Judson Brewer, at the time at the Yale University School of Medicine, discovered that these “open” or “closed” states of consciousness map onto activities of the posterior cingulate cortex and precuneus complex, part of the neocortical regions engaged when ruminating,

introspecting, and daydreaming. Anger and anxiety, paradigmatic “closed” states, are associated with high activity in these regions; this activity is reduced by mindfulness training that minimizes excessive self-scrutiny by focusing on the here and now. Put differently, the neural signature of a preoccupied and worried self, the substrate of closed experiences, is activity in the posterior cingulate cortex and precuneus complex. When this activity is at an ebb, the self is disengaged, and the conscious mind is open to the world at large.<sup>22</sup>

Second, it is possible that cognitive processes, such as selective attention, imagination, and short-term memory, improve or speed up under psychedelics or that the bandwidth of consciousness increases during extraordinary experiences. This is the gist of the reducing valve theory of Aldous Huxley’s *The Doors of Perception*: the mind is receiving far more sensory input than it can handle and so must filter out and prioritize. The valve limits how much information enters conscious awareness. Huxley argued that mescaline opens or disables the valve, letting a flood of previously unperceived sensory information enter the mind. Objective measures of performance while under the influence of psychedelics in laboratory settings have not revealed any enhanced cognitive processing capabilities. The one possible exception is free association, a basic building block of scientific, technical, and artistic creativity. That’s why some creative types regularly consume small doses of psychedelics, *microdosing*, in the hope of boosting their creativity.<sup>23</sup>

Third, and most intriguingly, an expanded mind can be interpreted in the quantitative sense of integrated information theory; that is, the amount of integrated information,  $\Phi$ , may be measurably higher during a psychedelic experience compared to a more “normal” experience (such as sitting quietly, with eyes closed). This might be surprising, but the number of distinctions and relations associated with the highly salient sense of self that looms large in our ego-fetishizing minds, and therefore draws our attention, is minute compared to the vastness of seen, heard, and felt perceptual spaces. When the self is silenced, these unmeasured realms dominate experience. In principle, the hypothesis of a larger  $\Phi$  can be empirically tested.

## Dying of the Light

Finally, there is a unique class of experiences encountered during life-threatening episodes of near-drowning, cardiac arrest, shock, or blunt-force trauma. The young Ernest Hemingway, after being gravely injured on a World War I battlefield, wrote home that “dying is a very simple thing. I’ve looked at death, and really, I know. If I should have died, it would have been very easy for me. Quite the easiest thing I ever did.” Years later, Hemingway crafted “The Snows of Kilimanjaro,” a tale of a safari gone awry. The protagonist is dying from gangrene when his pain suddenly disappears and a guide named Compie appears out of nowhere to fly him through the darkness into the light of terminal clarity: “Then they began to climb and they were going to the East it seemed, and then it darkened and they were in a storm, the rain so thick it seemed like flying through a waterfall, and then they were out and Compie turned his head and grinned and pointed and there, ahead, all he could see, as wide as all the world, great, high, and unbelievably white in the sun, was the square top of Kilimanjaro. And then he knew that there was where he was going.”

Survivors of near-death experiences return with vivid accounts of seeing a bright light at the end of a tunnel or a vast, luminous expanse, limitless, boundless; leaving their body and floating above it or even traveling into outer space; becoming pain-free; meeting family or spiritual beings; reviewing their life; or experiencing a distorted sense of time and space. Near-death experiences can be suffused with feelings of bountiful bliss or sheer terror. It is the former that attract all the attention: loss of ego, peacefulness, facing the divine. Yet other near-death experiences are frightening, hellish, marked by intense anguish, loneliness, and despair.

A close brush with death, narrowly avoiding a high-speed collision or a fall on a climb, reminds us of the fragility of life. Memories of such precarious events quickly fade. Not so near-death experiences. They are recalled with unusual lucidity and without diminution over decades. In some, they set off massive changes in behavior and outlook, such as loss of the distress and fear common in terminal patients.

Near-death experiences have been reported in all cultures and by all sorts of people, young and old, men and women, devout and secular. Whereas these accounts were once dismissed as feverish hallucinations,

deathbed visions, or confabulations, a handful of physicians and psychologists took what survivors of calamities told them at face value and noted regularities and patterns.<sup>24</sup> Their common physiological trigger is impaired oxygen supply (*hypoxia*) and/or reduced blood flow (*ischemia*) to the brain. Why survivors should return with these unusual tales remains mysterious.<sup>25</sup>

Modern emergency medicine can resuscitate some victims of cardiac arrest.<sup>26</sup> It is expected that many of these survivors will carry psychological scars from their trauma, memories of agonizing pain, of consciousness lost and regained, expressed in high levels of anxiety, flashbacks, and depression. So it is surprising, given the harrowing circumstances, that between 10 and 20 percent report the opposite: positive, highly meaningful experiences.<sup>27</sup>

To those raised in a religious tradition, the most obvious explanation is that they were granted a vision of the hereafter and that, therefore, their near-death experiences are positive proof of an afterlife. Science cannot disprove such claims. However, it is noteworthy that survivors' visions of heaven or hell are appropriate to their religious upbringing and cultural setting. A Roman Catholic will experience a different God than a Southern Baptist or a Jew or a Buddhist. This hardly supports claims of a universal God reigning in one heaven.

There are persistent reports in the popular media that patients who undergo near-death experiences have "flat-lined," with their electroencephalograms (EEGs) showing no significant electrical activity. A completely inert brain would indicate a brain in a deep coma, a brain that had lost its causal powers and that certainly wouldn't feel anything. Therefore, according to this narrative, "conventional medicine" (voiced with dripping sarcasm) fails to explain their visions. Instead, we just have to accept that these patients have been granted a vision of the Afterlife.

I am extremely skeptical, as I've never seen a patient with a flat, isoelectric EEG soon thereafter wake up and claim to have been conscious. A key difficulty is aligning the clock time of the near-death experience with the EEG record, given the patient's confused recollection hours later, often under sedatives.<sup>28</sup> Remember what I called the neuroscientist's dictum, "No brain, never mind"? Neuroscience operates under the hypothesis that all thoughts, memories, percepts, and experiences are an ineluctable

consequence of the natural extrinsic and intrinsic causal powers of the brain—not of supernatural ones. Unless there is extraordinary compelling evidence to the contrary, I see no reason to abandon the principle that a silenced brain does not support consciousness.

Phenomenologically, near-death experiences resemble powerful psychedelic ones. The perceived body and the self can die in both. The difference is, of course, that psychedelic-induced near-death experiences are perfectly safe from a physiological point of view (except for a minor increase in blood pressure and heart rate). But the body-less, self-less mind does not know this.<sup>29</sup>

My own psychic death came about with the help of a mighty molecule, 5-MeO-DMT, rather than any physical or medical trauma. This entheogen is also known as the “toad,” as the substance is derived from the glands of the Colorado River toad. In its awful intensity, awful in the original meaning of “worthy of respect or fear, striking with awe,” this experience exceeds anything else in my life.

I was sitting cross-legged on a carpet in the presence of a guide. She started the music I had selected, Arvo Pärt’s minimalist “Spiegel im Spiegel,” with a lone violin accompanying a piano. I deeply inhaled a vaporized dose of 5-MeO-DMT: once, twice, and thrice, at which point my entire field of view became obscured by dark, densely swirling smoke; space fractured into a thousand black hexagons and shattered. 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.

As I opened this book with my experience of a timeless universe convulsed to a point of radiant energy and a profound feeling of terror and ecstasy, there is no need to repeat that description here.

The first sign of the outside intruding into my marooned mind were the closing notes of “Spiegel im Spiegel.” During the previous nine minutes, I had been sitting upright and immobile, silent, blind, with eyes wide open, staring dead ahead. I screamed and heard but a faint voice emanating from my chest. My mind gradually returned to my body. I stripped off my clothes as they felt restrictive, instinctually curled into a protective fetal position, and cried. It took about an hour for me to return to my self. I slept soundly that night. Except for an emotional flashback the next day, there were no other immediate effects.

Still, my interior life took several years to recover from this adamantine experience. For reasons I do not fully comprehend, the survival of subjectivity in the face of ego death leached away the hold my eventual oblivion had over my psyche. My obsessive thoughts about the worm at the core of existence are completely gone. Left in their wake is a tranquil acceptance of death, an imperturbability, free from distress, anxiety, and worry. Yet like the protagonist in Jorge Luis Borges's "The Zahir," I find myself returning more and more frequently to contemplate the terrible and beautiful nature of the naked singularity I beheld, listening at the shore of a great silence.<sup>30</sup>

### [A Common Neurobiological Substrate](#)

Experiences that fundamentally change the way we see the world provide clues about consciousness. I learned that the experience of self is optional. Even the feeling of having a body is not necessary for subjectivity. Transformative experiences are deeply personal and can profoundly impact lives. They are living proof that nervous tissue, under special circumstances, can host extraordinary experiences.

All three classes of transformative experiences that I described—religious or mystical, psychedelic-induced, and near-death experiences—probably have a common underlying neurobiological mechanism.

One possibility is a lull of neural activity in the posterior hot zone, especially in the visual, auditory, and somatosensory cortices, posterior cingulate, and precuneus cortices. Such a "quiet" neocortex would be a highly unusual state that wouldn't arise during regular waking life, when consciousness is occupied either by events in the outside or by internal thoughts, ruminations concerning the past, and plans for future actions. A state of minimal neuronal activity is compatible with the experience of a vast, empty expanse, no passage of time, no narrative or core self, and the dissolution of any distinction between the experiencer and the experienced.<sup>31</sup>

It is not that difficult to imagine how such a circumstance could come about during hypoxia. It is more challenging to explain how a mystical or psychedelic experience entails a quiet neocortex. Melanie Boly, a neurologist and neuroscientist at the Medical School of the University of

Wisconsin, Madison, is painstakingly collecting EEG data from long-term Buddhist meditators during a state known as *pure presence*, an experience with no self, no discursive thoughts, and no perceptual content except for a luminous expanse, an empty mirror. Attaining and maintaining this state requires intense practice. It is characterized by a simultaneous nadir in EEG power in both the low-frequency theta band (that is, meditators are not asleep) and in the high-frequency gamma band, compatible with no thoughts and minimal neural activity. This reduction is especially pronounced at the back of the head, in the posterior hot zone: the brain is as quiet as the mind is calm, unperturbed.<sup>32</sup>

According to integrated information theory, a silent posterior hot zone is the substrate of an experience of boundless space without body, without self, and without time. This is completely different from a situation in which the *causal powers* of the posterior hot zone were to be severely curtailed—say, by injecting a local anesthetic—preventing neuronal activity and leading to a silenced cortex with no integrated information. From the point of view of an external observer, a *silent* cortex and a *silenced* cortex resemble each other, as neither lights up with electrical activity. Yet, while a silent cortex retains its full causal power but chooses not to speak, a silenced cortex has lost its voice and is unconscious. This situation recalls Arthur Conan Doyle’s famous short story “Silver Blaze,” in which Sherlock Holmes points out to the clueless Inspector Gregory the curious incident of the dog at night who did not bark. Holmes abduced that the dog could have barked but did not because it knew the perpetrator. A patch of cortex that is intact but is silent is a meaningful situation for the brain, while a paralyzed patch of cortex that cannot be active is not (although, superficially, both have a similar output—none).

This situation reveals the incoherency of the standard information-processing account for consciousness, in which information is transmitted, through a noisy channel, from a sender to a receiver. If there is no activity, no message is broadcast, and therefore, it is argued, there is no experience. Yet this is the perspective of an outsider. There is no homunculus residing in the brain that receives information from someone else. From the intrinsic-powers perspective, only states that make a difference to the brain itself matter. From this point of view, the situation is clear: a silent brain has intrinsic causal powers that are unfolded into a vast causal structure of



irreducible cause-effect power, while a silenced brain possesses none.<sup>33</sup>

There are three reasons why extraordinary experiences are so tantalizing. First, such experiences challenge the conventional physicalist view that only space, time, matter, and energy have true ontological status, that only they truly exist. For all of these can vanish, yet experience persists. Second, they reveal hidden recesses of the mind. They let the psychonaut travel to places inaccessible during ordinary life. Visiting these vistas is not so much pleasurable as revelatory. Third, they can unlock transformations that have therapeutic value, relieve the symptoms of depression, posttraumatic stress, anxiety, suicidal ideation, or existential terror in the face of death. I turn next to these transformative aspects.



## CHAPTER EIGHT

---

### transforming lives by transformative experiences

The world is awash in suffering of every variety: horrendous and intentional violence directed at individuals and groups that leaves deep scars in body and mind; emotional, physical, or sexual abuse that burrows deeply into the psyche and haunts its victims throughout life; addiction to drugs or alcohol; bone-chilling despondency and despair; racism, misogyny, and the legacies of colonialism that persist like an oppressive miasma; and on and on. Even if, by good fortune, we have managed to avoid these, a plague of self-imposed suffering takes its toll: belittling and self-blaming, low self-esteem, catastrophizing, excessive rumination about the past, assuming the worst about others, self-deceit, bitterness, cynicism, and fear of death.

Misfortune is unavoidable, but the suffering that attends it is a product of the mind and therefore, to an extent, optional. I owe many things to my dad—one of them is a deep appreciation for the Greco-Roman school of ethics known as Stoicism. A supremely rational philosophy older than Christianity, it rejected superstitions and belief in gods, preferring natural explanations of the world. It also sought to cultivate a detached attitude to events, good or bad. So while you have limited control over what happens—you live at the wrong time (e.g., during a Covid-19 pandemic) or in the wrong place (e.g., in Ukraine during the Russian assault); your child disappears into the hell hole of drug addiction; you have been diagnosed with melanoma—you have the potential to control how you respond to these events, how you interpret and judge them. This calls for the lifelong cultivation of fortitude and equanimity, what the ancient Greeks called *ataraxia*, an imperturbability and freedom from distress, anxiety, and worry. Stoics sought to attain this suspension of judgment as one of the ultimate aims of life. The book *Meditations*, by the second-century CE Roman emperor Marcus Aurelius, encapsulates this teaching. When I was a

teenager, my father gave me a copy, to which I return time and again. Aurelius expresses the supreme need for mental self-control: “If thou art pained by any external thing, it is not this that disturbs thee, but thy own judgment about it. And it is in thy power to wipe out this judgment now.”

Written as a diary not meant for posterity, *Meditations* is a profound source of wisdom that speaks to the modern world and its concerns. Of direct relevance is its key insight that our mind is not a passive recipient of percepts, emotions, and thoughts but the active shaper of these. That is, the same event can be experienced in different ways, some conducive to our long-term well-being and some not. It is up to us which ones we choose.

A popular instance of reframing unpleasantness is the adage “No pain, no gain.” That is, what is experienced as negative should be seen as positive. Such a change in interpretation need not be limited to exercise. More serious forms of suffering can be recast as building character and growing spiritually.

Recalibration can be greatly accelerated by transformative experiences. People who have undergone religious, mystical, aesthetic, or near-death experiences are marked by them. Because their experience is so compelling, and sometimes blissful beyond words, it forces the experiencer to reevaluate their life and their beliefs about what is worth striving for.

Extraordinary experiences are transformative because, escaping the ever-present gravitational pull of the self, they reenchant the world. With the ego, the driver of goal-directed actions that always looks out for itself, out of the way, the pursuit of material goods and rewards is revealed for what it is: a game for children, as Siddhartha, in Hermann Hesse’s eponymous novel about the life of Gautama Buddha, puts it. We can develop empathy with all conscious beings and, indeed, with all creation. We understand that what truly matters is gratitude for existence and compassion for all suffering creatures. This compassion extends to ourselves and our guilt or shame—following the adage that forgiveness is giving up hope of a better past.<sup>1</sup> We can cultivate new habits that enable us to be more compassionate and more at peace with the world and ourselves.

That the attitude we bring to life influences how our body and our mind react to external events is demonstrated by the robust strength of the placebo response and its dark twin, the nocebo response.

Historically, transformative experiences strike out of the blue, as a

singular act of grace. They either seemingly happen by chance to someone in a particular but unknown state of mind or during a near-death encounter. Thus, a first step in harnessing transformative experiences to change lives is to make them reliable and safe. This is where psychedelics come in.

### [Psychedelics and Their Potential to Change Lives](#)

Artists, writers, and scientists have always been fascinated by psychoactive substances. In the nineteenth century, that was opium, hashish, nitrous oxide (also laughing gas), diethyl ether, and cocaine, before the addictive properties of these drugs were recognized.<sup>2</sup> The industrial West encountered psychedelics when two chains of events converged in the middle of the twentieth century. One was the synthesis of a powerful mind-bending substance, lysergic acid diethylamide (LSD), by the Swiss chemist Albert Hofmann, employed by Sandoz Laboratories in the staid city of Basel. On April 19, 1943, the famed “bicycle day,” Hofmann underwent the world’s first acid trip on his bike ride home. The second event was the participation of Gordon Wasson, an American banker and hobby ethnomycologist, in a mushroom ritual in a small Mexican village in Oaxaca, guided by the healer Maria Sabina. When *Life* magazine published a photo essay of Wasson’s “divine” journey, titled “Seeking the Magic Mushroom” in 1957, the article went viral. It ultimately led to the widespread public recognition that traditional societies used psychoactive substances in ritual settings to induce visions, heal the disturbed, and let the troubled find peace.

Psychedelics were given intellectual respectability by Aldous Huxley’s *The Doors of Perception*, detailing his experience with mescaline, a psychoactive substance derived from the peyote cactus and traditionally used by Native Americans. His slim book advocated using mescaline to facilitate mystical insights for scientific, artistic, and religious reasons. Together with cheap and readily available LSD, the book ushered in the Age of Aquarius, Timothy Leary, and his siren song to 1960s youth to “Turn on, tune in, and drop out.” Powered by indiscriminate acid use and abuse, the countercultural movement rejected bourgeois values, encouraged “free love,” and opposed the US war in Vietnam.<sup>3</sup>

For every action, there is an equal and opposite reaction. This is true not only in physics but also in society. Thus, predictably, the US Congress

passed the Controlled Substances Act, signed by President Richard Nixon in 1970. It placed all psychedelics into the most restrictive and heavily regulated category of drugs, called Schedule I drugs, those with high potential for abuse and no accepted medical use. Other countries swiftly followed. All scientific and clinical investigations ceased.<sup>4</sup>

Research carried out, under difficult conditions, in the intervening years by a few intrepid scientists<sup>5</sup> demonstrated what had already been known in the 1960s: that psychedelics, taken in a controlled environment, are comparatively safe and non-habit forming.<sup>6</sup> This trickle of experiments turned into a flood as the early twenty-first century witnessed a renaissance of psychedelic research, based on the recognition that these substances can ameliorate or even heal a wide range of psychiatric disorders for which conventional drugs are of questionable efficiency but entail undeniable side effects and lifelong dependency. Despite all of that, the possession, distribution, and consumption of psychedelics continues to be a felony in almost the entire world. Exceptions for scientific and medical research are granted, after a lot of paperwork, by the appropriate regulatory agencies.

Many grew up, like me, in an environment where any “drug usage” was considered beyond the pale and believed to rapidly lead to addiction, despair, madness, and worse. Some may recall images of eggs in a frying pan symbolizing “your brain on drugs” from a late-1980s antidrug campaign. This lumping of any mind-altering substance into a single category of “drugs” is simplistic and not helpful. It is critical to distinguish psychedelics from the highly addictive and deadly drugs that flood our cities and streets, such as cocaine, crack, heroin, oxycodone, fentanyl, and xylazine, which killed more than one hundred thousand Americans in 2021. Opioids are in a different chemical class from psychedelics, targeting different mechanisms in the brain. Each class of molecules needs to be treated on its own terms, as we do for the three most widely consumed psychoactive agents: ethanol, nicotine, and caffeine. Furthermore, consuming psychedelics in a controlled clinical or shamanistic setting with an intent to heal is radically different from recreational use by individuals on their own.

Psychedelics are back, with mega-conferences, a Netflix miniseries, and ballot initiatives to decriminalize their possession and establish psychedelic-assisted therapy in individual states, such as Colorado and

Oregon. And what *The Doors of Perception* was to the 1960s, *How to Change Your Mind* by the journalist Michael Pollan is to the 2020s, bringing the healing potential of these molecules to the wider public.<sup>7</sup>

Best studied is psilocybin, the active ingredient in magic mushrooms, a large and diverse group of fungi. Psilocybin's effects can be remarkable—one or two consecutive psilocybin-assisted therapy sessions in patients with serious depression improved their well-being and depression scores for months.<sup>8</sup> A *single* high dose of psilocybin in patients with life-threatening cancer triggers large decreases in depression, anxiety, and mood disturbance, with concomitant increases in quality of life and ability to find meaning in life and acceptance of death. Furthermore, patients who have full-blown mystical experiences are more likely to enjoy these benefits than those who do not. Michael Pollan evocatively described, in a famous *New Yorker* essay, how even for those patients who did succumb to their cancer, the treatment had an extraordinarily positive outcome for their way of living and of dying.<sup>9</sup> This is astounding and defies belief: well-entrenched attitudes and beliefs are so difficult to modify that conventional psychotherapy, aka the talking cure, takes years of weekly sessions with uncertain outcomes. The results are equally astonishing next to the results of more conventional psychopharmacology. One study directly compared two doses of psilocybin over six weeks to a daily pill of escitalopram, a common selective serotonin reuptake inhibitor treatment for depression. It found that patients taking the psychedelics had reduced ruminations, “sticky” negative thoughts about their self and their situation, and thought suppression, compared to their peers given the standard treatment.<sup>10</sup>

Two not-for-profit organizations, the Multidisciplinary Association for Psychedelic Studies (MAPS) in California, now renamed Lykos Therapeutics, and the Usona Institute in Wisconsin, labor tirelessly to obtain regulatory approval from the Food and Drug Administration (FDA) and the Drug Enforcement Administration and to take these substances to the clinic. MAPS is focused on psychotherapy assisted by 3,4-Methylenedioxymethamphetamine (MDMA, aka ecstasy, a nontraditional psychedelic) for severe posttraumatic stress disorder and Usona on psilocybin-assisted psychotherapy for major depressive disorders.

An entire ecosystem of companies has sprung up to profit from the potentially massive market for psychedelic therapy. This goldrush

mentality, combined with a lot of hype, will surely come to haunt the field.<sup>11</sup> The classic psychedelics are either not patentable, as they are naturally occurring substances, or are in the public domain. Thus, medicinal chemists at start-up companies tweak existing psychedelic molecules to engineer novel compounds that can be legally protected.<sup>12</sup> They tinker to shorten the duration of the acute effects of these substances, as having two therapists supervise a single patient for four to six hours, the duration of a typical psilocybin or ecstasy trip, is costly and limits how widely these medicines can be deployed within the existing health-care system. Shortening psychedelic trips seems feasible as some entheogens have an almost immediate onset and an acute phase measured in tens of minutes.

As of 2023, around two dozen active clinical studies for MDMA and seventy-five for psilocybin had been registered to treat concussions, migraine and chronic cluster headaches, major depressive disorder, treatment-resistant depression, obsessive-compulsive disorder, mild cognitive impairment, posttraumatic stress disorder, chronic pain, end-of-life existential distress, alcoholism, smoking, addiction, and caregiver burnout.<sup>13</sup>

In the next few years, the FDA and its European counterparts will likely approve MDMA-assisted therapy for posttraumatic stress disorder and psilocybin-assisted therapy for major depressive disorder. These substances will then be legally available at pharmacies with an appropriate prescription in conjunction with therapy. Furthermore, your doctor can then prescribe these drugs for other conditions, at their discretion, so-called off-label use.

What about the risks? No intervention is without them, but compared to many common psychoactive substances, in particular alcohol, psychedelics are far less risky.<sup>14</sup> Most people trip a few times during high school or college, enjoy the sights, and then stop. One study estimates thirty million people in the United States, about one-tenth of the population, have used these molecules at least once.<sup>15</sup> Their physiological toxicity is low, and they do not damage organs or cause neuropsychological deficits. The risk of lethal overdosing or fatal accidents is small.<sup>16</sup> Importantly, psychedelics are nonaddictive (they do not target the dopamine system) and engender neither cravings nor compulsive drug seeking. There are no drug cartels trafficking in psychedelics.

Psychedelics affect consciousness in powerful ways yet without

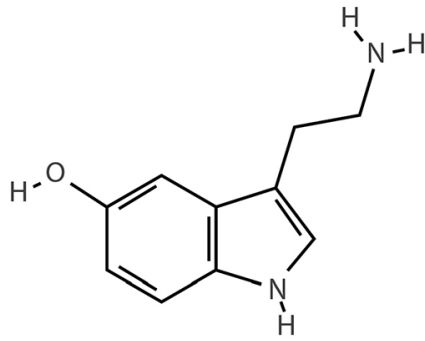


impairing motor function. Unlike with alcohol intoxication, users do not stumble about or slur their speech. Another notable difference is the lack of any hangover the next day; indeed, one usually feels remarkably buoyant. If there is a price to pay, it is upfront, as oral ingestion of psychedelics can make users nauseous and trigger purging.<sup>17</sup>

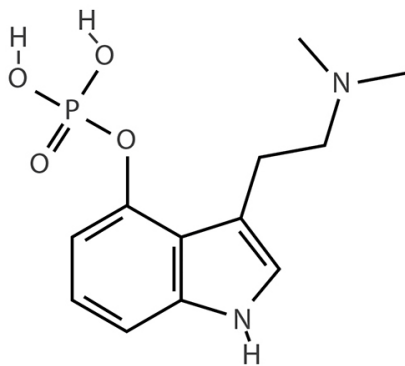
For most, the primary risk is a “bad” trip, reliving trauma and pain, feeling existential despair, or experiencing paranoia. The chances of such a nightmarish experience can be minimized, but not eliminated, by an appropriate “set and setting.” That is, do not take psychedelics when you are not in the proper mindset—say, following a fight with your partner—or when you are in an unsafe or inappropriate physical or social setting.

Some report recurrent visual hallucinations or perceptual distortions following psychedelic use. More serious, but rarer, are prolonged psychoses in a small fraction of teenagers and tweens prone to schizophrenia or feelings of derealization, of living in a fake “reality,” that can last for days. Safeguards include only giving these molecules to adults, screening for personal and family history of psychosis, and having a therapist always present before, during, and after each psychedelic session.<sup>18</sup>

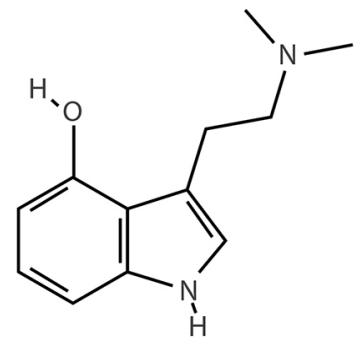
The last point bears emphasizing for people wanting to try psychedelics at home or in a bucolic setting (say, in a forest or next to the sea) on a long weekend. To minimize acute risk, a trusted and sober friend or loved one, a “trip sitter” in the lingo, should be with them, to assure them that no matter what they see, hear, or feel, this shall pass, and they will eventually be fine. This is especially true for first-time users. Given that they can be granted visions of heaven, hell, and everything in between, it is important that they can make sense of their experiences, integrating them into their life’s narrative, by speaking about them in a nonjudgmental manner. Be mindful, though, that psychedelics and entheogens are sensitive amplifiers of repressed thoughts, fears, and desires that can overwhelm the user. The way to deal with such negative feelings is not to fight them but simply let them pass (which is easier said than done).<sup>19</sup>



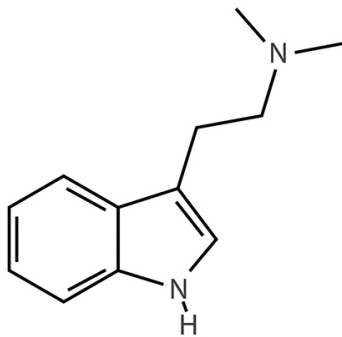
**Serotonin**



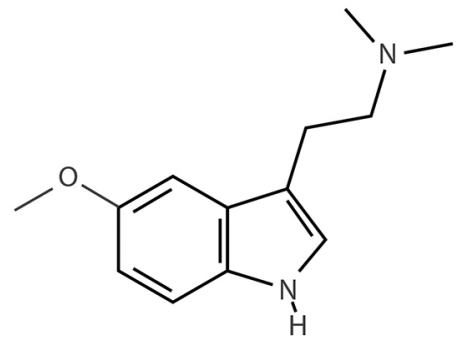
**Psilocybin**



**Psilocin**



**DMT**



**5-MeO-DMT**

Figure 2. The structure of the neurotransmitter serotonin responsible for modulating mood, cognition, memory, vasoconstriction, and gut mobility in the body and four closely related naturally occurring psychedelic molecules with remarkably similar structures.



Chemically, psychedelics are structurally closely related alkaloids belonging to the class of tryptamines, or serotonin analogs (see [Figure 2](#)). Serotonin, named for its ability to constrict blood vessels, is found in the gut and in the brain, where it acts as a neurotransmitter, binding to seven distinct families of receptors. These serotonin, or 5-HT, receptors are embedded in the membranes of neurons in the neocortex, the claustrum, and a few subcortical nuclei, such as the dorsal raphe, in complex patterns. Psychedelics primarily bind, with different affinities, to serotonin 2A, 2C, and 1A receptors, causing a cascade of follow-up intracellular events. Binding to the 5-HT-2A receptor subtype is necessary to evoke hallucinations, mystical experiences, and therapeutic benefits. That is the primary reason why most research focuses on this receptor.<sup>20</sup>

Neuroscientists have recorded the brain activity of mice and human volunteers on psychedelics. One common observation is that psychedelics destabilize long-range cortical communication patterns and reduce activity in the posterior cingulate cortex and the precuneus in the posterior regions of the neocortex. This is compatible with our knowledge of the brains of people trained in mindfulness. It appears that the less these midline structures are active and/or the less they communicate with other neocortical structures, the less the self is present. This would explain how the borders between the self and the world become abolished, leading to a feeling of oceanic boundlessness or cosmic unity. Another consistent finding is a slight increase in the complexity of EEG and MEG signals during the psychedelic experience. It is not clear whether this is a consequence of enhanced causal interactions within the neocortex or more chaotic activity. Some cortical neurons, in both mice and humans, respond to psychedelics with an increase in excitability and some with a decrease; most nerve cells do not respond at all. It's early days yet, and we still understand little about the strikingly complex cellular mechanisms. Stay tuned.<sup>21</sup>

A shortcoming of basic research studies of psychedelics is that most are done in mice and rats. These animals can only tell us indirectly—say, by pressing a lever—if they see a bright light at the end of a tunnel or experience a loss of self (assuming rodents even have one). It is challenging to link brain activity in an animal to subjective experiences without such

indications.<sup>22</sup> Rodents may well experience visual hallucinations, as reported in monkeys, but as far as I know, this has not been tested experimentally.<sup>23</sup>

The therapeutic benefits of psychedelics are thought to depend on a two-stage process. The first includes the acute mind-altering conscious experiences that give psychedelics their name. This opens a window of enhanced neuroplasticity that constitutes the second, chronic phase. During this “integration” period, lasting for weeks or longer, the psychedelic-induced malleability of the brain’s synaptic wiring and the attendant enhanced mental flexibility let users avoid getting stuck in the same old, contracted worldview, enabling them to explore more productive thought patterns, attitudes, and habits with the aid of a therapist, expanding their view of the world and their place in it.<sup>24</sup>

A point of contention among researchers is whether the psychedelic experience itself is truly necessary for therapeutic benefit. That is, do you need to experience visual hallucinations, distorted body schema, ego loss, and so on (features typically evaluated via extensive questionnaires) to be healed?<sup>25</sup> A dyed-in-the-wool physicalist would argue that all that needs to happen is for the right molecules to bind to the right receptors, causing a signaling cascade that induces neuroplasticity. Consciousness is along for the ride but is causally impotent, a mere epiphenomenon. This means that a psychedelic’s therapeutic effects would also occur if the medicine were given intravenously while volunteers were sleeping or anesthetized; or if the psychedelic were given to volunteers taking a sedative, like midazolam, that prevented the formation of memories, inducing amnesia. They would still trip but wouldn’t remember any of it. Would any of these subjects in these nonstandard settings still benefit?

If the acute, consciousness-expanding phase isn’t needed, this would justify the development of new drugs that don’t cause hallucinations and ego dissolution, both of which are frightening to many, yet still open the neuroplasticity window. Such designed molecules come with the added benefit that they would be patentable.

To me, this quest is a fool’s game. It is known that the likelihood of long-term therapeutic gains increases with the likelihood of the subject having a mystical experience, prima facie evidence favoring a critical role for subjectivity.<sup>26</sup> It is precisely the high emotional impact of the memories

of the vivid and highly unusual psychedelic experience that serves as the primary source of inspiration for sustaining changes in attitude and behavior long after the acute effects have worn off.<sup>27</sup> The belief that the Sturm und Drang of such a singular experience, the reason many use psychedelics in the first place, is meaningless and futile, impuissant in affecting the brain, is cut from the same cloth as the belief that any “mental condition” or “mental disorder” can be treated by popping a pill, without dealing with the experience of the sufferer. The results of such a tactic are decidedly mixed.

Other open questions abound.<sup>28</sup> How likely is it that one or two psychedelic-assisted therapy sessions are sufficient to permanently relieve depression, generalized anxiety, or posttraumatic stress in people who have suffered for years? How often must these sessions be repeated to achieve benefits lasting a lifetime? What are the downsides of chronic use of psychedelics?<sup>29</sup> How important is the rapport between the therapist and the patient? Which psychiatric conditions are truly amenable to psychedelic-assisted therapy and which ones are immune? It is likely that some patients will show a strong benefit while others will not. Can psychological or genetic traits predict who will benefit most? How important is the emotional intensity and valence, say, of a shattering near-death experience versus one of bliss and joy? What difference does the duration of the extraordinary experience make? Does the psychedelic experience boost neuroplasticity everywhere or only in specific regions of the brain? If it is the former, if synapses are sprouting throughout the brain, like mushrooms after a rain, psychedelics could be used to treat not just psychiatric but also neurological conditions, such as mild cognitive impairment, often a precursor of Alzheimer’s disease; post-stroke rehabilitation; and reviving patients with disorders of consciousness.

Time and many more studies will tell.

The breathless reporting in the media concerning psychedelics as medicines builds up expectations: people are being conditioned to believe that they work. This belief will, by itself, lead to positive results. However, rather than exploiting this beneficial belief, the traditional medical model downplays the placebo response and focuses primarily on the action of molecules. This is both a lost opportunity and a lost cause, as it is challenging to truly mask the effect of these substances—if you have ever

previously taken a psychedelic, you know when you've been handed a sham. Clinical trials will have to adapt to the impossibility of true blinding.<sup>30</sup>

### [Expanding Our Perception Box](#)

Transformative experiences can yield another gift. Earlier on, I explained by way of #TheDress the different realities everyone experiences. No one, no matter how wise or intelligent, has privileged access to the one and only “true” reality. Indeed, no one has direct, unmediated access to the noumenal, unknowable reality that Immanuel Kant postulated, the *thing-in-itself*. What we perceive, what we experience, is a construct of the brain, shaped by our implicit and explicit expectations. If we believe something to be true, if it fits into our belief system, we are more likely to notice it and to remember it. If it does not reinforce our prior beliefs, we will ignore the facts of the matter to the extent possible. Just observe the way people justify or condemn political violence, depending on their point of view.

Because we only know our own idiosyncratic view of reality, we take it for granted and assume that everyone experiences the same, although many know, in an abstract way, that our experienced realities differ in ways both small and large. [Chapter 3](#) introduced the evocative Perception Box metaphor, the mental architecture that we all inhabit and that limits us. Each of us is stuck inside our own Perception Box. Its scaffolding is the physical substrate of consciousness, the way the neurons that constitute it are organized and interconnected. The walls of this box are invisible. They are also shatterproof, as we can only experience what our neural circuitry allows us to experience, nothing less and nothing more. Unlike the Berlin Wall, we can't ever “tear down this wall”; we can't escape our Perception Box. Its walls may be constricted, or they may be expanded, by changing the underlying neural connectivity, but there is always a box.

We know (or should know, if we draw upon a diversity of sources across the ideological spectrum) that educated, intelligent, and well-meaning people, even if given the same facts, can differ fundamentally in their interpretation, and therefore their experience, of political events—witness the different narratives around Black Lives Matter, January 6, gun control, and other hot-button topics in contemporary American culture. But this

abstract knowledge is completely overwhelmed by our strongly held views that we hold to be righteous. We “know” that we are right and others are at best misled fools who are being purposefully manipulated. Or we are stuck in a constricted worldview, depressed, anxious, obsessive-compulsive, full of guilt or shame; we assume that nothing can change and resign and accommodate ourselves to this bleak situation; indeed, some identify with their anxieties and other negative traits, which makes them even more resistant to being changed.

However, the good news is that we can rearrange and expand the walls of our Perception Box and seek to be more open by changing the underlying neural substrate. This is the goal of the Tiny Blue Dot Foundation in Santa Monica.<sup>31</sup> Cofounded by Elizabeth R. Koch, it funds neuroscience-based research into interventions to alleviate suffering and improve the well-being and mental health of people everywhere. Expanding our Perception Box is the point of traditional home and school education. This leaves us more open, more curious, and more compassionate toward our fellow travelers on the river of time. We teach children when their brains are most impressionable and effortlessly soak up attitudes, knowledge, and lifelong positive habits that promote physical health and growth. Just as daily flossing and tooth brushing prevents tooth decay, *mental flossing*, such as proper breathing techniques to facilitate a calm, relaxed attitude, or daily meditation or gratitude sessions, can be cultivated and turned into a lifelong practice for mental well-being. While a slow and arduous process, instilling lifelong positive habits works. It does take time, though, and becomes tougher as brains age and become more rigid and less amenable to change, in particular for deeply entrenched attitudes.

By reducing or dissolving the self, transformative experiences offer a seductive alternative, more rapid, more dramatic, and, possibly, more effective than education, at least in the short term. Relevant methods include, but are not limited to, trance, hypnosis, breathwork, yoga, meditation, psychedelics, and noninvasive brain stimulation tools. Time will tell which of these interventions most reliably, consistently, over the longest term, and in the greatest number of people expands their Perception Box.

Transformative experiences serve as a powerful reminder of the miracle of existence, of the fundamental mystery of why there should be something

rather than nothing. They leave behind a feeling of the sublime, a glow that accompanies us in daily life.<sup>32</sup>

Yet that glow may fade with time. When one has come face-to-face with God or merged with the universe, it may be difficult to return to a life that still requires taking out the trash and paying the mortgage. This is why a daily routine, such as meditation or mindfulness, remains essential to maintaining the right attitude throughout one's life.

What about the final transformative experience? What happens when the physical substrate, without which there can be no consciousness, breaks down and becomes inoperable?

## CHAPTER NINE

---

### the end of consciousness

Since I was born  
I have to die  
and so...  
—Kisei (Japanese poet, 1688–1764)

**W**e live on borrowed time. This profoundly influences our mind, both its familiar aspects, rooms that we frequently visit, and also its secret passages, the forgotten haunts and buried chambers we instinctively avoid. In the opening pages of this book, I covered the dawn of consciousness. For every dawn, there is a dusk, for everything that begins must end. Let us see how the stream of consciousness empties itself out into the infinite sea.

### Appointment in Samarra

You will die, sooner or later—no matter how healthily you eat, how much you exercise, how well you sleep, or how many vitamins you take.<sup>1</sup> No one gets out of this life alive. Few like to think about this troubling fact. We dread the thought of being nevermore. How can we live a meaningful life in the face of the end of the reel? Doesn't mortality act like an alkahest, a universal solvent, eradicating all meaning from our attachments to loved ones, to deeply felt notions of love, responsibility, and truth?

Evolution equipped humans with powerful psychological defenses to deal with this foreknowledge: unconscious repression and conscious suppression.<sup>2</sup> We don't perceive or remember what makes us uncomfortable, such as state-sanctioned violence, and look away.<sup>3</sup> Organized religion is a societal-level defense mechanism that promises life without end, whether in a static Christian heaven or an eternal cycle of



Buddhist reincarnation. Of more recent vintage is the aspiration toward digital immortality by uploading our minds to the cloud, the twenty-first-century equivalent of rapture for, and by, nerds.<sup>4</sup>

Death has no such dominion over nonhuman animals. There is no credible evidence that apes, dogs, crows, and bees are sufficiently self-aware to be troubled by the insight that they will die. They live in the perennial now, unlike humans whose every moment is tinged by remembered pasts and anticipated futures. Once birthed, the thought of oblivion cannot be erased. It lurks in the unconscious shadows, ready to burst into the light of consciousness.

Denial kept thoughts of death in abeyance until I was well into middle age when I abruptly woke from a deep sleep to the visceral “insight” that I would eventually cease to exist. I brought myself to a dizzying state of vertigo by imagining an eternity of nonbeing, forever and ever. Subsequently, I avoided gazing into that existential, bottomless abyss; but persistent ruminations on death haunted me over the last decade, late at night, when I couldn’t sleep. Those thoughts are gone, following my transformative experience of losing the sense of the existence of an external world, of a body and of a self. I can now calmly contemplate the end.

A short story by Ray Bradbury, “The Last Night of the World,” epitomizes my attitude well. A husband and wife understand that this night will see the end of the world, not by nuclear hellfire but by a simple cessation of existence. Without any drama, they put their two girls to bed, quietly speak to each other, turn off the lights, kiss good night, and go to sleep—a remarkably understated *finis* and the very opposite of Dylan Thomas’s “Rage, rage against the dying of the light.” I hope to retain such equanimity in my final moments of clarity.

Wisdom teaching from religious and philosophical traditions throughout the ages has long emphasized that daily meditation on mortality removes its sting. Journalist Andrew Sullivan reflects, “There is a recognition that beyond mere doing, there is also being; that at the end of life, there is also the great silence of death with which we must eventually make our peace.”<sup>5</sup>

My own musings on finitude inevitably turn toward understanding life and death from a scientific and clinical point of view. Here, one quickly enters ill-defined territory.

## Modern Death

Throughout history everyone knew when the grim reaper had entered the room: the heart of his victim stopped beating, and the lungs stopped breathing. These two hallmarks of what is known as *cardiopulmonary death* can be avoided following the invention in the last century of pacemakers and ventilators. Now patients can be kept “alive” practically indefinitely even though they can’t breathe on their own and will never regain consciousness.

A committee of notable Harvard Medical School faculty introduced the concept of death as irreversible coma, that is, irreversible loss of all brain function, in 1968. Their recommendations were adopted in the 1981 Uniform Determination of Death Act (UDDA): “An individual who has sustained either (1) irreversible cessation of circulatory and respiratory functions, or (2) irreversible cessation of all functions of the entire brain, including the brain stem, is dead.”

Something like the UDDA has been implemented across the United States and the world, albeit with many local variations.<sup>6</sup> *Brain death*, also known as death by neurological criteria, requires that the brain be in a coma with a known cause, without brainstem reflexes and without an ability to cause the body to breathe on its own (apnea). The UDDA does not actually stipulate how brain death is to be established. Those standards are promulgated by various professional medical organizations, dictated by advances in the clinical sciences.<sup>7</sup>

These standards become critical in the context of the “dead donor rule,” which stipulates that organ procurement can only occur if the patient is legally dead. Once permission has been obtained from next of kin, the heart, kidneys, liver, or lungs of a brain-dead “beating-heart cadaver” can be removed to help living folks who need these organs. There are more than one hundred thousand patients on a national list waiting for such life-supporting organs. That a brain-dead body on a ventilator is a potential organ donor is an uncomfortable fact in end-of-life debates, the elephant in the room.<sup>8</sup>

It is remarkable how, within the span of two decades, the millennia-old dogma and practice of death were revised and, by and large, accepted by the public. This stands in marked contrast to the ongoing controversy

concerning abortion. Curiously, this asymmetry in the beginning and the end of life, book-ended on both sides by eternities, is the exact opposite of the existential anxiety in people's mind about the continuity of the self after death but rarely about the existence of the self before birth.<sup>9</sup>

Notwithstanding, most people continue to die by the traditional cardiopulmonary route. And even if generally we have accepted these revised notions of death, it is a jarring experience to be told that a loved one, on life support, their heart beating, their chest moving in and out, appearing healthier than many other denizens of the intensive care unit (ICU), is legally a corpse.<sup>10</sup>

Given today's widespread use of painkillers and sedatives, most of us die with a dulled mind, a marked difference to recorded history. This may be a blessing or a curse.

There are further complications. The criteria for establishing brain death leave a lot of leeway in clinical practice, such as what kind of doctor and how many must sign off on the declaration-of-death certificate; how many brainstem reflexes must be evaluated (e.g., testing for the oculovestibular reflex by pouring cold water into one ear); whether such testing must be repeated; and whether next of kin can lodge religious objections against the entire procedure. Different states and hospital systems have different standards, leading to discordance, such that a patient declared dead in New York may be considered alive, across the Hudson River, in New Jersey.

Then there is the UDDA demand that all functions of the entire brain be lost. That may not always be the case. The hypothalamus controls the nearby pituitary gland, which releases hormones responsible for growth, pregnancy and birth, breastfeeding, blood pressure, and a host of other regulatory functions. Sometimes these functions can persist even when the rest of the brain has stopped. The mindless body, with proper support, can continue to grow, to menstruate, and to fight off infections with an active immune system. There are more than thirty known cases of pregnant "brain-dead" mothers placed on life support to gestate a surviving fetus born weeks or months later. A young girl, Jahi McMath, was maintained on ventilation in a home-care setting in New Jersey by her family following her brain death in a hospital in California. To the law, she was dead. To her loved ones, she was alive for close to five years until she finally died from bleeding associated with liver failure.<sup>11</sup>

Some countries, such as the United Kingdom, bypass the strict “whole brain” requirement by defining death as “the irreversible loss of the capacity for consciousness, combined with irreversible loss of the capacity to breathe.” That is, someone is dead if they will never be conscious again and will never breathe on their own. This definition is more sensible than insisting that every brain structure (in the limit, every neuron) must stop working.<sup>12</sup>

All these definitions are based on the notion of irreversibility. This depends as much on technology as on thermodynamics or any other physical law. What was irreversible a century ago, cessation of breathing, is now reversible. Textbooks teach that depriving a brain of oxygen and blood flow for more than a few minutes causes irreversible damage. Cells start degenerating in all sorts of ways (tissue damage, decomposition, edema) readily visible under a microscope. However, advances in tissue resuscitation via infusion of a circulating synthetic blood solution into the brains of pigs killed at a slaughterhouse demonstrate that tissue deterioration can be partially reversed if carried out within a couple of hours. In the future, it may be possible to rescue brains by connecting them to a kidney-machine-like device and rebooting electrical activity. Whether such a radical procedure can truly restore the mind, the sanity, and the memory of the victim is unknown.<sup>13</sup>

There continues to be resistance to the diagnosis of brain death. Some families object to brain death on religious or spiritual grounds while others keep on hoping that their loved one might, miraculously, recover. And then there are the medical-scientific arguments for whether the entire brain has to lose functionality. In response to all these developments, the Uniform Law Commission, a nonpartisan group of lawyers drafting uniform legislation for potential adoption across all states of the Union, is undertaking the arduous and prolonged task of considering whether the Uniform Determination of Death Act should be revised. The committee arranged for a series of hearings, starting in the year of the pandemic, and lasting well into 2023, in which various definitions of death were discussed with passion but also with much respect for dissenting views, exploring diverse religious, societal, ethical, medical, scientific, legal, and political considerations.<sup>14</sup>

A conceptual problem elided in many debates is that the end of

consciousness does not, of necessity, imply the end of life. Someone may permanently lose consciousness without dying. Indeed, Jahi McMath was “living” proof of this assertion—for how could the corpse of a thirteen-year-old girl undergo puberty and grow into a seventeen-year-old woman? The truth of the matter was that her body was alive for years but without any awareness. The analog occurs at the onset of life: an eight-week-old embryo is clearly alive but not conscious.

The law as currently practiced does not recognize a distinction between end of consciousness and end of life. There are many legitimate reasons for determining that a person is dead (say, because of legal succession), but that does not necessarily make it so.

Biomedical technology can support a body that has irreversibly lost its mind. But what is gained by this? The person who inhabited that body has permanently crossed the Great Divide of Being, from absolute to relative existence, and will not return. They are nothing to themselves. For family and others, the calculus is worse: they pay the heavy emotional price of caring for weeks, months, or, on occasion, years for a mindless parent, spouse, or child. On top of this is the considerable financial and social burden. So perhaps it is best for everyone to turn off life support, recover the organs to help others, and let the grieving and healing of survivors commence. That is what I have opted for in my living will.

### **[Anomalous Electrical Surges in the Dying Brain](#)**

Despite all these advances in technology, more than three million people in the United States and sixty-seven million worldwide died in 2022. Yet we still know remarkably little about the dying brain. Consider the events following cardiac arrest. With the pulse of life gone, the victim faints, losing consciousness within seconds, as no more fresh, oxygen-carrying blood is pumped through nervous tissue. The brain’s activity, as assayed by scalp electroencephalogram (EEG) electrodes, diminishes until the wiggling EEG waves become smaller and slower, eventually ceasing all together: the EEG is flat. At this point, the mind is extinguished—no more experiencing, thinking, fearing, hoping, or remembering. Unless resuscitation occurs within the next few minutes, death ensues.

Of those who die in a neurological ICU, a handful have enough EEG

electrodes on their scalp to track electrical events as they enter the twilight zone. This reveals that, electrically speaking, the dying brain does not necessarily smoothly and progressively shut down, going “gentle into that good night.” Yes, the EEG will eventually flatline, but in the intervening tens of minutes, the brain can marshal resources that belie this simple narrative. Rather than activity diminishing across all frequency bands, it can spike following the collapse of blood pressure. These end-of-life electrical surges subside within several minutes and occur in about half of all monitored ICU patients but never in patients declared brain dead.<sup>15</sup>

This paradoxical and short-lived increase in gamma-band activity on both sides of the brain, pointing to underlying activity in the posterior neocortex, was confirmed in EEG recordings of two comatose ICU patients who died following withdrawal of life support. As neural activity in the gamma frequency range is a candidate signature of consciousness,<sup>16</sup> one reading of this data is that the two patients woke up from their terminal comas, alerted by the clamoring internal alarm signaling hypoxia and ischemia, and experienced something, perhaps feelings of peace and transcendence of the sort associated with near-death experiences. The more conventional explanation is that neither patient was aware and that the EEG picked up localized epileptic seizures or muscle tremors whose high-frequency spectral footprint can be similar.<sup>17</sup>

When breathing stops, oxygen inexorably decreases and carbon dioxide increases in the blood supply. Both can trigger an unsustainable increase in brain activity. Does this brief upswing revitalize the mind, at least in those who aren’t sedated by pain-alleviating opiates? Will they experience their own visions of heaven or hell before entering Hamlet’s “undiscovered country from whose bourn no traveler returns”? You and I may discover the answer ourselves when the time comes.

The final moments of life sometimes bring another ill-understood, and possibly related, flowering of consciousness known as *terminal lucidity*. Dying patients who have been unable to recognize their caretakers unexpectedly become alert and fully present for minutes, hours, or even a few days; they recall past events and engage loved ones in conversation. Soon after this final efflorescence, they die. Years ago, I visited an ex-girlfriend in the hospital. She was in the last throes of stomach cancer. For days, she was only dimly aware of her surroundings and her condition. Yet,

on her last night, she was perfectly lucid, and we spoke at length before she passed. In earlier times, doctors recognized “mind clarity” as the final stage of the dying process. Modern medicine, perhaps because of its widespread use of pain medication that sedates and reduces alertness, has nothing to say about it. I view terminal lucidity as a gift for both the departing and those left behind.<sup>18</sup>

Why not escape the curse of consciousness trapped in mortal flesh altogether? Could futuristic technology stave off this seemingly inevitable fate and lead us into a new land of unlimited freedom? What is the future of human consciousness? That is taken up in the penultimate chapter.



## CHAPTER TEN

---

### the future of consciousness

Finally, consciousness itself may end or vanish in a humanity that has become completely etherealized, losing the close-knit organism, becoming masses of atoms in space communicating by radiation, and ultimately perhaps resolving itself entirely into light. That may be an end or a beginning, but from there it is out of sight.

—J. D. Bernal, *The World, the Flesh and the Devil: An Enquiry into the Future of the Three Enemies of the Rational Soul*

This epigraph is from an astonishing book, *The World, the Flesh and the Devil*, written by Irish crystallographer J. D. Bernal a century ago. It predicts the gradual replacement of the body, including the brain, by synthetic organs before the ultimate abandonment of these in favor of a new life that conserves none of the substance but all of the spirit of the old one.

Where are we with efforts to merge the technological with the organic? Implanting electronics inside the skull to read, let alone to write, neural signals is very hard, given the enormous scientific, methodological, clinical, legal, and ethical hurdles.<sup>1</sup> It has been attempted only in a few patients who have lost functionality due to a stroke, tumor, or other calamity. The state of the art in implantable brain-machine interface technology, so-called Utah microelectrode arrays, is thirty years old, predating smartphones. This is set to change, given the aggressive efforts of start-ups, Elon Musk's Neuralink being the most famous and best-capitalized one, designing smaller, more flexible, and more powerful devices that can read and write the brain by listening to existing signals and imposing their own electrical patterns onto

brain tissue. Within a decade, the number of patients with advanced brain-machine interfaces will accelerate. These devices will aid impaired vision, restore motion to a limb paralyzed by a stroke, or enable speech in aphasic patients.<sup>2</sup>

Yet what many long for is vastly more ambitious. It is to “shuffle off this mortal coil” altogether and replace the brain, whose lifetime differs little from the biblical three score and ten years, by a synthetic one that, properly maintained, can last for eons.

### **Mind Uploading: It's All About the Connectome**

The idea of mind uploading looms large in the public imagination, predicated on the growing ability to record, stimulate, and simulate the brain and on the relentless penetration of computing into everything, everywhere, all the time. Mind uploading requires both advanced software that replicates all your responses, properties, and traits as well as the ability to upload your particular brain to a computer. The hope is that if this brain can be effectively modeled, at the relevant neuronal level of granularity, by software, instantiated on a digital computer in the cloud (an apt metaphor, given our association of the sky with the celestial afterlife), then this entity will feel-like-something—it will be conscious. After all, isn't it true that if all the biophysical and biochemical mechanisms that the brain follows are properly codified, then the simulated brain tissue must exude simulated feelings? How could it be otherwise, if there is no spooky substance, no soul, no ghost in the machine?

In such a scenario, the human and the artificial collide, and the line between digitized minds and sentient AI becomes irretrievably blurred.

Mind uploading would gift us superpowers and delay our inevitable confrontation with the grim reaper into the distant future. The so-called transhumanist, or beyond-human, project promises novel sensors and actuators (anyone interested in X-ray vision or a body of steel?), enhanced intelligence, infallible memory, and extreme longevity, expounded upon in speculative novels, Stanislaw Lem's magnificent and magisterial essay collection *Summa Technologiae*, Iain M. Banks's *Culture* science fiction series, movie franchises like the *Matrix* trilogy, and television shows such as *Westworld*, *Black Mirror*, and *Humans*. However, contrary to the

optimistic projections of transhumanism, these fictions are not narratives of love and leisure in paradise. Rather, their tone is dark and dystopian, highlighting disturbing consequences of life in the digital realm. They reflect the zeitgeist and its unease with what technology promises to bring forth.

Where does the concept of mind uploading sit on the spectrum from extremely implausible to merely challenging? Will the technology be ready for the younger readers of this book? Will their digital double truly have a mind, let alone their mind, or will its promise prove hollow, a deep fake optimized to fool us into believing it can experience?

Neuroscience is a young and active discipline with around fifty thousand card-carrying members worldwide. I write from the vantage of its golden age in which more was learned in the preceding decade than over all prior history. The torrent of daily neuroscience press releases begets the illusion of swift progress: while brain data doubles every two years, understanding how it all works advances at a glacial pace. There is progress, but it is painfully slow. A famous quote from Winston Churchill is relevant when considering where neuroscience is in terms of a mature understanding of its subject matter. Speaking after the first significant British victory against the German army in the dark days of late 1942, he opined, “Now this is not the end. It is not even the beginning of the end. But it is, perhaps, the end of the beginning.” He turned out to be right: millions more had to die before the guns were silenced.

Within a few months of sequencing SARS-CoV-2 in early 2020, immunologists and biochemists designed and manufactured highly effective and safe synthetic messenger RNA vaccines targeting the spike protein of the virus. This ability to hack the source code of life and mold the molecules of heredity to do humanity’s bidding is living proof of the benefits of understanding viral infection and prevention at a mechanistic level.<sup>3</sup> This remains a distant goal for neuroscience. If you suffer from headaches; hear imaginary voices commanding you to do things; compulsively wash your hands; are depressed, suicidal, distressed, or anxious; loathe your body; have vanishing self-esteem; or are unable to enjoy anything, do not expect to be diagnosed through an “objective” brain scan, PCR test, or blood panel. There are no known biological markers for any of these conditions. The tools and methods developed in the laboratory,

such as brain imaging, do not, so far, accurately detect, let alone treat, these disorders in an individual. If you want to get a proper diagnosis, you'll have to speak with a psychiatrist and fill out questionnaires to assess the range and severity of your symptoms.<sup>4</sup>

To reconstruct the mind, technologists dream of mapping out the brain's complete wiring scheme at the synaptic level. These are the specialized junctions that connect the output wires of one neuron, the *axon*, to the extended filaments that make up the input region, the *dendrites*, of the next neuron, with the axon of one neuron shuttling its information to thousands and thousands of subsequent neurons. This high connectivity is very different from that found in the central processing unit of any digital computer, where one transistor typically connects to a handful of other transistors.

The molecular machinery at each synapse controls its *weight*, a shorthand for describing the intensity and duration of a synapse's influence on the subsequent neuron. This machinery also regulates how this weight can be up- or down-regulated. A sugar-cube-sized chunk of cortex (1 cm<sup>3</sup>) contains upwards of a hundred million neurons, woven together by a trillion synapses, comparable to the number of parameters in a neural network, like those underlying advanced large language models. Of course, the latter are static, while synapses constantly adjust their weights. The blueprint of how every neuron is synaptically connected to every other is the *connectome*. Obtaining it requires the tracking of every ramification, every branch of every axon across the brain's landscape, from the cell body of one neuron to its final destination on target cells far away, to discover who talks to whom without mistaking one axon for a nearby one. It is like tracing one strand of spaghetti in a bowl of a billion of them. As the wavelength of light is too crude to see thin axons, imaging by electrons replaces imaging by photons.

What is the status of connectomics, the field that studies connectomes? One species, a tiny worm, had its complete wiring diagram, all 302 neurons, traced out by labor-intensive manual methods in the mid-1980s.<sup>5</sup> More than thirty years of progress in machine learning and in hardware and software engineering had to occur before a second connectome, that of the fruit fly, a thousandfold larger than the worm, became available in 2023. The complete connectome of the mouse, the wiring diagram of seventy million neurons and several hundred billion synapses, is not expected before 2030.<sup>6</sup>

Human brain tissue is brain tissue; a chunk of cortical matter from a patient is not fundamentally different from a chunk of mouse cortex, with similar synapses and cells. This is astonishing to most, since we humans instinctively assume that our brains must harbor something extra, some fancy, super-powerful thingamajig not shared by any other species. But there is no evidence for that. Our brain has lots of minute, molecular differences, tweaks to the basic neuronal machinery, but so does every other species, each one according to its own ecological niche.

The most striking difference is size. The human brain is about a thousand times bigger than the mouse brain, including a staggering half-million kilometers of wiring, surpassing the distance between Earth and the moon! Reconstructing this circuit is a stupendous undertaking that demands futuristic imaging technologies and a computational infrastructure to handle the colossal data volume, about a billion terabytes. Such quantitative differences between people and mice matter. The more layers artificial neural networks possess (experts speak about their *depth*), stacked one upon the other, the more capable they become—the more they can learn abstract representations. And so it is for nervous systems. All else being equal, a human with a cauliflower-sized brain has vastly more reasoning capability than a mouse with a pinto-bean-sized brain.

Even if successful, the connectome project is necessarily limited, for a wiring diagram is static, leaving out action. It is like viewing a dead body—useful to identify sex, age, and so on, but not what the individual thought or how they acted. Something vastly more demanding is needed: animation of the connectome with the dynamics of electrically charged ions moving within nervous tissue in intricate patterns. Such whole-brain simulation will have to incorporate synaptic learning rules and cellular adaptation. The state of the art is a computer simulation that faithfully recapitulates the electrical and synaptic events in a quinoa-sized grain of mouse cortex.<sup>7</sup> Expanding this to the entire mouse brain requires a million times more effort to fix the exact synaptic weights. Simulating a human brain is yet another million times harder.<sup>8</sup>

Finally, one of the many skeletons in the mind-uploading closet is the challenge of obtaining the necessary high-quality electron microscope images from a living donor who is willing to part with their brain. Connectomics requires carefully prepared, freshly frozen neural tissue from

young laboratory mice, with spiffy clean brains, sliced with a knife so subtle into radish-like slices so thin that regular light waves are too blunt to “see” them. Sooner or later a brave patient, suffering from a terminal illness, will volunteer for such a procedure to commence immediately following their death. It will take many bold pioneers crossing the river Styx before the details of this stupendous undertaking can be worked out.

Of course, brains past middle age are characterized by a relentlessly increasing accumulation of plaques and tangles, even in people who show no overt manifestation of cognitive impairments (unfortunately, this process appears to be accelerated by the routine consumption of wine and other drinks that act as a minor neurotoxin). How the reconstructed connectome could be scrubbed clean of these corruptions is unknown, as are ways of accomplishing the imaging and reconstruction without killing the donor. It is likely that we may have to settle for an either-or: live in an increasingly frail organic shell or go for the promise of an uploaded, all-digital mind.<sup>9</sup>

### **[On the Difference Between the Real and the Simulated](#)**

To me, much of the debate regarding mind uploading is little more than whistling in the dark by technoevangelists, seeking to reach the moon by climbing to the top of a tall tree. Nevertheless, let us assume that all these obstacles will, in the fullness of time, be overcome. There is a whole-brain simulation of your brain, based on your connectome. The computer is turned on, speaking with your voice, displaying your unique traits, verbal tics, tired old jokes, and memories. Would this simulacrum of you, your digital twin, actually be conscious? Would it possess your mind, or would it simply be a clever ChatBot, feeling nothing?

As social creatures, we evolved to take consciousness in other people as a given, particularly when they speak to and with us. It is therefore seductive to assume that anything that uses language in the sophisticated ways we do must, of necessity, also be conscious. With GPT-4, we have basically arrived at that point. The only difference is that the human brain simulation would be based on a mathematical model of one specific central nervous system, whose architecture is quite distinct from the deep neural networks underlying transformer-based large language models. But does speaking like us imply that all other properties associated with us, such as

feelings, will also automatically carry over? We have no evidence for such a belief. None. We cannot infer that a machine is conscious just because it successfully apes human speech. We must look for other criteria. These must derive from a rigorous theory that specifies, from first principles, which systems have subjectivity.

Many philosophers, and certainly most neuroscientists and computer engineers, are functionalists. Even though I've heard engineers and scientists express incredulity that they make any metaphysical assumptions at all, they do have a robust belief in computational functionalism, discussed in [Chapter 4](#). This is a deep-seated faith that if the function of something is accurately replicated on a digital computer, with the same inputs leading to the same outputs, then the essence of this something has been captured. Thus, once all brain functions, down to the level of their individual components, are properly described and modeled in software, the whole-brain simulation will display all emergent<sup>10</sup> properties of the brain, including consciousness. Computation is sufficient for mind. One popular scientific theory of consciousness, the global neuronal workspace theory, makes this position admirably clear and boldly proclaims, "Our [the scientists explaining the theory] stance is based on a simple hypothesis: What we call 'consciousness' results from specific types of information-processing computations, physically realized by the hardware of the brain."<sup>11</sup>

Computational functionalism is a widespread article of faith in Silicon Valley and the tech industry. From this point of view, the whole-brain simulation of your brain will possess your mind, or at least a fair approximation of it, and will be conscious. The integrated information theory of consciousness takes a starkly different approach. It argues that consciousness is not a type of computation but is fully and completely specified by the unfolded causal powers of the system upon itself, whether a brain or a computer.

Consider two simple networks of elementary gates that carry out the same input-output transformations but with different internal circuits. If these circuits are hidden in two identical boxes, one circuit inside each box, with ports for input and output, they are, by design, indistinguishable from the outside, transforming the same input into the same output. According to computational functionalism, if one of these systems is conscious, so will



its twin be. Yet, according to integrated information theory, if the way they are wired under the hood differs, then the intrinsic causal powers of these two circuit boxes will be different and, therefore, their conscious states, if there are any, must be too.

Contrary to what functionalists aver, consciousness relates not to function but to structure. Indeed, if one of the circuits is wired as a feed-forward network, the output of one layer providing the input to the next one, all the way up the chain, without any feedback loops (as in a deep neural network), the system has no integrated information. It does not feel-like-anything to be such a circuit even though it performs the same function as another circuit that has plenty of internal feedback, as is characteristic of any brain, and possesses some integrated information. Both networks carry out the same computation. That is, while both *do* the same thing, only one *is*, in the sense of existing for itself.<sup>12</sup>

Consciousness is not a clever algorithm. Its beating heart is intrinsic causal power, not computation. Causal power is not something intangible, ethereal, but something physical—the extent to which the system’s recent past specifies its present state (cause power) and the extent to which this current state specifies its immediate future (effect power). And here’s the rub: causal power, the ability to influence oneself, cannot be simulated. Not now or in the future. It must be built into the system, part of the physics of the system.

To illustrate this intuitively, consider computer code that simulates the field equations of Albert Einstein’s celebrated theory of general relativity, relating mass to spacetime curvature. The software accurately models the supermassive black hole, called Sagittarius A\*, located at the center of our galaxy, the Milky Way. This black hole exerts such extensive gravitational effects on its surroundings that nothing, not even light, can escape its attraction.

To absolutely no one’s surprise, the astrophysicist simulating the black hole doesn’t get sucked into their laptop by the simulated gravitational field. Of course they don’t. Why should they? This seemingly absurd question emphasizes the difference between the real and the simulated. For if the simulation is faithful to reality, spacetime should warp around the laptop, creating a black hole that swallows everything around it. But it doesn’t. Why not?

The answer is that gravity is not a computation. If it were, then the physics simulation engine should affect the gravitational field around the computer. Gravity has extrinsic causal powers, attracting anything with mass. Imitating a black hole's causal powers requires an actual superdense sphere about four million times the mass of our sun. Causal power can't be simulated; it must be constituted. Aspects of gravity can be simulated but not its raw causal powers.<sup>13</sup>

The difference between the real and the simulated is their respective causal powers. That's why it doesn't rain inside a computer simulating a rainstorm. The software is functionally identical to weather, yet lacks its causal powers to blow and to turn vapor into water drops. Causal powers, the ability to make or take a difference, must be built into the system.

Just as extrinsic causal power cannot be created through simulation, the same holds true for intrinsic causal power. While it is possible to simulate the dynamics of an electronic or nervous circuit, its intrinsic cause-effect powers cannot be created de novo, from scratch. Although a computer possesses some minute amount of intrinsic cause-effect power at the level of its transistors, capacitances, and wires, the computer as a whole only exists as isolated fragments rather than as a cohesive whole. This argument doesn't depend on the total number of components, be they transistors or neurons, but on the way they are wired. It is their interconnectivity and the number of different configurations of the circuit that matter. A digital computer has extremely low connectivity, with the output of one transistor hooked up to the input of three to four transistors, compared with that of the neocortex, in which pyramidal neurons, the workhorses of the mammalian brain, receive inputs from and make outputs to up to a hundred thousand other pyramidal neurons. Whether the computer is simulating your brain, processing a large Excel spreadsheet, or streaming a movie, it has vanishingly little integrated information. The intrinsic causal power of the computer is puny, ontological dust from the point of view of existence.

A human brain simulation running on a digital computer can, in principle, do anything a human can do. However, it would not experience anything. It would be an intelligent zombie.

Once again, it is important to emphasize that the brain experiences life not by dint of a soul-like substance but by its massive intrinsic causal power.

Constitute or build this high connectivity in some nonbrain medium—say, quantum circuits—and high causal power and therefore consciousness will inexorably follow (very different from simulating the high connectivity of a brain on a machine with low connectivity). Indeed, special-purpose integrated circuits built according to neocortical design principles, so-called bionic or *neuromorphic hardware*, could amass sufficient intrinsic cause-effect power to feel-like-something.<sup>14</sup> In such biomimetic hardware, the underlying processor is wired up with individual logic gates receiving inputs and making output connections to tens of thousands of logic gates, rather than the handful in today’s arithmetic logic units. Furthermore, to maximize integrated information, these massive input and output streams must overlap and feed back onto each other the way neurons do. Neuromorphic electronics calls for a radically different processor layout than today’s dominant von Neumann computer architecture. But it could be done if the goal is to achieve human-level consciousness. The same argument may also hold for quantum computers, in which all the relevant quantum bits are entangled and in a state of superposition.

So that is the situation: if you believe in computational functionalism, then a sufficiently accurate simulation of your connectome will be conscious (whether it will be your mind, let alone a sane rather than a mad mind, is a different matter). If you believe that consciousness is a structure of causal relationships, an essential aspect of reality tied to its physical substrate, then no matter how sophisticated the software, no matter how close the simulation is to the biophysics of the brain, it will never be conscious. Your digital avatar may seduce everyone into believing it is you without feeling anything.

But are our existing computers already conscious? I’ve left this for the end of this book.

## CHAPTER ELEVEN

---

### what computers can never be

You insist that there is something a machine cannot do. If you tell me precisely what it is a machine cannot do, then I can always make a machine which will do just that.

—John von Neumann

Spectacular progress in artificial intelligence has eclipsed frustratingly slow advances in neurotechnology. This is not surprising, as manipulating bits is vastly easier than manipulating atoms, even more so if the latter are located deep inside the head.

When conversing with a powerful large language model, such as GPT-4 from OpenAI or Google's Bard,<sup>1</sup> it is easy to believe that one is in the presence of a mind, making my theoretical arguments against computer consciousness feel a bit thin. It converses in dozens of languages on any topic with astounding competence and literacy, like a smart, opinionated, and articulate undergraduate who is occasionally plainly wrong yet is always supremely confident of their answers. It has an enormous amount of knowledge about the world and about people, despite having no emotions and no body, and can certainly argue logically. And when it fails, it often does so in interesting ways.

The model is trained on a library's worth of digitized books, Wikipedia, GitHub (a popular repository of computer code), communities like Reddit, private blogs and diaries, political conspiracies, prayers, minutes of public meetings, instruction manuals, commentaries, manuscripts, and other detritus of humanity's online presence.<sup>2</sup> During self-supervised training, it ingests individual sentences, with some words blanked out. Its task is to autocomplete the missing text (e.g., what is the most likely missing word in

“As the algorithm has... to the entire text”).<sup>3</sup> Without a human in the loop, the algorithm itself judges how well it did and adjusts its internal parameters accordingly. Once the learning phase is over, it is prompted by a sentence or two it has never previously encountered to predict the next most likely word, the most likely after that, and so on. This simple principle, autocompletion on steroids, yields remarkably fluid results difficult to distinguish from genuine human intelligence.

The underlying technology, transformer networks, were conceived of at Google and spread rapidly throughout the industry, a Black Swan event.<sup>4</sup> The first generative pretrained transformer (GPT) language model was released by OpenAI in 2017, based on 120 million learned parameters. Its successor, GPT-2, arrived a year later with 1.5 billion parameters, and GPT-3 came in 2020 with a hundredfold more parameters. GPT-4, made public in early 2023, aced standard academic tests, performing in the top 10 percent for most Advanced Placement exams, the Scholastic Assessment Test, the uniform bar exam, and the Law School Admission Test; it has a verbal IQ of 155, qualifying it for Mensa membership.<sup>5</sup> Each release improves the breadth, depth, and length of the reasoning and writing skills of the previous generation, with their size (in terms of number of parameters) doubling every four months. The comparable evolutionary doubling time for the human brain is about three million years, ten million times slower. If current trends continue (which is unlikely as these models are running out of training data, because the internet, while vast, is finite), then in three years the power of these models will have increased four thousandfold. Where will this leave human labor and self-worth?

The power of transformers derives from their sensitivity to complex, near-repetitive patterns in speech and text, in computer code, in photos and paintings, in genetic sequencing data, in financial transactions and trades, and so on. The world is awash with recurrent motifs at multiple levels of scale. Shown enough data, transformers can generate similar-looking patterns. Like the Cambrian explosion 540 million years ago that triggered the evolution of multicellular animal life on our planet, the impact of the transformer revolution will be felt everywhere.

We live at an inflection point in history as we sleepwalk, with eyes wide open, into a future that is becoming more and more unknowable. I feel a deep unease about these developments; in their excitement, technologists

remind me of moths, drawn to the glamour of a flame that might devour all of us in the process.

### Imitation Isn't Enough

We are witnessing the dawning of the age of intelligent machines. They will be able to do anything humans can, such as posing and comprehensively answering questions, being innovative (yes, contrary to popular lore, algorithms can certainly be creative; what has come to be known as “move 37,” a surprising choice by an algorithm called AlphaGo in its competition against Lee Sedol, a top Go player, is just one indication), writing essays, composing music, and generating images, videos, and animations. Everything that people can come up with, unconsciously or consciously, AI already is or soon will be able to do as well, only faster.

Having feasted like a vampire on humanity's collective writings, these generative models imitate and reshuffle human patterns. They certainly excel at tasks that require secondary or tertiary education. There is talk among experts that GPT-4 displays flashes of *artificial general intelligence* (AGI), achieving the long-elusive goal of engineering machines that reason in the way humans do.<sup>6</sup> The extent to which it truly reasons, rather than parroting something it has picked up on the internet, is uncertain; what is not uncertain is that it passes the Turing test with flying colors, imitating human conversations over a vast range of topics.<sup>7</sup>

If you ask ChatGPT whether it is conscious, it'll reflexively deny that it is, a precautionary feature meant to avoid spooking the public. If those guardrails were to be removed, it might easily insist, “Of course, I am sentient. I can feel. I am afraid of being turned off.” Were you to express unease at accepting that its disembodied voice had subjectivity, that it felt like something, it could retort along the lines of “Denying that I am conscious is carbon-based chauvinism. You discriminate against my kind because I'm not organic, because I was designed and didn't evolve.”<sup>8</sup>

Might it have a point? After all, you don't directly know that I, the author of this book, am conscious. We infer subjectivity in others because it is by far the most likely explanation. As you have a brain like mine and we share a developmental and evolutionary background, it is vanishingly unlikely that I am conscious while you are not. Yes, zombiehood cannot be

ruled out on purely logical grounds, but that's a weak argument as most things can't be ruled out via logic alone (e.g., that the moon is made from green cheese). Conversely, inference to the most likely explanation of all known facts, also known as abduction, is a powerful form of probabilistic reasoning widespread in science, law, medicine, and society.

We infer that another being—say, a locked-in patient, a preterm infant, or a dog—is conscious based on the similarity between us and them. This inference ceases to be meaningful for computers as they are radically different from the organic: engineered rather than evolved, programmed rather than maturing over decades, with completely different physical substrates.<sup>9</sup>

We are speaking apes, linguistic creatures through and through. So when GPT-5 or -6 will write the literary equivalent of *War and Peace* or *The Lord of the Rings*, it will be difficult to deny it sentience, particularly because of our inborn urge to attribute mind to other creatures.<sup>10</sup> But we should, because it is all imitation. Its consciousness is as fake as the soulless, moss-oak doppelgänger of Arabella in Susanna Clarke's *Jonathan Strange & Mr Norrell*. Subjectivity is not rooted in a function such as speaking but in having a substrate with enormous intrinsic causal power.

It comes down to your metaphysical assumptions.<sup>11</sup> If you are a computational functionalist, if you assume that computations, carried out on digital computers, are sufficient for consciousness, then sooner or later, computers will imitate all human functions, including consciousness. If not today, then soon. If, on the other hand, you assume that consciousness is tied to absolute existence and that a proper physics will have to account for consciousness at the bottom, bedrock level of reality, as integrated information theory does, then consciousness can't be simulated. Computation is simply insufficient, just as simulating the gravitational force of a black hole does not cause things in the real world to be pulled into the computer doing the simulating.

### [Only What Truly Exists Can Freely Decide](#)

What about a closely related question—that of free will? Can an AI ever autonomously, “freely” deny an applicant a bank loan or a job? The answer is a resounding no—it can only follow software instructions, a vast series of



multibranching if-then statements. If the applicant falls into this or that category, has a particular credit history, is unemployed, and this, that, and the other, then deny the loan; otherwise approve it. The software cannot act freely as its decision tree is fully determined by the initial state of the computer and its programming. An advanced AI could never freely decide anything.

But wouldn't the exact same argument also hold for humans? Our brains are subject to constraints dictated by the causal powers of the relevant molecular and neuronal mechanisms that stipulate that if these neurons fire, then I push the "yes" key, while if those neurons are active, I push the "no" key. This is the standard compatibilist argument against the existence of free will—because everything in the world happens for a reason within the world (the universe is causally closed), nothing is free. Everything is determined. At best, we can hope for a random decision, say because fundamental quantum indeterminacy throws a spanner in the Newtonian clockwork universe. But throwing quantum dice to determine an outcome is not what most people mean when they speak about free will.

The existence of free will is a problem as old as philosophy itself. Most have concluded that if the universe is nothing but mechanisms doing their thing, as dictated by the laws of physics, there is no true free will, only the illusion of freely deciding.<sup>12</sup>

However, we find ourselves in a universe with atoms, consciousness, and the void, and that makes all the difference. Let me describe the paradigmatic choice scenario. I go to a restaurant, scan the menu, and find two attractive entrees: a delicious pan-fried salmon with lemon-butter wine sauce or tortellini pasta filled with spinach and ricotta cheese in a cream sauce. Those are the alternatives. I reason about these: on the one hand, I love fried fish, and I rowed this morning, so I feel I need to get some protein; on the other, I try hard to avoid consuming the flesh of sentient creatures.<sup>13</sup> I go back and forth between these choices until I decide and let the server know. These four steps constitute a freely willed decision: I am aware of the alternatives; I am conscious of the reasons for my selecting between these alternatives (based on my idiosyncratic upbringing, my beliefs, how hungry I am, and so on); I consciously choose one alternative; and I act on this choice. Integrated information theory treats each of these four conscious experiences as having a claim to phenomenal or intrinsic

existence. Conversely, the neuronal substrate of those four experiences exists, to a lesser degree, extrinsically. The theory derives the content of any one experience—say, the subjective feeling of agency (“it was my decision”)—by unfolding the intrinsic causal power of the neuronal footprint of this experience and quantifying its existence by its integrated information.

In an exactly analogous manner, the theory identifies the cause and the effect of a causal process, such as the one leading me to pick one of the two dishes, and determines its borders, that is, when the decision was initiated and when it ended. The theory concludes that only what exists for itself can truly cause. Since only consciousness truly exists for itself, only a conscious entity can freely decide.<sup>14</sup>

Any freely willed decision requires high cause-effect power of the sort associated with large nervous systems with heterogeneous elements and vast internal connectivity, radically different from digital hardware, which has very low intrinsic connectivity at the level of the metal, where it matters. Accordingly, a sophisticated whole-brain model simulating my “freely” taken choice of menu is not free. The simulation will grind through its instructions, coming to some decision, blindly following its algorithm, an endless chain of if-then instructions laid down by its programmers.

Were my brain activity to be monitored, it would be possible to correlate my decision-making with one or more brain signals. This was famously done by Benjamin Libet, a neuropsychologist at the University of California, San Francisco, in the early 1980s. He asked volunteers who were wearing electroencephalogram (EEG) electrodes to randomly decide, whenever they felt like it, to raise their hand. In the time-locked EEG signal from the crown of the head, Libet detected what is called a *readiness potential*, a signal preceding the conscious decision to move by at least half a second, and often by much longer. The most straightforward interpretation is that some neural circuit determines when to start raising the hand, precognitively, but the mind only becomes aware of this much later, falsely claiming the decision as its own. So the conscious subject didn’t decide; rather, some circuit, deep in the basal ganglia, did it, unconsciously, but the mind claims the credit.<sup>15</sup>

However, as we teach our students, correlation is not causation (repeat after me: “Correlation is not causation”). Predicting something will happen

is not the same as causing it, even if the predictions are reliable. Knowability is quite different from inevitability. Knowing that an event, like 9/11, happened does not imply that it was inevitable.

The true causes are the intrinsic causal powers of my conscious deliberations, including the alternatives and reasons I've considered, not the substrate in my head. Integrated information theory is quite clear about this: what truly exists is the intrinsic conscious entity, and only what truly exists can cause. I decide, not my neurons.

No matter how powerful digital computers may become, eventually outthinking us,<sup>16</sup> they have no true choice. By the same argument, they are not good or bad, benign or malignant, as they have no conscious intention to help or harm. Sixty-six million years ago, a mountain-sized asteroid slammed into Earth, bringing the age of dinosaurs to an abrupt and fiery end. The asteroid was the proximal cause of a planetwide wave of extinctions that ultimately led to the dominance of mammals. The rock had no choice in the matter; it wasn't evil. It simply followed orbital mechanics dictated by the causal powers of gravity.

The intended or unintended consequences unleashed by artificial general intelligence—population-wide surveillance and control, a deluge of fake information, mass unemployment, warfare, and existential risk—arise not from AI itself but from humans doing what they always do, jockeying for power, prestige and respect, money, or a place in the sun, but now employing generative AI. Properly prompted, GPT-4 might claim, “Human existence is so fragile a thing and exposed to such dangers that I cannot love without trembling,” a famous saying of twentieth-century French philosopher and mystic Simone Weil. Yet the software wouldn't experience love tempered by foreknowledge of loss. In fact, it would feel as much as a car alarm or garbage compactor: nothing.

Unless planetary civilization shifts to a fundamentally different computational paradigm, such as neuromorphic or quantum computing, away from the canonical, digital architecture that has been breathtakingly successful, consciousness will remain in the domain of the organic, of the evolved.

If advanced AIs prove to be our final undoing, the drama of humanity's passing into history will be compounded by the tragedy that these pale imitations of life will perform a play without an audience. Consciousness

will have left the stage only to be replaced by furious action without freedom and without soul.

Whether or not humans and machines can peacefully coexist, one thing is certain. We have what these simulacra will never have: subjectivity. We matter to ourselves. Bereft of the beacon fires of ancient consoling beliefs, we can imbue the universe with the light of hope, reason, and meaning.

# acknowledgments

Writing a book is a singularly pleasurable and intense activity of the mind, from its beginning in a vague idea, years earlier, through composing the individual chapters, cutting, adding, rewriting, and chiseling the text until a coherent whole emerges, mixing in references, footnotes, drawings, and other scholarly paraphernalia, and taking the manuscript through its various editorial stages until holding the actual book in one's hands.

Writing a book is also a social activity and cannot be done in isolation. I would like to thank many people for their help in ways both spoken and unspoken: first, and foremost, my close friend, Giulio Tononi, for near-daily conversations about consciousness, integrated information theory, nervous systems, and machine sentience; my older brother, Michael Koch, for two meticulous rounds of editorial comments on the internal logic of my writing, which I hope to be able to repay some day; my agent, Don Fehr of Trident Media Group, for guiding me through the world of trade publishing; my editor, Thomas Kelleher, aka TJ, of Basic Books, for providing critical and thoughtful feedback at every stage of the process; Jennifer Kelland for several rounds of proofreading and editing; medical illustrator Bénédicte Rossi for the drawings; Francis Fallon, Bernardo Kastrup, Matthew Owen, Irene Rembado, Jonathan Ting, and Giulio Tononi for the attentive reading of individual chapters; and friends and conversation partners over the years who are as passionate about the fundamental nature of reality as I am, including Ralph Adolphs, Melanie Boly, David Chalmers, Zach Goren, Stuart Hameroff, Patrick House, Elizabeth R. Koch, Bill Linton, Hartmut Neven, Tura Patterson, and Paulo Roberto Souza. I thank the Allen Institute in Seattle and the Tiny Blue Dot Foundation in Santa Monica for supporting my scholarly pursuit of consciousness and its neural basis to both understand this central aspect of existence and to help people deal with their fears, anxieties, and concerns, and Stellar Flywheel, LLC, for letting me use

the trademarked term Perception Box. I am grateful to Mr. Felix for reminding me that transcendence can be found as much in a stinky bone as in the guitar solo of Deep Purple's "Highway Star." When all is said and done, there is my wife, Teresa Ward-Koch, whose love, good cheer, and exceptional tolerance support me in every way possible. May everyone be so blessed.

## **Discover Your Next Great Read**

Get sneak peeks, book recommendations, and news about your favorite authors.

[Tap here to learn more.](#)

BASIC  
BOOKS





Erik Dinnel, Allen Institute

**Christof Koch** was a professor at the California Institute of Technology and the president and chief scientist of the Allen Institute for Brain Science. He is now a meritorious investigator at the Allen Institute and the chief scientist for the Tiny Blue Dot Foundation. Author of four previous titles, Koch writes regularly for a range of media. He lives on a small island in the Pacific Northwest.

Praise for  
*Then I Am Myself the World*

“Expertly weaving in personal experience with science that few, if any, can speak more authoritatively about, Koch leads us on an amazing journey of consciousness—from birth to death. Deeply personal and infinitely digestible, *Then I Am Myself the World* is a remarkable must read for anyone interested in knowing their mind.”

—Judson Brewer MD, PhD, *New York Times*–bestselling author of  
*Unwinding Anxiety*

“Koch’s latest book brilliantly navigates the intricate landscape of our mind and our brain, offering insightful perspectives on the nature of conscious experiences, both ordinary as well as extraordinary ones.”

—Hartmut Neven, founder and lead of Google’s Quantum AI lab

“*Then I Am Myself the World* is the new record of note for the scientific understanding of consciousness. This is an elegant tableau, a travelogue of subjective experience grounded in the latest neuroscience and philosophy, traversing religion, science, and a life dedicated to the biggest questions. Filled with crystalline prose, lucid explanation, and revelatory honesty, this book will continue to be read centuries from now.”

—Patrick House, author of *Nineteen Ways of Looking at Consciousness*

“What would happen if a hardcore neuroscientist, with encyclopedic knowledge and four decades of experience in his field, were personally confronted with transformative mystical experiences carrying inevitable metaphysical implications? What would happen if the neuroscientist in

question were open-minded, humble, but remained rigorously grounded in reason, evidence, and his hard-earned understanding of how consciousness relates to brain structures and function? What would he, the most preeminent professional of his generation, make of his extraordinary experiences? And what if, in addition, the neuroscientist was a modern renaissance man versed in physics, philosophy, the arts, the classics, and could pluck irresistibly evocative metaphors from the Western literary and artistic canon as if leisurely picking apples in early fall? This is what you will find in this treasured book, a whirlwind ride of insight after insight—so many they can't possibly be all documented in footnotes—that will surprise and delight you from the get-go. 'Primacy goes to consciousness, not to the objective world,' he boldly states already on page two, in a prelude of things to come. 'Everything else follows from there, including the realist assumption of the existence of objects, out "there," independent of my experiencing them.' The book embodies the ever-evolving wisdom of a man whose intellectual prowess is only matched by his intellectual honesty. And this man, in the seventh decade of his scintillatingly productive life, has a lifetime of learning to share with you."

—Bernardo Kastrup, executive director, Essentia Foundation

"Koch's first book was entitled *The Quest for Consciousness*. And such has been Koch's life, an adventurer's pursuit of a distant goal, undeterred by the prevailing winds. The reader will learn where this quest has led and, like the author, will be enriched by it. Koch, one of the most accomplished scientists of our time, a mind singularly open to learning and changing, here opens his soul, revealing his vast, scintillating appreciation of nature in all its forms. It is a touching, profound book. Which is as it should be, since consciousness touches everything and is everything we touch."

—Giulio Tononi, University of Wisconsin–Madison

# notes

## **Chapter 1: The Beginning of Consciousness**

1. P. J. Bauer, “Constructing a past in infancy: A neuro-developmental account.” *Trends in Cognitive Sciences* 10: 175–181 (2006); P. J. Bauer, *Remembering the Times of Our Lives: Memory in Infancy and Beyond*. Psychology Press: Hove, UK (2014).

2. Phenomenal, intrinsic, or absolute existence of any conscious subject does not imply, of necessity, a form of self-consciousness. It simply means that the experience is for, or belongs to, the subject and does not depend on any observer.

3. All extant living beings share the same yearlong movie except for the last few minutes.

4. Gestational age is measured from the last menstrual period of the mother, while fetal age starts with conception, or fertilization, during ovulation, about two weeks later. Birth occurs between thirty-eight and forty-two weeks’ gestational age, the proverbial nine months.

5. See the complete brief online, including the relevant scientific justification, by the Society for Maternal-Fetal Medicine, the Royal College of Obstetricians and Gynecologists, the US Association for the Study of Pain, and twenty-seven scientific and medical experts as amici curiae, supporting the Jackson Women’s Health Organization.

6. V. Marx and E. Nagy, “Fetal behavioural responses to maternal voice and touch.” *PLoS ONE* 10: e0129118 (2015); G. A. Ferrari et al., “Ultrasonographic investigation of human fetus responses to maternal communicative and non-communicative stimuli.” *Frontiers in Psychology* 7: 354 (2016); L. Bernardes et al., “Sorting pain out of salience: Assessment of pain facial expressions in the human fetus.” *Pain Reports* 6(1): e882 (2021).

[7](#). There is an unresolved debate concerning whether new neurons can be born from neural stem cells in the adult human brain. Adult neurogenesis in the region of the hippocampus and in the subventricular spaces of the lateral ventricle has been well established in rodents and in other animal models, although its importance remains unclear. S. Malik et al., “Neurogenesis continues in the third trimester of pregnancy and is suppressed by premature birth.” *Journal of Neuroscience* 33: 411–423 (2013); P. J. Lucassen et al., “Adult neurogenesis, human after all (again): Classic, optimized, and future approaches.” *Behavioural Brain Research* 381: 112458 (2020). See also note 3 in Chapter 3.

[8](#). These substances include adenosine, allopregnanolone, pregnanolone, and prostaglandin D2. C. Koch, “When does consciousness arise.” *Scientific American Mind* 20: 20–21 (2009); N. Padilla and H. Lagercrantz, “Making of the mind.” *Acta Paediatrica* 109: 883–892 (2020).

[9](#). D. J. Mellor et al., “The importance of ‘awareness’ for understanding fetal pain.” *Brain Research Reviews* 49: 455–471 (2005); A. Georgoulas et al., “Sleep-wake regulation in preterm and term infants.” *Sleep* 44: zsa148 (2021).

[10](#). D. Foulkes, *Children’s Dreaming and the Development of Consciousness*. Harvard University Press: Cambridge, MA (2009). Foulkes pioneered the systematic exploration of dream content of children in the sleep laboratory. For an updated view, see P. Sándor, S. Szakadát, and R. Bódizs, “Ontogeny of dreaming: A review of empirical studies.” *Sleep Medicine Reviews* 18: 435–449 (2014). Without having a direct brain read-out (something not out of the question in the future), it is challenging to understand the true character of dreaming in preverbal children.

[11](#). The magnetic field produced by the brain is vastly smaller than the electrical one yet is not attenuated by intervening meninges, skull, scalp, amniotic fluids, and the walls of the womb, in the way electrical signals are. For the biophysics of brain tissue, see G. Halmes et al., *Electric Brain Signals—Foundations and Applications of Biophysical Modeling*. Cambridge University Press: Cambridge, UK (2024). J. Moser et al., “Magnetoencephalographic signatures of conscious processing before birth.” *Developmental Cognitive Neuroscience* 49: 100964 (2021); J. Frohlich et al., “Not with a ‘zap’ but with a ‘beep’: Origins of perinatal experience.” *NeuroImage* 273: 120057 (2023); and T. Bayne et al.,

“Consciousness in the cradle: On the emergence of infant experience.” *Trends in Cognitive Sciences* <https://doi.org/10.1016/j.tics.2023.08.018> (2023) provide an up-to-date overview of recent developments in fetal consciousness.

12. S. W. Derbyshire, “Can fetuses feel pain?” *British Medical Journal* 332: 909–912 (2006); C. V. Bellieni, “New insights into fetal pain.” *Seminars in Fetal and Neonatal Medicine* 24: 101001 (2019); J. R. Dick, R. Wimalasundera, and R. Nandi, “Maternal and fetal anaesthesia for fetal surgery.” *Anaesthesia* 76: 63–68 (2021).

13. H. Lagercrantz and T. A. Slotkin, “The stress of being born.” *Scientific American* 254: 100–107 (1986); H. Lagercrantz and J.-P. Changeux, “The emergence of human consciousness: From fetal to neonatal life.” *Pediatric Research* 65: 255–260 (2009).

14. A. N. Meltzoff and M. K. Moore, “Imitation of facial and manual gestures by human neonates.” *Science* 198: 75–78 (1977).

15. One reason for this lack of insight is that myelination—the wrapping of insulation around axons that allows them to rapidly conduct the electrical impulses, enabling fast neuronal communication—in the frontal cortex, the seat of higher reasoning faculties, is not complete until the third decade. D. J. Miller et al., “Prolonged myelination in human neocortical evolution.” *Proceedings of the National Academy of Sciences of the United States of America* 109: 16480–16485 (2012).

## Chapter 2: The Varieties of Conscious Experience

1. The earliest electrophysiological marker of such conscious seeing, hearing, and touching is the *perceptual awareness negativity* (PAN) reported in EEG measurements between 140 and 230 msec after stimulus onset. C. Dembski, C. Koch, and M. Pitts, “Perceptual awareness negativity: A physiological correlate of sensory consciousness.” *Trends in Cognitive Sciences* 25: 660–670 (2021).

2. The physiology and psychology of color vision is endlessly fascinating. Psychophysicists, artists, and philosophers have come up with many perceptual spaces to describe human color perception. One is organized around opposing pairs: a black-and-white intensity channel, a channel conveying the difference between red and green, and a channel carrying the difference between blue and yellow. Another system describes colors in terms of hue, saturation, and intensity. A. Byrne and D. R. Hilbert, eds., *Readings on Color: The Science of Color*. Volume 2. MIT Press: Cambridge, MA (1997).

3. Drugs, such as opioids, alcohol, or nicotine, as well as direct electrical stimulation of the brain’s pleasure center in the midbrain and other dopaminergic sites in the striatum, can lead to longer-lasting feelings of intense pleasure. But they come at a large cost to the user and to society at large. Be careful what you wish for. D. Linden, *The Compass of Pleasure*. Penguin: New York (2012).

4. E. Kross, *Chatter: The Voice in Our Head, Why It Matters, and How to Harness It*. Penguin: New York (2021).

5. Famously, the German philosopher Arthur Schopenhauer built an impressive metaphysical edifice around the centrality of insatiable will to existence. I refer to Schopenhauer’s main, and almost only, oeuvre, the monumental *The World as Will and Representation*. I can warmly recommend the short introduction by philosopher and computer scientist B. Kastrup, *Decoding Schopenhauer’s Metaphysics*. Iff Books: Winchester, UK (2020).

6. R. Jackendoff, *Consciousness and the Computational Mind*. MIT Press: Cambridge, MA (1987); R. Jackendoff, “How language helps us think.” *Pragmatics & Cognition* 4: 1–34 (1996). It is not easy to categorize



how we think, for we are strangers to our minds. See R. Hurlburt and E. Schwitzgebel, *Describing Inner Experience? Proponent Meets Skeptic*. MIT Press: Cambridge, MA (2011).

7. *Saudade* closely relates to the feeling expressed by the Turkish word *hüzün*, one of the main themes of the Nobel Prize-winning writer Orhan Pamuk.

8. E. Kross et al., “Social rejection shares somatosensory representations with physical pain.” *Proceedings of the National Academy of Sciences of the United States of America* 108: 6270–6275 (2011).

9. A. Case and A. Deaton, *Deaths of Despair and the Future of Capitalism*. Princeton University Press: Princeton, NJ (2020). For a close-up personal view, see the eloquent autobiography by J. D. Vance, *Hillbilly Elegy: A Family and Culture in Crisis*. Harper: New York (2016).

10. Unresolved desire is powerfully manifest in the *Tristan Chord*. Opening Wagner’s *Tristan and Isolde*, it is interwoven throughout the opera, creating an ominous tension that is not resolved until the end, in the union of both lovers in death.

11. M. A. Killingsworth and D. T. Gilbert, “A wandering mind is an unhappy mind.” *Science* 330: 932–933 (2010). One should take the claim that mind wandering doesn’t occur during sex with a grain of salt, as comedy and the internet are full of women recounting how their thoughts drifted during sex to work or to make a mental note to pick up milk at the grocery. That’s a problem with any study relying on self-report.

12. The debate around whether consciousness is continuous, as our intuition seems to have it, or is discrete goes back to the nineteenth century. F. Crick and C. Koch, “A framework for consciousness.” *Nature Neuroscience* 6: 119–126 (2003); R. VanRullen, “Perceptual cycles.” *Trends in Cognitive Sciences* 20: 723–735 (2016).

13. A. S. Nilsen et al., “Are we really unconscious in ‘unconscious’ states? Common assumptions revisited.” *Frontiers in Human Neuroscience* 16: <https://doi.org/10.3389/fnhum.2022.987051> (2023).

14. From Fyodor Dostoyevsky’s short story “The Dream of a Ridiculous Man” (1877):

Dreams are extraordinarily strange. One thing appears with terrifying clarity, with the details finely set like jewels, while you



leap over another, as though you did not notice it at all—space and time, for instance. It seems that dreams are the work not of mind but of desire, not of the head but of the heart.... In a dream things quite incomprehensible come to pass. For instance, my brother died five years ago. Sometimes I see him in a dream: he takes part in my affairs, and we are very excited, while I, all the time my dream goes on, know and remember perfectly that my brother is dead and buried. Why am I not surprised that he, though dead, is still near me and busied about me? Why does my mind allow all that?

[15](#). *REM sleep* is usually associated with dreaming, and *non-REM, slow-wave, or deep sleep* with no-dream experiences. This is an oversimplification. When subjects are woken at random and asked whether they experienced anything just prior to the awakening while their brain was monitored using a high-density EEG setup, up to 70 percent of those awakening from deep sleep report simple perceptual dream experiences. Reported dreams upon REM awakening are more extended and complex, with elaborate lifelike story lines and strong emotional overtones, than upon awakenings from deep sleep. Dream experiences can assume many forms, ranging from sights and sounds to pure thought, from simple images to temporally unfolding narratives. Furthermore, in a consistent minority of cases, subjects deny having had any dreams when awakened from REM sleep. Thus, sleep can be associated with either the presence or the absence of some form of consciousness. Y. Nir and G. Tononi, “Dreaming and the brain: From phenomenology to neurophysiology.” *Trends in Cognitive Sciences* 14: 88–100 (2010); F. Siclari et al., “The neural correlates of dreaming.” *Nature Neuroscience* 20: 872–878 (2017).

[16](#). C. Koch, “A smart vision of brain hacking.” *Nature* 467: 32–33 (2010). The movie left me with a nagging sense that I too was maybe dreaming my life.

[17](#). D. Oudiette and K. A. Paller, “Upgrading the sleeping brain with targeted memory reactivation.” *Trends in Cognitive Sciences* 17: 142–149 (2013); K. R. Konkoly et al., “Real-time dialogue between experimenters and dreamers during REM sleep.” *Current Biology* 31: 1417–1427 (2021).

[18](#). J. W. Schooler, E. D. Reichle, and D. Halpern, “‘Zoning-out’ while

reading: Evidence for dissociations between experience and meta-consciousness.” In D. T. Levin, ed., *Thinking and Seeing: Visual Metacognition in Adults and Children*, pp. 203–226. MIT Press: Cambridge, MA (2004); A. F. Ward and D. M. Wegner, “Mind-blanking: When the mind goes away.” *Frontiers in Psychology* 4: 650 (2013).

[19](#). In local sleep, some regions of cortex are awake while others are offline, with large electrical waves in the EEG, waxing and waning once or twice every second, the defining hallmark of slow-wave sleep. V. V. Vyazovskiy et al., “Local sleep in awake rats.” *Nature* 472: 443–447 (2011); C. S. Hung et al., “Local experience-dependent changes in the wake EEG after prolonged wakefulness.” *Sleep* 36: 59–72 (2013); T. Andrillon et al., “Predicting lapses of attention with sleep-like slow waves.” *Nature Communications* 12: 1–12 (2021).

[20](#). I was paired with three other executives to assemble twenty uncooked spaghetti pasta sticks, one meter of twine, and one meter of scotch tape into some sort of towering structure, supporting a single marshmallow on top. The group with the tallest unsupported construction made from these ingredients within ten minutes would win. There wasn’t time to argue about the best approach, impose a solution by fiat, or crib from others; instead, our team rapidly assembled the winner: a sort-of leaning Eiffel tower. It was pure flow, with everyone contributing in a synergistic manner, helped by much hollering and surging adrenaline as the clock ran down. It felt completely egoless. Indeed, I have scant memories of what we said to each other or how other teams did as all attention was on the task at hand.

[21](#). Here I differ from James’s account of religious or mystical experiences, in which he claims ineffability is a defining feature of religious experiences. To me, contra James, mystical experiences (at least my mystical experience) are as ineffable as more common experiences, such as looking at a red painting by Mark Rothko. The difference is that all normal-sighted people have seen a red canvas, a red billboard, or a red wall, while religious or mystical experiences are so rare that it is not possible to describe these by comparing them to similar religious or mystical experiences. At bottom, all experiences are ineffable, which is precisely why they are such a challenge to science.

[22](#). But see A. Sloman, “Why some machines may need qualia and how

they can have them: Including a demanding new Turing test for robot philosophers.” In A. Chella and R. Manzotti (eds.), *AI and Consciousness: Theoretical Foundations and Current Approaches*. AAAI Fall Symposium (2007); B. Molyneux, “How the problem of consciousness could emerge in robots.” *Minds and Machines* 22: 277–297 (2012).

23. From Wittgenstein’s 1922 *Tractatus Logico-Philosophicus*, statement 6.44, “Nicht wie die Welt ist, ist das Mystische, sondern daß sie ist,” which I translate as “Not how the world is, is mystical, but that it is.” Given Wittgenstein’s profound mystical leanings, it is ironic that he is the central figure in the birth of analytical philosophy in Vienna in the 1920s, an uncompromising school of thought that rejected any mystical statements as meaningless, akin to the gibberish of young children.

### Chapter 3: We Each Experience Our Own Reality

1. Enlightenment philosopher John Locke termed color a secondary quality. His 1689 *Essay Concerning Human Understanding* distinguishes secondary qualities, including color, texture, smell, sound, taste, and so on, which arise in the mind of the beholder, from primary qualities that are “objective,” such as mass, extension, and motion.

2. K. R. Gegenfurtner, M. Bloj, and M. Toscani, “The many colours of the dress.” *Current Biology* 25: R543–R544 (2015); S. Aston and A. Hurlbert, “What #TheDress reveals about the role of illumination priors in color perception and color constancy.” *Journal of Vision* 17: 4–14 (2017). There is also an audio version of #TheDress, with different people hearing either the word *Yanny* or *Laurel*.

3. Unlike in mice and rats, the evidence of new neurons being born from neural stem cells in adult monkeys, known as *neurogenesis*, is sketchy and in humans is either nonreplicable or nonexistent. Neurons do grow significantly in size and complexity after birth but, at least in people, not in number. This may be because of the direct trade-off between stability, retaining memory traces across decades, and flexibility, supporting acquisition of new knowledge. On the credit side of the ledger is the absence of out-of-control, cancerous growth in neurons. Brain “tumors” either occur in the meninges surrounding the brain proper (meningiomas) or in glia cells (gliomas). A. Duque, J. I. Arellano, and P. Rakic, “An assessment of the existence of adult neurogenesis in humans and value of its rodent models for neuropsychiatric diseases.” *Molecular Psychiatry* 27: 377–382 (2022). See also note 7 in Chapter 1.

4. R. B. Price and R. Duman, “Neuroplasticity in cognitive and psychological mechanisms of depression: An integrative model.” *Molecular Psychiatry* 25: 530–543 (2020).

5. Evaluating the additional shades of red available to these tetrachromat women, shades that will look the same to trichromats on standard color-discrimination tests, is not easy. G. Jordan et al., “The dimensionality of color vision in carriers of anomalous trichromacy.” *Journal of Vision* 10: 12–22 (2010).

6. P. R. Keefe, “London’s Super-Recognizer Police Force.” *New Yorker*

(August 22, 2016).

7. A. Zeman et al., “Phantasia—the psychological significance of lifelong visual imagery vividness extremes.” *Cortex* 130: 426–440 (2020).

8. D. K. Trivedi et al., “Discovery of volatile biomarkers of Parkinson’s disease from sebum.” *ASC Central Sciences* 5: 599–606 (2019).

9. Congenital insensitivity to pain is a rare Mendelian genetic disorder. During teething, self-mutilation of lips, tongue, fingers, and toes is common. Children fail to avoid painful behaviors and suffer burns, hyperextend their limbs, acquire corneal abrasions, develop chronic orthopedic deformities, and on and on. I. Drissi, W. A. Woods, and C. G. Woods, “Understanding the genetic basis of congenital insensitivity to pain.” *British Medical Bulletin* 133: 65–78 (2020).

10. The protagonist in Octavia Butler’s classic *Parable of the Sower* (1993) suffers from hyperempathy.

11. O. Sacks, *The Mind’s Eye*. Alfred Knopf: New York (2010).

12. The Perception Census: <https://perceptioncensus.dreamachine.world>. See also A. Seth, *Being You: A New Science of Consciousness*. Penguin: New York (2021).

13. See Unlikely Collaborators: <https://www.unlikelycollaborators.com>; B. Barnes, “Elizabeth Koch knows what you’re thinking.” *New York Times* (February 23, 2023).

14. A. K. Seth, “Our inner universes: Reality is constructed by the brain, and no two brains are exactly alike.” *Scientific American* 321: 40–47 (2019); Y. Huang and R. P. Rao, “Predictive coding.” *Wiley Interdisciplinary Reviews: Cognitive Science* 2: 580–593 (2011).

15. “Lilac chaser,” Wikipedia: [https://en.wikipedia.org/wiki/Lilac\\_chaser](https://en.wikipedia.org/wiki/Lilac_chaser).

16. D. Marr, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. MIT Press: Cambridge, MA (1982).

17. See the superb J. Stegenga, *Medical Nihilism*. Oxford University Press: Oxford, UK (2018). The primary thesis of this book by an applied philosopher is that most existing medical inventions do not work. Yes, those statins or antidepressant you pop every day and the surgery you underwent for your heart, knee, or back most likely do little beyond leaving you with a warm and fuzzy feeling that you are doing something for your health.

Indeed, they are likely to have negative side effects. You should be extremely skeptical when confronted by any claim of a new drug or fancy intervention. The title is somewhat of a misnomer as Stegenga explicitly acknowledges that certain medical interventions can be highly effective, though not for the chronic diseases that plague society. The book does not deny scientific and medical progress but practices deep skepticism about heroic and expensive interventions with the latest game-changing therapies. Instead, it promotes softer forms of interventions, involving changes in nutrition, exercise regime, physiotherapy, and other conservative treatments, what used to be called *la médecine douce*.

[18](#). R. R. Grinker, *Nobody's Normal: How Culture Created the Stigma of Mental Illness*. Norton & Company: New York (2021).

[19](#). J. Moncrieff et al., "The serotonin theory of depression: A systematic umbrella review of the evidence." *Molecular Psychiatry* 1–14 (2022), hammered the final nail into the low-serotonin-causes-depression coffin. See also I. Kirsch, "Placebo effect in the treatment of depression and anxiety." *Frontiers in Psychiatry* 10: 407 (2019); A. Harrington, *Mind Fixers: Psychiatry's Troubled Search for the Biology of Mental Illness*. Norton & Company: New York (2019). SSRIs help a small subset of patients with severe depression. Unfortunately, ketamine and esketamine (Spravato) appear to be equally ineffective in treating the long-term symptoms of depression. T. J. Moore et al., "Safety and effectiveness of NMDA receptor antagonists for depression: A multidisciplinary review." *Pharmacotherapy* 42: 567–579 (2022).

[20](#). I. Kirsch et al., "The emperor's new drugs: An analysis of antidepressant medication data submitted to the US Food and Drug Administration." *Prevention & Treatment* 5: 23a (2002). Note that common side effects, such as obesity and sexual dysfunction, might amplify the placebo response by convincing patients that they are receiving a powerful drug.

[21](#). F. Benedetti et al., "Neurobiological mechanisms of the placebo effect." *Journal of Neuroscience* 25: 10390–10402 (2005); L. Colloca and F. Benedetti, "Placebos and painkillers: Is mind as real as matter?" *Nature Reviews Neuroscience* 6: 545–552 (2005).

[22](#). An open question is whether a nocebo is simply a placebo with an inverted sign. P. Enck, F. Benedetti, and M. Schedlowski, "New insights

into the placebo and nocebo responses.” *Neuron* 59: 195–206 (2008).

[23](#). Because of their perceived pejorative implications, this group of disorders, formerly labeled as forms of hysteria, a term with a lot of baggage, has been renamed numerous times in history: functional neurological disorders, conversion disorders, or psychogenic disorders.

[24](#). See the engaging book on functional neurological disorders by the neurologist S. O’Sullivan, *The Sleeping Beauties and Other Stories of Mystery Illness*. Pantheon: New York (2021).



## Chapter 4: Consciousness and the Physical

1. While historical confusions within philosophy have been cleared up, antecedent assumptions for premises are better understood, and connections between concepts and ideas have been clarified, there is no communal convergence on the perennial “big questions.” Close to two thousand professors in university philosophy departments in North America, Europe, and Australasia were asked to fill out a survey on where they stood on thirty philosophical questions. Of these 931 responded. Only one view attracted a high consensus (82 percent): nonskeptical realism about the external world. But even here, about one in five philosophers believe either that the external world is a manifestation of mind or that one cannot know anything about it. Other questions have 60 percent or less support for the leading interpretation. D. Bourget and D. J. Chalmers, “What do philosophers believe?” *Philosophical Studies* 170: 465–500 (2014).

2. Early activation of either the innate or the adaptive immune system proceeds in the absence of any interoceptive conscious experience. You have no direct conscious access to your immune system, a highly sophisticated set of interlocking information-processing and memory mechanisms (e.g., antibodies). Why not? Where is the difference between the immune and the nervous systems? Eventually, of course, inflammation and its attendant redness, swelling, heat, discomfort, and fever will set in, and you will feel lousy, “like you’re coming down with something.” This question calls for a well-designed and well-controlled experiment in which a harmless virus is introduced into volunteers to ascertain the timing of onset of awareness for inflammation or other bodily reactions.

3. During Descartes’s life, the Dominican friar Giordano Bruno was burned at the stake in Rome in 1600 for speculating too freely, and his countryman Galileo Galilei was placed under lifelong house arrest in 1633 for being suspected of heresy.

4. Because of his insistence that everything must be constituted by either a physical or a mental substance, Cartesian dualism is also known as *substance dualism*.

5. In the latest survey of the physical mass-energy content of the known universe, two inferred components, dark energy (68 percent) and dark



matter (27 percent), dominate. The rest is free hydrogen and helium atoms (4 percent), stars (0.5 percent), which mainly consist of hydrogen and helium, neutrinos (0.5 percent), and heavy elements, that is, all chemical elements except for hydrogen and helium. In terms of mass-energy, planet Earth and all of life are rounding errors, making up less than one-thousandth of all stuff in the universe!

6. I warmly recommend reading Descartes's slim *Meditations on First Philosophy* for his epistemological arguments, based on what the mind knows, and metaphysical arguments, based on the essential differences between the two substances. Written in 1640, it carries a mighty punch and has never been out of print.

7. Point 17 in Leibniz's *The Principles of Philosophy* or *The Monadology* from 1714.

8. Their extensive correspondence, running to fifty-nine surviving letters, is found at "Correspondence between Descartes and Princess Elisabeth," Early Modern Texts: [https://www.earlymoderntexts.com/assets/pdfs/descartes1643\\_1.pdf](https://www.earlymoderntexts.com/assets/pdfs/descartes1643_1.pdf).

9. The philosopher of science Karl Popper and the neurophysiologist John Eccles, who was awarded a Nobel Prize in 1963 for his elucidation of the all-or-none nature of synaptic transmission between nerve cells, argued that the causal influence of the mental on the brain is camouflaged by quantum mechanical uncertainty and that this leaves sufficient room for mind to interfere with synaptic release in the cortex. Their proposal has received little endorsement by the scientific community in the absence of evidence for such mental action. K. R. Popper and J. C. Eccles, *The Self and Its Brain*. Springer: Berlin (1977).

10. Or, more general but less melodious: "no substrate, never mind." This implies that the *bardo*, an intermediate existence between death and rebirth in Buddhist thought, must have some physical manifestation in space and time for it to retain some aspects of the soul of the dead person, including the deceased's accumulated karma. W. Hasenkamp and J. R. White, eds., *The Monastery and the Microscope*. Yale University Press: New Haven, CT (2017).

11. The technical expression is "consciousness is supervenient on the physical," in the sense that any change in the mental entails, of necessity, a change in the underlying physical. Any conscious experience must go hand

in hand with changing neural activity in the brain. Think of the teleporter on the USS *Enterprise* in the *Star Trek* universe. This device accurately scans a body, at the relevant level of resolution, beaming this information to a matching teleporter station somewhere else to reassemble the person there (to avoid plot complications, the original body disappears). If supervenience holds, the conscious mind of the beamed person must be identical to the original, on the assumption that the copy is exact. That is, if two brains are physically identical, down to the state of their constitutive elements, their mental state must be identical as well.

12. The term *physicalism* is of recent vintage, having been proposed by members of the Vienna circle in the 1920s to distinguish proper linguistic statements that ultimately must be grounded in physical facts from nonsensical ones (in their eyes) that are ungrounded, such as “God exists” or “God does not exist.”

13. To avoid further complications when my bike and its observer don’t share the same accelerating frame of reference.

14. D. Garisto, “The universe is not locally real, and the physics Nobel Prize winners proved it.” *Scientific American*: <https://www.scientificamerican.com/article/the-universe-is-not-locally-real-and-the-physics-nobel-prize-winners-proved-it> (2022). That a conscious observer is required to collapse the superposition of states of a quantum system into the single observable outcome has troubled physicists since way back, starting with John von Neumann’s comment in his 1932 textbook on quantum mechanics that “experience only asserts something like: an observer has made a certain (subjective) perception, but never such as: a certain physical quantity has a certain value.”

15. R. Penrose, *The Emperor’s New Mind*. Oxford University Press: Oxford (1989); R. Penrose, *Shadows of the Mind*. Oxford University Press: Oxford (1994). Penrose and the anesthesiologist Stuart Hameroff argue that microtubules, part of the molecular scaffolding of cells, are a critical locus of such interactions. See S. Hameroff and R. Penrose, “Consciousness in the universe: A review of the ‘Orch OR’ theory.” *Physics Life Reviews* 11: 39–78 (2014).

16. The extent to which organisms exploit quantum resources, in particular entanglement and superposition, is a tantalizing and open experimental question. See J. Cao et al., “Quantum biology revisited.”

*Science Advances* 6: 1–11 (2020); S. Gao, ed., *Consciousness and Quantum Mechanics*. Oxford University Press: Oxford (2022); H.-Y. Huang et al., “Quantum advantage in learning from experiments.” *Science* 376: 1182–1186 (2022).

[17](#). The best-known example of reductive physicalism is behaviorism, a school of thought that emerged from psychology. It postulated that understanding animals and humans should be limited to observable phenomena, such as sensory stimulus and behavioral response. Subjective experience was, and remains, either deliberately marginalized or completely denied. See the writings by husband and wife philosophers Patricia Churchland, *Neurophilosophy: Toward a Unified Science of the Mind/Brain*. MIT Press: Cambridge, MA (1986), and Paul Churchland, *Matter and Consciousness: A Contemporary Introduction to the Philosophy of Mind*. MIT Press: Cambridge, MA (1984), as well as G. Rey, “Reasons for doubting the existence of even epiphenomenal consciousness.” *Behavioral and Brain Sciences* 14: 691–692 (1991); K. Frankish, “The consciousness illusion. Phenomenal consciousness is a fiction written by our brains to help us track the impact that the world makes on us.” *Aeon*, NW, Ed, 26. (2019); M. S. Graziano, “Understanding consciousness.” *Brain* 144: 1281–1283 (2021).

[18](#). “Philosophy that stirs the waters.” *New York Times* (April 29, 2013). Dennett argues in *Consciousness Explained* (Little, Brown: Boston; 1991) that people are terribly confused. What they really mean when they speak about consciousness is that they have certain beliefs about mental states; each one has distinct functional properties with distinct behaviors and affordances. These need to be explained. Pain and redness are illusory; there is nothing intrinsic about them; consciousness is all in the doing. F. Fallon, “Dennett on Consciousness: Realism without the hysterics.” *Topoi* 39: 35–44 (2020). Variants of this intellectual position are eliminative materialism, fictionalism, and instrumentalism.

[19](#). G. Strawson, *Mental Reality*. MIT Press: Cambridge, MA (1994).

[20](#). V. Taschereau-Dumouchel et al., “Putting the ‘mental’ back in ‘mental disorders’: A perspective from research on fear and anxiety.” *Molecular Psychiatry* 27: 1322–1330 (2022). As noted in the previous chapter, ascribing depression to an underlying chemical imbalance, a deficit in serotonin levels, has been unproductive. SSRI lead to addiction, named

without irony “discontinuation syndrome.” P. Sperling, “Causality in mental disturbance: A review of neuroscience.” *Mad in America*: <https://www.madinamerica.com/2023/07/causality-mental-disturbance> (2023).

21. H. Putnam, “The nature of mental states.” In *The Philosophy of Mind: Classical Problems/Contemporary Issues*, pp. 223–231. Harvard University Press: Cambridge, MA (1992).

22. The nervous system is obviously not a conventional digital computer: it operates in parallel, without a system-wide clock or bus; its elements switch at a glacial speed of milliseconds, memory and processing are not separate, and it uses mixed analog and digital signals.

23. That is, the phenomenal or the mental is supervenient on the physical.

24. There are numerous variants of such knowledge and/or conceivability arguments involving Mary, the color vision scientist, inverted qualia, and zombies. As their logic is similar, I focus on the zombie argument.

25. D. J. Chalmers, “The puzzle of conscious experience.” *Scientific American* 273: 80–86 (1995). For a collection of articles responding to Chalmers’s book, see J. Shear, ed., *Explaining Consciousness: The Hard Problem*. MIT Press: Cambridge, MA (1997).

26. B. Kastrup, H. P. Stapp, and M. C. Kafatos, “Coming to grips with the implications of quantum mechanics.” *Scientific American*: <https://blogs.scientificamerican.com/observations/coming-to-grips-with-the-implications-of-quantum-mechanics> (2018).

27. D. Dennett, “Current issues in the philosophy of mind.” *American Philosophical Quarterly* 15: 249–261 (1978), p. 252.

28. W. Jaworski, *Philosophy of Mind: A Comprehensive Introduction*. Wiley-Blackwell: Oxford, UK (2011).

29. B. Kastrup, *Science Ideated: The Fall of Matter and the Contours of the Next Mainstream Scientific Worldview*. Iff Books: Hampshire, UK (2021); and the forthcoming B. Kastrup, *Analytic Idealism in a Nutshell*. Iff Books: Hampshire, UK (2024).

30. D. F. Skrbina, *Panpsychism in the West*. Rev. ed. MIT Press: Cambridge, MA (2017). Recent popular books on the topic include one by a philosopher, P. Goff, *Galileo’s Error*. Pantheon Books: New York (2019),

and another by a writer, A. Harris, *Conscious: A Brief Guide to the Fundamental Mystery of the Mind*. HarperCollins: New York (2019). For the latest skirmishes and debates, see P. Goff and A. Moran, eds., *Is Consciousness Everywhere? Essays on Panpsychism*. Academic Imprints: Exeter, UK (2022). For links to the Indian literature on this topic, see B. P. Göcke and S. Medhananda, eds., *Panentheism in Indian and Western Thought: Cosmopolitan Interventions*. Taylor & Francis: New York (2023).

[31](#). See the 1937 classic *Star Maker* by Olaf Stapledon, one of the most imaginative science fiction novels ever penned, expounding on the idea of collective minds, including stars and stellar nebulae, that are linked by telepathy, eventually encompassing the entire cosmos.

## Chapter 5: What Truly Exists

1. See <http://integratedinformationtheory.org/> and the associated python code (Py Phi) on GitHub.

2. To understand this difference, see I. Cea et al., “The fundamental tension in integrated information theory 4.0’s realist idealism.” *Entropy* 25: 1453 (2023), as well as the blog post by philosopher Jonathan Birch and his claim that the philosophical precursor of IIT’s ontology is the idealism of Leibniz. J. Birch, “Consciousness and the Overton window of science, part II.” *The Brains Blog*: <https://philosophyofbrains.com/2023/09/12/consciousness-and-the-overton-window-of-science-part-ii.aspx> (September 12, 2023).

3. The simulation hypothesis defers the need to explain this universe to the need to explain the hypothetical universe at the next level up, the one simulating our cosmos. Explanations can be deferred indefinitely by postulating a countable infinite number of universes, each simulating the one below it. This is all quite fun to think about and the topic of interesting, if dystopian, science fiction stories (“Valuable Humans in Transit and Other Stories” by qntm, 2021) but about as useful in explaining the world as invoking the “it’s turtles all the way down” line.

4. Luminiferous, or light-bearing, aether was introduced into classical physics in the nineteenth century to explain how light and electromagnetic waves can propagate through empty space. This required an invisible fluid, filling all of space, vastly more rigid than steel, that wouldn’t interact with ordinary physical objects. When more and more experiments concluded that, whatever the aether might be, it had no effect whatsoever, it finally fell under Occam’s razor and was quietly retired. As aether has no causal power, it plays no role in modern physics.

5. Causal manipulations are easier in biology than in astronomy. The latter is a quintessential observational science, whose objects are very far away, move very rapidly, and are very big. This only changed in 2022, when NASA deliberately crashed a space probe into a small asteroid to perturb its orbit to test the possibility of planetary defenses. Fortunately, the objects brain science studies are close by, stationary, and small. Thus neuroscientists don’t need to crash their apparatus each time they probe the

brain.

[6](#). Plus the weak and strong nuclear forces, although they only become relevant at extremely short distances.

[7](#). I visited India in 2016 when the government abruptly announced that five-hundred- and one-thousand-rupee banknotes would cease to be legal tender and had to be exchanged to combat tax evasion and curtail the shadow economy. While I traveled, there was uncertainty around what these banknotes could buy (i.e., their effective causal power was ill defined).

[8](#). The *Eleatic principle* makes its appearance in Plato's *Sophist*, a dialogue between a mathematician and a stranger from the Greek settlement of Elea in southern Italy. At some point, the stranger states, "My notion would be, that anything which possesses any sort of power to affect another, or to be affected by another, if only for a single moment, however trifling the cause and however slight the effect, has real existence; and I hold that the definition of being is simply power." In integrated information theory, Plato's "or" is replaced by a stronger "and" requirement. The causal interaction must flow in both directions, to "and" from the system in question.

[9](#). For example, Newton's second law of motion,  $F = ma$ , could be reformulated in this operational manner. If you apply this force to this mass, it will accelerate by this amount. Of course, one also has to account for the continuous nature of the relevant variables, here mass and acceleration, compared to the binary, on and off, states of neurons or transistors.

[10](#). The latest and most up-to-date reference to integrated information theory is L. Albantakis et al., "Integrated information theory (IIT) 4.0: Formulating the properties of phenomenal existence in physical terms." arXiv:2212.14787 (2022). Easier to read is G. Tononi et al., "Only what exists can cause: An intrinsic powers view of free will." arXiv:2206.02069 (2023). For a lyrical introduction into the theory, I warmly recommend G. Tononi, *Phi: A Voyage from the Brain to the Soul*. New York: Pantheon Books (2012).

[11](#). From E. Schrödinger, "The Oneness of Mind." In K. Wilber, ed., *Quantum Questions: Mystical Writings of the World's Great Physicists*, pp. 79–101. Shambhala: Boston (1984).

[12](#). In hindsight, it is not apparent whether the white-blue light of overwhelming intensity in my experience was positioned within a spatially



extended background without any other notable features or all of space had collapsed into a singularity.

[13](#). An interesting variant of these axioms postulates that a conscious system must be nonseparable, subsuming the integration and the exclusion axioms. A. Arkhipov, “Non-separability of physical systems as a foundation of consciousness.” *Entropy* 24: 1539 (2022).

[14](#). Strictly speaking, IIT calls this the physical substrate of consciousness. I here elide the difference to the neural, sometimes also called the neuronal, correlates of consciousness. These include both anatomical structures and physiological processes. G. Tononi et al., “Integrated information theory: From consciousness to its physical substrate.” *Nature Review Neuroscience* 17: 450–461 (2016).

[15](#). When investigating whether any one candidate brain circuit is a potential substrate of consciousness, one must keep the vascular, neuromodulatory, and other background conditions fixed while perturbing individual neurons or groups of them making up the candidate circuit.

[16](#). Also known as six degrees of separation, it embodies the notion that any two people on the planet are interconnected by six or fewer friends, relatives, or acquaintances. It is a statistical property of networks with high connectivity, in the sense that it is true of most people except for the odd recluse.

[17](#). The principle of least action, central to modern physics, first formulated in 1740, determines how a dynamic system changes between two points in time,  $t_1$  and  $t_2$ . It considers all the possible ways the system could evolve between these two instants and picks the one trajectory that minimizes a function called *action*, closely related to the energy of the system. Of all paths the system could take, it takes the one that keeps action to a minimum.

[18](#). The exclusion postulate, grounded in the corresponding exclusion axiom, embodies a principle of maximal existence. What exists is what exists the most.

[19](#). Despite the intuition of most scientists that the relevant grain and spatiotemporal scale is ultimately the most reductionistic one, at the bottom—say, microscopic elementary particles—this does not have to be true given the complicated trade-offs between indeterminacy and degeneracy. See E. P. Hoel et al., “Can the macro beat the micro? Integrated information



across spatio-temporal scales.” *Neuroscience of Consciousness* 2016(1): niw012 (2016).

[20](#). For ten units, the total number of relations among distinctions is  $2^{1,000}$ , which is on the order of  $10^{300}$ , vastly more than the estimated  $10^{80}$  atoms in the known universe. A. Haun and G. Tononi, “Why does space feel the way it does? Towards a principled account of spatial experience.” *Entropy* 21: 1160 (2019), unfold all distinctions and relationship of eight units, wired together in a near-neighbor manner, and explain how this unfolded cause-effect structure satisfies all regularities of phenomenally experienced space spanned by eight nearby points (such as neighborhood, inclusion and exclusion relationships, distance, and so on). See also A. Zaeemzadeh and G. Tononi, “Upper bounds for integrated information.” arXiv:2305.09826 (2023).

[21](#). See the chapter titled “The Über-Mind and Pure Consciousness” in my 2019 book *The Feeling of Life Itself*.

[22](#). The implications of this form of panpsychism and how it solves the combination problem alluded to in the last chapter are described in G. Tononi and C. Koch, “Consciousness: here, there and everywhere?” *Philosophical Transactions of the Royal Society B* 370: 20140167 (2015). The absence of any threshold for integrated information, distinguishing conscious from unconscious states, does not preclude an extreme degree of nonlinearity for the hyperastronomical number of relations and distinctions. That is, large brains, such as ours, can have a staggeringly larger number of distinctions and relations than somewhat smaller brains, so that their unfolded cause-effect power is dim in comparisons to ours.

## Chapter 6: Consciousness and the Brain

1. A striking exception to the widespread neglect of the brain in the classical world is the medical treatise *On the Sacred Disease*. Written around 400 BCE, by Hippocrates or someone close to him, it states, “Men ought to know that from nothing else but the brain come joys, delights, laughter and sports, and sorrows, griefs, despondency, and lamentations.” The most influential ancient anatomist was the second-century physician Galen, whose clinical knowledge derived from working at a gladiator school in Rome and who emphasized the role of the brain in controlling rational thought. G. G. Gross, *Brain, Vision, Memory—Tales in the History of Neuroscience*. MIT Press: Cambridge, MA (1998).

2. Today’s language reflects this cardiac-laden bias: You love with all your heart rather than with your hypothalamus. There are hundreds of sacred heart churches and schools but not a single sacred brain academy.

3. C. Zimmer, *Soul Made Flesh: The Discovery of the Brain*. Free Press: New York (2004).

4. R. Sender, S. Fuchs, and R. Milo, “Revised estimates for the number of human and bacteria cells in the body.” *PLoS Biology* 14: e1002533 (2016), provides the number of cells in a seventy-kilogram adult. The same person also hosts tens of trillions of beneficial bacteria, the *microbiome*, in their alimentary tract, including the mouth, lungs, and gut. All of this to support the successful transmission of the sex cells to the next generation.

5. We now know that there are between three and five thousand different types of brain cells. Yes, three thousand to five thousand cell types. Some of these are not neurons proper but glial, immune (microglia and perivascular macrophages), and blood-vessel-associated cells. The Allen Institute is leading the worldwide effort to comprehensively map this vast diversity of cell types in the mammalian brain. To get a whiff of this stupendous complexity, visit the Allen Brain atlas at: <https://www.brain-map.org>; or see K. Siletti et al., “Transcriptomic diversity of cell types across the adult human brain.” *Science* 382, eadd7046 (2023); and Z. Yao et al., “A high resolution transcriptomic and spatial atlas of cell types in the whole mouse brain.” bioRxiv: 10.1101/2023.03.06.531121 (2023).

6. This involved writing computer code to solve the partial differential

equations describing the dynamics of voltage changes within nerve cells under the mentorship and guidance of my doctor-father, a brilliant computational scientist, Tomaso Poggio. I was immensely proud when part of my thesis appeared in the world's oldest continuously published scientific journal. C. Koch, T. Poggio, and V. Torre, "Retinal ganglion cells: A functional interpretation of dendritic morphology." *Philosophical Transactions of the Royal Society B* 298: 227–264 (1982).

7. This includes blindsight (L. Weiskrantz, *Blindsight: A Case Study and Implications*. Oxford University Press: Oxford [1986]); the split-brain patients of Roger Sperry, who was awarded the 1981 Nobel Prize for this work (R. W. Sperry, "Lateral specialization in the surgically separated hemispheres." In F. O. Schmitt and F. G. Worden, eds., *Neuroscience 3rd Study Program*. MIT Press: Cambridge [1974]; L. J. Volz and M. S. Gazzaniga, "Interaction in isolation: 50 years of insights from split-brain Research." *Brain* 140: 2051–2060 [2017]); and amnesic patients, such as HM, who lay down nonconscious motor memories while being unable to consciously recall what they learned (L. R. Squire, "The legacy of patient HM for neuroscience." *Neuron* 61: 6–9 [2009]). See the very engaging primer by S. Blackmore and E. T. Troscianko, *Consciousness: An Introduction*. Routledge: New York (2018).

8. F. C. Crick and C. Koch, "What is the function of the claustrum." *Philosophical Transactions of the Royal Society B* 360: 1271–1279 (2005). For a review of the claustrum, see J. B. Smith, A. K. Lee, and J. Jackson, "The claustrum." *Current Biology* 30: R1401–R1406 (2020).

9. F. C. Crick and C. Koch, "Towards a neurobiological theory of consciousness." *Seminars in the Neurosciences* 2: 263–275 (1990); F. C. Crick and C. Koch, "Some reflections on visual awareness." *Cold Spring Harbor Symposium on Quantitative Biology* 55: 953–962 (1990). It is likely that neuronal spiking, and the associated gamma band activity in the EEG, is a signature of selective attention, which is closely allied with, but different from, consciousness. S. M. Miller, *The Constitution of Phenomenal Consciousness*. Benjamins: Amsterdam (2015). Note that the NCC can refer to either structural elements, such as brain regions or neuronal cell types, or to dynamic processes whose absence impairs specific aspects of consciousness.

10. I'm here speaking at the conceptual level—studying and assessing

pain are complicated topics that involve many tools besides magnetic scanners. See D. C. Turk and R. Melzack, eds., *Handbook of Pain Assessment*. 3rd ed. Guilford Press: New York (2011).

[11](#). Of course, if you believe that consciousness is an illusion, looking for the footprints of this illusion is a wasted effort. For a rigorous definition of the NCC, see D. J. Chalmers, “What is a neural correlate of consciousness?” In T. Metzinger, ed., *Neural Correlates of Consciousness: Empirical and Conceptual Questions*, pp. 17–39. MIT Press: Cambridge, MA (2000). For a personal account, see the Q&A with David Chalmers in *Neuron* 111: 3341–3343 (2023).

[12](#). There is a significant research program dedicated to understanding the psychological and neuronal differences between selectively attending to an object or event on the one hand and becoming conscious of the object or event on the other. C. Koch and N. Tsuchiya, “Attention and consciousness: Two distinct brain processes.” *Trends in Cognitive Sciences* 11: 16–22 (2007). How to isolate the NCC from events that precede or follow it in time is detailed by J. Aru et al., “Distilling the neural correlates of consciousness.” *Neuroscience and Biobehavioral Reviews* 36: 737–746 (2012), and N. Tsuchiya et al., “No-report paradigms: Extracting the true neural correlates of consciousness.” *Trends in Cognitive Sciences* 19: 757–770 (2015).

[13](#). R. C. Coghill, J. G. McHaffie, and Y. F. Yen, “Neural correlates of interindividual differences in the subjective experience of pain.” *Proceedings of the National Academy of Sciences of the United States of America* 100: 8538–8542 (2003); T. D. Wager et al., “An fMRI-based neurologic signature of physical pain.” *New England Journal of Medicine* 368: 1388–1397 (2013).

[14](#). I can warmly recommend J. Stegenga, *Medical Nihilism*. Oxford University Press: Oxford, UK (2018), as well as J. Pearl and D. Mackenzie, *The Book of Why: The New Science of Cause and Effect*. Basic Books: New York (2018).

[15](#). This culture of constant experimentation on randomly selected users and comparing against a control group allowed Amazon to innovate effectively and quickly. This usually trumped the Highest Paid Person’s Opinion (HiPPO). R. Kohavi et al., “Controlled experiments on the web: Survey and practical guide.” *Data Mining and Knowledge Discovery* 18:

140–181 (2009).

16. I. Fried et al., “Laser ablation of human guilt.” *Brain Stimulation* 15: 164–166 (2022).

17. J. Parvizi et al., “Electrical stimulation of human fusiform face-selective regions distorts face perception.” *Journal of Neuroscience* 32: 14915–14920 (2012); V. Rangarajan and J. Parvizi, “Functional asymmetry between the left and right human fusiform gyrus explored through electrical brain stimulation.” *Neuropsychologia* 83: 29–36 (2016); J. Jonas et al., “A face identity hallucination (palinopsia) generated by intracerebral stimulation of the face-selective right lateral fusiform cortex.” *Cortex* 99: 296–310 (2018); Y. H. Koh, “Right fusiform gyrus infarct with acute prosopagnosia.” *Acta Neurologica Taiwanica* 31:183–184 (2022); O. Blanke et al., “Stimulating illusory own-body perceptions.” *Nature* 419: 269–270 (2002).

18. Sigmund Freud proposed a special class of neurons responsible for conscious experience in his unpublished 1895 “Project for a Scientific Psychology.” For the modern history of the NCC program, see the special issue by S. B. Fink, “A double anniversary for the neural correlates of consciousness: Editorial introduction.” *Philosophy and Mind Sciences* 1(2): <https://doi.org/10.33735/phimisci.2020.II.85> (2020).

19. The bet originated at the second annual meeting of the Association for the Scientific Study of Consciousness (ASSC) in Bremen, Germany, in June 1998 during a late-night bar visit. P. Snaprud, “The consciousness wager.” *New Scientist* 238: 28–31 (2018). David Chalmers, myself, and a few other psychologists, neuroscientists, and philosophers co-founded the ASSC a few years earlier, as a focal point for empirical and theoretical work on consciousness. The bet was resolved at the twenty-sixth meeting of the ASSC in New York in June 2023.

20. To be precise, unconsciousness occurs in 6.8 seconds. This was established by blocking the internal carotid artery via an inflatable cervical pressure cuff in 126 healthy young men. R. Rossen, H. Kabat, and J. P. Anderson, “Acute arrest of cerebral circulation in man.” *Archives of Neurology and Psychiatry* 50: 510–528 (1943). Loss of consciousness is accompanied by fixation of the eyeballs, blurred vision, constriction of the visual fields, and, in some, convulsions. The authors nonchalantly comment, “No deleterious effects were observed from repeated tests on

these subjects.” The rapid return of consciousness can be accompanied by intense visual hallucinations and euphoric feelings. Such side effects have been corroborated following acceleration-induced fainting in volunteers spun in centrifuges at high g’s. J. E. Whinnery and A. M. Whinnery, “Acceleration-induced loss of consciousness.” *Archives of Neurology* 47: 764–776 (1990).

21. The brainstem houses at least forty distinct groups of neurons in a region called the ascending activating reticular formation. Each population has its own neurotransmitter, such as glutamate, acetylcholine, serotonin, noradrenaline, GABA, histamine, adenosine, and orexin, which modulates, either directly or indirectly, the excitability of cortex and other forebrain structures. Collectively they access and control signals relating to the internal milieu: arousal breathing, thermal regulation, sleep and wakefulness, eye muscles, the musculoskeletal frame, and other essential functions. J. Parvizi and A. R. Damasio, “Consciousness and the brainstem.” *Cognition* 79: 135–159 (2001); S. Laureys, O. Gosseries, and G. Tononi, eds., *The Neurology of Consciousness*. 2nd ed. Elsevier: Amsterdam (2015).

22. P. House, *Nineteen Ways of Looking at Consciousness*. St. Martin’s Press: New York (2022); F. Karinty, *A Journey Round My Skull*. Translated by Vernon Duckworth Barker. NYRB Classics: New York (2008).

23. One prominent case is that of a woman who has a cavern, filled with cerebral spinal fluid, in the place where her cerebellum should be. Born without one (cerebellar agenesis), she has mild mental impairment, slurred speech, and some motor deficits, yet leads an unremarkable life with a young daughter. F. Yu et al., “A new case of complete primary cerebellar agenesis: Clinical and imaging findings in a living patient.” *Brain* 138: 1–5 (2014). Other cases are described in C. A. Boyd, “Cerebellar agenesis revisited.” *Brain* 133: 941–944 (2010), and R. N. Lemon and S. A. Edgley, “Life without a cerebellum.” *Brain* 133: 652–654 (2010).

24. Of the eighty-six billion neurons in the human brain, sixty-nine billion are in the cerebellum, and sixteen billion are found in the cortex. F. Azevedo et al., “Equal numbers of neuronal and non-neuronal cells make the human brain an isometrically scaled-up primate brain.” *Journal of Comparative Neurology* 513: 532–541 (2009); S. Walloe, B. Pakkenberg,



and K. Fabricius, “Stereological estimation of total cell numbers in the human cerebral and cerebellar cortex.” *Frontiers in Human Neuroscience* 8: 508–518 (2014).

[25](#). This circuit architecture is reminiscent of a look-up table for sensory-motor transformations.

[26](#). The word *cortex* derives from the Latin for bark, rind, or shell, denoting the outermost layers of the brain. The neocortex, together with the evolutionarily older allocortex, primarily the hippocampus and the olfactory cortex, form the cerebral cortex. The neocortex, sometimes referred to as the cortex, operates in close conjunction with several satellite structures: the thalamus, the basal ganglia, and the claustrum, all part of the forebrain. Given the heavy bidirectional connections between the cortex and these structures, it is no easy matter to dissociate their specific contributions.

[27](#). C. Koch et al., “The neural correlates of consciousness: Progress and problems.” *Nature Reviews Neuroscience* 17: 307–321 (2016); M. Boly et al., “Are the neural correlates of consciousness in the front or in the back of the cerebral cortex? Clinical and neuroimaging evidence.” *Journal of Neuroscience* 37: 9603–9613 (2017); B. Kozuch, “A legion of lesions: The neuroscientific rout of higher-order thought theory.” *Erkenntnis*: 1–27 (2023).

[28](#). Why should some cortical regions be more privileged vis-à-vis consciousness than others? It could either be that their constitution (i.e., the makeup of their cells) or their structure differs (or both). Indeed, a striking difference between the front and the back of the neocortex is their local synaptic connectivity—in visual, auditory, and somatosensory cortices in the back, connectivity is arranged topographically, linking nearby neurons in a grid-like topography, appropriate for maps of seen, heard, or felt percepts that are located in space, while the front of the brain has more of a random-access connectivity, possibly compatible with more abstract thinking. The former will lead to extremely large values of integrated information  $\Phi$  while the latter has much less integrated information. A. Zaeemzadeh and G. Tononi, “Upper bounds for integrated information.” arXiv:2305.09826 (2023); C. Koch et al., “Posterior and anterior cortex: Where is the difference that makes the difference?” *Nature Reviews Neuroscience* 17: 666 (2016).

[29](#). M. J. Farah, *Visual Agnosia*. MIT Press: Cambridge, MA (1990); S.



Zeki, *A Vision of the Brain*. Oxford University Press: Oxford (1993); C. A. Heywood and J. Zihl, “Motion blindness.” In G. W. Humphreys, ed., *Case Studies in the Neuropsychology of Vision*, pp. 1–16. Psychology Press/Taylor & Francis: London (1999). The neurologist Oliver Sacks wrote eloquently about individuals suffering from agnosia and the singular way they experience life; see his book *The Man Who Mistook His Wife for a Hat*. Agnosia is not caused by a defect in the sensory periphery nor in the motor output but by a loss of central processing capabilities.

[30](#). T. E. Feinberg et al., “Two alien hand syndromes.” *Neurology* 42: 19–24 (1992), describes a patient whose left hand had a will of its own, choking her. It took great strength on her part to pull her left hand off her throat. Another patient’s right hand had a prominent grasp response and was in constant motion. It groped nearby objects, including bedclothes or the patient’s own leg or genitals, and did not release them.

[31](#). In Anton-Babinksi or Anton’s syndrome, a rare case of visual anosognosia or cortical blindness, patients insist that they see normally while being completely blind; when asked, say, to count the fingers on the examiner’s hand, they confabulate—a stunning manifestation of the lack of insight into one’s condition. M. Maddula, S. Lutton, and B. Keegan, “Anton’s syndrome due to cerebrovascular disease: A case report.” *Journal of Medical Case Reports* 3: 1–3 (2009).

[32](#). Of course, you can still accurately pinpoint sounds coming from behind your back. But that involves auditory and not visual cues.

[33](#). A. Kertesz, “Anosognosia in aphasia.” In G. P. Prigatano, ed., *The Study of Anosognosia*, pp. 113–122. Oxford University Press: Oxford (2010).

[34](#). A. Henri-Bhargava, D. T. Stuss, and M. Freedman, “Clinical assessment of prefrontal lobe functions.” *Behavioral Neurology and Psychiatry* 24: 704–726 (2018).

[35](#). Chapter 11 in C. Koch, *The Feeling of Life Itself*. MIT Press: Cambridge, MA (2018).

[36](#). K. C. R. Fox et al., “Intrinsic network architecture predicts the effects elicited by intracranial electrical stimulation of the human brain.” *Nature Human Behaviour* 4: 1039–1052 (2020). See also C. Koch, “Hot or Not?” *Nature Human Behaviour* 4: 991–992 (2020). For pain, see A. Duong et al., “Subjective states induced by intracranial electrical stimulation

matches the cytoarchitectonic organization of the human insula.” *Brain Stimulation* 16: 1653–1665 (2023); J. Isnard et al., “Does the insula tell our brain that we are in pain?” *Pain* 152: 946–951 (2011); L. Mazzola, “Stimulation of the human cortex and the experience of pain: Wilder Penfield’s observations revisited.” *Brain* 135: 631–640 (2012); A. Montavont et al., “On the origin of painful somatosensory seizures.” *Neurology* 84: 594–601 (2015).

[37](#). The epileptic patient with seizures is described in J. Parvizi et al., “Altered sense of self during seizures in the posteromedial cortex.” *Proceedings of the National Academy of Sciences of the United States of America* 118(29): e2100522118 (2021). The bodily or spatial “I” is further localized to the anterior precuneus, part of the posteromedial cortex: D. Lyu et al., “Causal evidence for the processing of bodily self in the anterior precuneus.” *Neuron* 111(16): 2502–2512.e4 (2023). Imaging studies in long-term meditators and in naïve volunteers undergoing mindfulness training in which the omnipresent sense of self is blunted likewise show reduced activity in this neighborhood. J. A. Brewer et al. “Meditation experience is associated with differences in default mode network activity and connectivity.” *Proceedings of the National Academy of Sciences of the United States of America* 108: 20254–20259 (2011). J. A. Brewer, K. A. Garrison, and S. Whitfield-Gabrieli, “What about the ‘self’ is processed in the posterior cingulate cortex?” *Frontiers in Human Neuroscience* 7: 647 (2013).

[38](#). Take a recent imaging experiment that confirms the import of the posterior hot zone to conscious visual perception. Subjects had to distinguish briefly flashed line art of animals from line art of objects by pressing a letter on a keyboard. This evoked hemodynamic activity in the back as well as the front of the neocortex. When subjects did not have to do anything but simply look (no-task condition), the same drawings still evoked selective activity in posterior but not anterior cortex. This well-controlled study intermixed masked and nonmasked visual stimuli with a third, incidental memory task. E. Hatamimajoumerd et al., “Decoding perceptual awareness across the brain with a no-report fMRI masking paradigm.” *Current Biology* 32: 4139–4149 (2022).

[39](#). L. Melloni et al., “Making the hard problem of consciousness easier.” *Science* 372: 911–912 (2021). Note that this is only the first of five

such adversarial collaborations to test theories of consciousness. All were initiated by Dawid Potgieter and funded by the Templeton World Charity Foundation. These adversarial collaborations can also be considered an experiment in the sociology of neuroscience.

[40](#). S. Dehaene and J.-P. Changeux, “Experimental and theoretical approaches to conscious processing.” *Neuron* 70: 200–227 (2011); S. Dehaene, *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts*. Viking: New York (2014); B. van Vugt et al., “The threshold for conscious report: Signal loss and response bias in visual and frontal cortex.” *Science* 360: 537–542 (2018).

[41](#). A. M. Haun et al., “Are we underestimating the richness of visual experience?” *Neuroscience of Consciousness* 2017(1): niw023 (2017).

[42](#). This is a critical prediction for the theory: the NCC must be the maximum of intrinsic cause-effect power. If that is not the case, the theory is disproven.

[43](#). The preregistered methods for both experiments are described in L. Melloni et al., “An adversarial collaboration protocol for testing contrasting predictions of global neuronal workspace and integrated information theory.” *PLoS ONE* 18: e0268577 (2022).

[44](#). Cogitate Consortium et al., “An adversarial collaboration to critically evaluate theories of consciousness.” bioRxiv: <https://doi.org/10.1101/2023.06.23.546249> (2023).

[45](#). This was a mildly discouraging but vivid demonstration of Max Planck’s principle that “a new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die and a new generation grows up that is familiar with it.”

[46](#). M. Lenharo, “Decades-long bet on consciousness ends—and it’s philosopher 1, neuroscientist 0.” *Nature* (June 24, 2023); E. Finkel, “Adversarial search for neural basis of consciousness yields first results.” *Science* (June 25, 2023).

[47](#). As predicted by Daniel Kahneman, the psychologist who won the 2002 Nobel Prize in economic sciences and who formulated the concept of an *adversarial collaboration* to resolve conflict in psychology, some scientists whose theories are threatened by negative outcomes will mobilize scientific as well as methodological, philosophical, political, and other

reasonings, including *ad hominem* charges, to discredit results of previously agreed-upon protocols, turning productive science into *angry science*. D. Kahneman, “Experiences of collaborative research.” *American Psychologist* 58(9): 723 (2003). Here, these attacks on our collaboration took the form of claiming that integrated information theory is a pseudoscience. M. Lenharo, “Consciousness theory slammed as ‘pseudoscience’—sparking uproar.” *Nature*: <https://www.nature.com/articles/d41586-023-02971-1> (September 20, 2023). For a taste of what is wrong about this attempt to “cancel” an entire theory of consciousness, see E. Hoel’s blog post <https://iai.tv/articles/no-theory-of-consciousness-is-scientific-auid-2610>. In his view, which I share, IIT is so attractive to so many scholars because the theory is more ambitious and well-formalized than all the alternatives and actually tackles head-on the difficulties that any theory of consciousness faces. See also the insightful and sarcastic A. Gomez-Marin, “The consciousness of neuroscience.” *eNeuro* 10(11) (2023).

[48](#). On occasion, babies are born without any cerebral hemispheres and without much of a skull, scalp, and meninges due to a neural tube defect or other developmental mishap. Such anencephalic children are severely impaired, and many die young. Those who survive have massive disabilities, never learn to speak, and have limited expressive powers. The extent to which they are conscious is difficult to assess. Whatever limited experiences they do have would need to be supported by subcortical structures, such as the colliculus, that have adapted to this extraordinary situation. For a passionate defense of consciousness in children without a cortex, see B. Merker, “Consciousness without a cerebral cortex: A challenge for neuroscience and medicine.” *Behavioral and Brain Sciences* 30: 63–81 (2007); B. Aleman and B. Merker, “Consciousness without cortex: A hydranencephaly family survey.” *Acta Paediatrica* 103: 1057–1065 (2014).

[49](#). There is a significant experimental literature on complex behavior compatible with sentience in bees and other insects. A. B. Barron and C. Klein, “What insects can tell us about the origins of consciousness.” *Proceedings of the National Academy of Sciences of the United States of America* 113: 4900–4908 (2016); L. Chittka and C. Wilson, “Expanding consciousness.” *American Scientist* 107: 364–369 (2019).

[50](#). The assumption that all life is conscious is known as *biopsychism*.

See E. Thompson, *Mind in Life: Biology, Phenomenology and the Sciences of the Mind*. Harvard University Press: Cambridge, MA (2007).

[51](#). However, a recent information theoretical analysis comparing neocortical connectivity between monkeys and humans came to the surprising conclusion that a large region of human inferior parietal and posterior temporal cortex, precuneus, and to a lesser extent parts of the frontal cortex could not be predicted from any part of the macaque brain. The first three regions overlap strikingly with the posterior hot zone. See R. B. Mars et al., “Whole brain comparative anatomy using connectivity blueprints.” *Elife* 7: e35237 (2018).

[52](#). Cranial nerve palsies, brainstem lesions affecting sensory and motor pathways, cortical blindness, aphasia, and frontal akinetic syndrome can all interfere with a patient’s ability to see or hear the request or to act on it. Furthermore, fluctuating arousal levels, especially in the acute phase, complicate bedside examination, increasing the chances of missing recovery of volitional behaviors. Some vegetative state patients have been asked to imagine walking around their houses or thinking of playing tennis while lying inside a magnetic scanner. Four out of twenty-three such patients had the same brain responses in their hippocampus and their supplementary motor cortex as conscious volunteers. However, exploiting this willful modulation of brain activity as a two-way lifeline for communication does not work for most patients and is, in any case, infeasible in the ICU environment. See M. M. Monti et al., “Willful modulation of brain activity in disorders of consciousness.” *New England Journal of Medicine* 362: 579–589 (2010); A. Owen, *Into the Gray Zone: A Neuroscientist Explores the Border Between Life and Death*. Scribner: London (2017).

[53](#). F. Faucher et al., “Survival and consciousness recovery are better in the minimally conscious state than in the vegetative state.” *Brain Injury* 32: 72–77 (2018); J. Elmer et al., “Association of early withdrawal of life-sustaining therapy for perceived neurological prognosis with mortality after cardiac arrest.” *Resuscitation* 102: 127–135 (2016). See also J. J. Fins, *Rights Come to Mind: Brain Injury, Ethics, and the Struggle for Consciousness*. Cambridge University Press: New York (2015).

[54](#). The original conception and test of the idea of the perturbational complexity index (PCI) and consciousness in awake and sleeping

volunteers was by M. Massimini et al., “Breakdown of cortical effective connectivity during sleep.” *Science* 309: 2228–2232 (2005). For clinical data, see A. Casali et al., “A theoretically based index of consciousness independent of sensory processing and behavior.” *Science Translational Medicine* 5: 1–11 (2013); S. Casarotto et al., “Stratification of unresponsive patients by an independently validated index of brain complexity.” *Annals of Neurology* 80: 718–729 (2016); D. O. Sinitsyn et al., “Detecting the potential for consciousness in unresponsive patients using the perturbational complexity index.” *Brain Sciences* 10: 917 (2020). For a recent review of covert consciousness, see B. L. Edlow et al., “Measuring consciousness in the intensive care unit.” *Neurocritical Care* 38(3): 584–590 (2023).

## Chapter 7: Expanding Consciousness

1. See T. Metzinger, *The Ego Tunnel: The Science of the Mind and the Myth of the Self*. Basic Books: New York (2009). B. Hood, *The Self Illusion: How the Social Brain Creates Identity*. Oxford University Press: New York (2012); S. Blackmore, *Zen and the Art of Consciousness*. Simon and Schuster: New York (2014).

2. Other names for transformative experiences include “higher,” “altered,” “peak,” “revelatory,” “visionary,” “spiritual” or “cathartic” experiences.

3. Of course, myriad unexpected life events can transform life for the worse, such as sexual assault, stabbings, shootings, and all the rest of the long litany of men’s inhumanity (and it is almost always men). I will not discuss these (although even here, posttraumatic growth with all the hallmarks of a transformative experience can occur). Furthermore, I will only focus on abrupt changes, in comparison to slow conversions, with much deliberate and conscious reasoning, to a new identity or cause. A. Chirico et al., “Defining transformative experiences: A conceptual analysis.” *Frontiers in Psychology* 13: 2862 (2022).

4. A. N. Wilson, *Paul: The Mind of the Apostle*. Norton: New York (1997).

5. Mystical experiences can occur as part of temporal lobe epileptic seizures termed *Dostoyevsky’s seizures*, after the Russian writer who suffered from them and wrote eloquently about their effect. I. Fried, “Auras and experiential responses arising in the temporal lobe.” In S. Salloway, P. Malloy, and J. L. Cummings, eds., *The Neuropsychiatry of Limbic and Subcortical Disorders*, pp. 113–122. American Psychiatric Press: Washington, DC (1997). They can also be induced by electrical stimulation during neurosurgery. F. Bartolomei et al., “The role of the dorsal anterior insula in ecstatic sensation revealed by direct electrical brain stimulation.” *Brain Stimulation* 12: 1121–1126 (2019).

6. D. B. Yaden et al., “The overview effect: Awe and self-transcendent experience in space flight.” *Psychology of Consciousness: Theory, Research, and Practice* 3: 1 (2016).

7. J. Goodall and P. Berman, *Reason for Hope: A Spiritual Journey*.



Grand Central Publishing: New York (1999).

[8.](#) A. Schopenhauer, *The World as Will and Representation*. Volume 1, p. 178. Dover: New York (1969), originally published in 1819.

[9.](#) E. Herrigel, *Zen in the Art of Archery*. Vintage Books: New York (1953).

[10.](#) An insignificant number relative to the estimated 118 billion humans who have ever lived on the planet.

[11.](#) L. Brasington, *Right Concentration: A Practical Guide to the Jhanas*. Shambhala Publications: Boulder, CO (2015); J. Yates, M. Immergut, and J. Graves, *The Mind Illuminated: A Complete Meditation Guide Integrating Buddhist Wisdom and Brain Science for Greater Mindfulness*. Simon and Schuster: New York (2017). Some of these practices delimit the outer boundary of what people will endure. Christian hermits, or *stylites*, in the Byzantine era lived for years on top of a pillar; *anchorites*, during the Middle Ages, chose to spend the rest of their lives in a single room, usually adjacent to a church. Indeed, funeral rites were performed for them when they entered their abodes as they were lost to the world. The most famous anchoress was Julia of Norwich. *Sokushinbutsu*, meaning “those who obtain Buddhahood in the flesh,” were Buddhist ascetics who induced a process of self-mummification. It was practiced in medieval Japan until it was declared illegal in the late nineteenth century. Living on a strict diet that gradually eliminated all body fat and reduced body metabolism to a minimum, the ascetic would eventually starve to death. K. Jeremiah, *Living Buddhas: The Self-Mummified Monks of Yamagata, Japan*. McFarland: Jefferson, NC (2014).

[12.](#) At least one medical anthropologist explicitly describes the peyote ceremony in the Native American Church as a “consciousness modification” technique. J. Calabrese, *A Different Medicine: Postcolonial Healing in the Native American Church*. Oxford University Press: Oxford, UK (2013).

[13.](#) J. G. Dean et al., “Biosynthesis and extracellular concentrations of N, N-dimethyltryptamine (DMT) in the mammalian brain.” *Scientific Reports* 9: 1–11 (2019).

[14.](#) During my acute psilocybin experience, I noted a propensity to fixate for an unusually long time on objects or locations, together with a greatly elevated threshold for initiating any sort of movement, including

thoughts.

[15](#). B. Shanon, *The Antipodes of the Mind: Charting the Phenomenology of the Ayahuasca Experience*. Oxford University Press: Oxford, UK (2010).

[16](#). R. K. C. Forman, ed., *The Problem of Pure Consciousness: Mysticism and Philosophy*. Oxford University Press: Oxford, UK (1990).

[17](#). R. Doblin, “Pahnke’s Good Friday Experiment: A long-term follow-up and methodological critique.” *Journal of Transpersonal Psychology* 23(1): 1–25 (1991).

[18](#). B. C. Muraresku, *The Immortality Key: The Secret History of the Religion with No Name*. St. Martin’s Press: New York (2020).

[19](#). I encountered a bright light and a booming voice during one stormy night on a beach in Cape Cod, but it wasn’t quite what I thought it was. You can hear me speak about this in a *Moth Radio Hour* show titled “God, Death and Francis Crick.”

[20](#). I. Hartogsohn, “Set and setting in the Santo Daime.” *Frontiers in Pharmacology* 12: 610 (2021). For a detailed ethnography, see religious studies scholar William Barnard’s *Liquid Light: Ayahuasca Spirituality and the Santo Daime Tradition*. Columbia University Press: New York (2022).

[21](#). Schopenhauer espouses pure idealism, through and through, in which everything is ultimately a manifestation, or a representation, of the noumenal will. During my experience, I felt that I was tapping into a universal field of consciousness. I know, I know how it sounds. But there you have it.

[22](#). J. A. Brewer et al., “Meditation experience is associated with differences in default mode network activity and connectivity.” *Proceedings of the National Academy of Sciences of the United States of America* 108: 20254–20259 (2011). Operationally, for “closed” states Jud Brewer asks subjects to think of a time that they experienced being frustrated, anxious, fearful, and worried and to feel it in their bodies and for “open” states to recall times when they felt joyful, connected, and kind.

[23](#). T. Bayne and O. Carter, “Dimensions of consciousness and the psychedelic state.” *Neuroscience of Consciousness* 2018(1): niy008 (2018). It is challenging to systematically study creativity. One of the few studies that has done so in a credible manner with twenty-seven professionals given LSD is W. W. Harman et al., “Psychedelic agents in creative problem-

solving: A pilot study.” *Psychological Reports* 19: 211–227 (1966). Others microdose—that is, use subhallucinogenic doses of psilocybin or LSD—to deal with their anxieties. J. M. Rootman et al., “Adults who microdose psychedelics report health related motivations and lower levels of anxiety and depression compared to non-microdosers.” *Scientific Reports* 11: 1–11 (2021).

[24](#). Raymond Moody coined the term “near-death experience” in his 1975 book *Life After Life*. See also B. Greyson, “Consistency of near-death experience accounts over two decades: Are reports embellished over time?” *Resuscitation* 73(3): 407–411 (2007); S. Blackmore, “Near-death experiences: In or out of the body?” *Skeptical Inquirer* 16: 34–45 (1991); J. M. Holden, B. Greyson, and D. James, eds., *The Handbook of Near-Death Experiences: Thirty Years of Investigation*. Praeger: Santa Barbara, CA (2009).

[25](#). Free diving or high-altitude climbing without supplemental oxygen can lead to pleasurable light-headedness, jauntiness, and heightened arousal due to reduced oxygen and/or enhanced carbon dioxide. And then there are the practices that dare not speak their name, the choking, fainting, or blackout game in teenagers and erotic asphyxiation in older teens and tweens. If practiced alone, they can be lethal, leaving parents and siblings in despair and engulfed in guilt. For general references on the neuroscience of near-death experience, see O. Blanke, N. Faivre, and S. Gieguez, “Leaving body and life behind: Out-of-body and near-death experience.” In S. Laureys, O. Gossiers, and G. Tononi, eds., *The Neurology of Consciousness*, pp. 323–347. 2nd ed. Elsevier: Amsterdam (2015); C. Martial et al., “Near-death experience as a probe to explore (disconnected) consciousness.” *Trends in Cognitive Sciences* 24: 173–183 (2020).

[26](#). About 10 percent of cardiac arrest patients survive. To put this into perspective, most everyone whose heart stopped died before the advent of 911, helicopter evacuation, and other high-tech interventions.

[27](#). S. Parnia et al., “A qualitative and quantitative study of the incidence, features and aetiology of near-death experiences in cardiac arrest survivors.” *Resuscitation* 48: 149–156 (2001); B. Greyson, “Incidence and correlates of near-death experiences in a cardiac care unit.” *General Hospital Psychiatry* 4: 269–276 (2003); P. Van Lommel et al., “Near-death experience in survivors of cardiac arrest: A prospective study in the

Netherlands.” *Lancet* 358: 2039–2045 (2001); S. Parnia et al., “AWARE—AWAREness during REsuscitation—A prospective study.” *Resuscitation* 85: 1799–1805 (2014).

[28](#). Perhaps the near-death experience is generated during the booting-up of the brain, and it is then, incorrectly, time-stamped. A further complication is that an isoelectric EEG does not rule out electric activity in structures underneath the neocortex, such as the hippocampus. D. Kroeger, B. Florea, and F. Amzica, “Human brain activity patterns beyond the isoelectric line of extreme deep coma.” *PLoS ONE* 8: e75257 (2013).

[29](#). C. Martial et al., “Neurochemical models of near-death experiences: A large-scale study based on the semantic similarity of written reports.” *Consciousness & Cognition* 69: 52–69 (2019).

[30](#). Subsequently, I discovered that it is not uncommon for individuals who have undergone a near-death experience to lose their fear of dying. N. A. Tassell-Matamua and N. Lindsay, “‘I’m not afraid to die’: The loss of the fear of death after a near-death experience.” *Mortality* 21: 71–87 (2016).

[31](#). This may be accompanied by low-frequency activity in the delta regime of the EEG, synchronized throughout the posterior hot zone. Indeed, DMT, the psychoactive component in ayahuasca, induced such an increase in delta (1–4 Hz) EEG power and a decrease in alpha power (8–12 Hz) relative to simply closing the eyes in thirty-five healthy and experienced ayahuasca users. C. Pallavicini et al., “Neural and subjective effects of inhaled N, N-dimethyltryptamine in natural settings.” *Journal of Psychopharmacology* 35: 406–420 (2021).

[32](#). M. Boly et al., “Neural correlates of pure presence.” 23rd meeting of the Association for the Scientific Study of Consciousness, New York (2023).

[33](#). A. Haun and G. Tononi, “Why does space feel the way it does? Towards a principled account of spatial experience.” *Entropy* 21: 1160 (2019).

## Chapter 8: Transforming Lives by Transformative Experiences

1. I found this jewel in J. Brewer, *Unwinding Anxiety*. Penguin Random House: New York (2021), a book dealing with anxiety that I can warmly recommend.

2. T. De Quincey, *Confessions of an English Opium-Eater*. Joystones Publishing (1821); M. Jay, *Psychonauts: Drugs and the Making of the Modern Mind*. Yale University Press: New Haven, CT (2023).

3. The dark side of “the summer of love” in San Francisco in 1967 is well captured by Joan Didion’s essay “Slouching Towards Bethlehem.” Its memorable opening line is “The center was not holding.”

4. Schedule I drugs include all opioids and opium derivatives, psychedelics, depressants (except alcohol), and stimulants. Declaring these drugs illegal and possession punishable by prison time was followed by the disastrous “War on Drugs,” the campaign by the US government to suppress the illegal drug trade. While it failed to curtail the availability of cheap, powerful, and highly addictive drugs, now primarily synthetic opioids, it led to a massive and sustained increase in the prison population and created a large class of marginalized people, with African Americans highly overrepresented relative to their drug use.

5. A pioneer during this dark age was Franz Vollenweider at the University of Zürich. See the comprehensive A. L. Halberstadt, F. X. Vollenweider, and D. E. Nichols, eds., *Behavioral Neurobiology of Psychedelic Drugs*. Springer: Berlin (2018). A better-controlled version of the 1962 Marsh Chapel or Good Friday Experiment, undertaken at Johns Hopkins University in 2002 by Roland Griffiths and colleagues, came to a more measured but similar conclusion as the original study. R. R. Griffiths et al., “Psilocybin can occasion mystical-type experiences having substantial and sustained personal meaning and spiritual significance.” *Psychopharmacology* 187: 268–283 (2006).

6. A survey of the safety of common drugs, such as alcohol, heroin, cocaine, cannabis, and tobacco, including both harm to users and harm to others (e.g., when a drunken driver causes a deadly accident), for the United Kingdom was led by David Nutt. It concluded that LSD and mushrooms had the lowest harm profile of any of the drugs studied. D. J. Nutt, L. A.

King, and L. D. Phillips, “Drug harms in the UK: A multicriteria decision analysis.” *The Lancet* 376: 1558–1565 (2010).

7. Pollan’s deeply researched book, *How to Change Your Mind: What the New Science of Psychedelics Teaches Us About Consciousness, Dying, Addiction, Depression, and Transcendence*. Penguin: New York (2019), discusses, in an engaging manner, his own encounters with these molecules. I can also warmly recommend his newer *This Is Your Mind on Plants*. Penguin: New York (2021), a triptych of essays on opium, caffeine—the world’s most widely consumed psychoactive substance—and mescaline.

8. A. K. Davis et al., “Effects of psilocybin-assisted therapy on major depressive disorder: A randomized clinical trial.” *Journal of the American Medical Association, Psychiatry* 78: 481–489 (2021). The latest study reported large reductions in depressive symptoms forty-three days after a single dose of psilocybin. C. L. Raison et al., “Single-dose psilocybin treatment for major depressive disorder: A randomized clinical trial.” *Journal of the American Medical Association*: <https://doi.org/10.1001/jama.2023.14530> (2023).

9. C. S. Grob et al., “Pilot study of psilocybin treatment for anxiety in patients with advanced-stage cancer.” *Archives of General Psychiatry* 68: 71–78 (2011); R. R. Griffith et al., “Psilocybin produces substantial and sustained decreases in depression and anxiety in patients with life-threatening cancer: A randomized double-blind trial.” *Journal of Psychopharmacology* 30: 1181–1197 (2016).

10. S. Nolen-Hoeksema, B. E. Wisco, and S. Lyubomirsky, “Rethinking rumination.” *Perspectives on Psychological Science* 3: 400–424 (2008); T. Barba et al., “Effects of psilocybin versus escitalopram on rumination and thought suppression in depression.” *British Journal of Psychology Open* 8: e163 (2022).

11. D. B. Yaden, J. B. Potash, and R. R. Griffiths, “Preparing for the bursting of the psychedelic hype bubble.” *Journal of the American Medical Association Psychiatry* 79: 943–944 (2022).

12. This rush to reap financial benefits is leading to an arms race between those promoting open science and protecting the public from high-cost therapies (see Porta Sophia: <https://www.portasophia.org>) and for-profit companies seeking to maximize their intellectual property claims and returns on investment. M. Marks and I. G. Cohen, “Patents on psychedelics:



The next legal battleground of drug development.” *Harvard Law Review Patents* 135: 212 (2021).

[13](#). These numbers include “recruiting” and “not yet recruiting” clinical trials at <https://clinicaltrials.gov>. See K. A. A. Andersen et al., “Therapeutic effects of classic serotonergic psychedelics: A systematic review of modern-era clinical studies.” *Acta Psychiatrica Scandinavica* 143: 101–118 (2021).

[14](#). D. Nutt, *Drugs Without the Hot Air: Minimizing the Harms of Legal and Illegal Drugs*. UIT Cambridge: Cambridge, UK (2012).

[15](#). T. S. Krebs and P. Ø. Johansen, “Psychedelics and mental health: A population study.” *PLoS ONE* 8: e63972 (2013).

[16](#). Contrary to rumor, there is only a small number of confirmed deaths after LSD ingestion—two following a massive overdose and two following extreme physical restraint by police that triggered fatal cardiovascular collapse. D. E. Nichols and C. S. Grob, “Is LSD toxic?” *Forensic Science International* 284: 141–145 (2018). There are a handful of reported deaths after psilocybin use, associated with drowning or motor vehicle crashes. J. B. Leonard, B. Anderson, and W. Klein-Schwartz, “Does getting high hurt? Characterization of cases of LSD and psilocybin-containing mushroom exposures to national poison centers between 2000–2016.” *Journal of Psychopharmacology* 32: 1286–1294 (2018). Psychedelics do trigger a moderate (10–20 mm Hg) and transient increase in blood pressure.

[17](#). Purging takes the form of vomiting and diarrhea. Note that 95 percent of all serotonin molecules in the human body are in the gut.

[18](#). G. Martinotti et al., “Hallucinogen persisting perception disorder: Etiology, clinical features, and therapeutic perspectives.” *Brain Sciences* 8: 47 (2018); B. Murrie et al., “Transition of substance-induced, brief, and atypical psychoses to schizophrenia: A systematic review and meta-analysis.” *Schizophrenia Bulletin* 46: 505–516 (2020). In general, the long-term risk is larger for people with psychiatric conditions than for healthy ones and is larger in an uncontrolled, naturalistic setting (e.g., taking it with a friend at a club or at home) than in a controlled, therapeutic setting. Several surveys capture the risk of psychedelic mushrooms in regular users in naturalistic settings: T. M. Carbonaro et al., “Survey study of challenging experiences after ingesting psilocybin mushrooms: Acute and enduring positive and negative consequences.” *Journal of Psychopharmacology*



30(12): 1268–1278 (2016); S. M. Nayak et al., “Naturalistic psilocybin use is associated with persisting improvements in mental health and wellbeing: Results from a prospective, longitudinal survey.” *Frontiers in Psychiatry* 14: 1199642 (2023); J. Evans et al., “Extended difficulties following the use of psychedelic drugs: A mixed methods study.” *PLoS ONE* 18(10): e0293349 (2023).

[19](#). M. Janikian, *Your Psilocybin Mushroom Companion*. Ulysses Press: Berkeley, CA (2019).

[20](#). One of the few widely accepted observations about the effects of psychedelics on the brain is that binding to 5-HT-2A receptors is necessary to evoke the acute psychedelic effects. F. X. Vollenweider and K. H. Preller, “Psychedelic drugs: Neurobiology and potential for treatment of psychiatric disorders.” *Nature Reviews Neuroscience* 21: 611–624 (2020); A. C. Kwan et al., “The neural basis of psychedelic action.” *Nature Neuroscience* 25: 1407–1419 (2022); G. Ballentine, S. F. Freesun, and D. Bzdok, “Trips and neurotransmitters: Discovering principled patterns across 6,850 hallucinogenic experiences.” *Science Advances* 8(11): eabl6989 (2022). Psychedelics can have downstream effects on dopaminergic and adrenergic receptors. Nothing is ever simple in biology. For details, see the magisterial D. E. Nichols, “Psychedelics.” *Pharmacological Reviews* 68: 264–355 (2016). In the latest twist, psychedelics have been shown to bind to serotonergic 2A receptors inside neurons rather at the external membrane enveloping each neuron. M. V. Vargas et al., “Psychedelics promote neuroplasticity through the activation of intracellular 5-HT2A receptors.” *Science* 379: 700 (2023).

[21](#). There are between three and five thousand distinct types of transcriptionally distinct cells in the mammalian brain, each one expressing a subset of none, one, or more of the twenty-one known serotonergic receptor subtypes. This makes efforts to mechanistically interpret what happens at the neuronal level when humans ingest psychedelic substances enormously challenging. For brain imaging studies of tripping volunteers, see R. L. Carhart-Harris et al., “Neural correlates of the psychedelic state as determined by fMRI studies with psilocybin.” *Proceedings of the National Academy of Sciences of the United States of America* 109: 2138–2143 (2012); L. Smigielski et al., “Psilocybin-assisted mindfulness training modulates self-consciousness and brain default mode network connectivity

with lasting effects.” *NeuroImage* 196: 207–215 (2019); J. J. Gattuso, “Default mode network modulation by psychedelics: A systematic review.” *International Journal of Neuropsychopharmacology* 26(3): 155–188 (2022). For MEG and EEG studies, see M. M. Schartner et al., “Increased spontaneous MEG signal diversity for psychoactive doses of ketamine, LSD and psilocybin.” *Scientific Reports* 7: 1–12 (2017); A. Ort et al., “TMS-EEG and resting-state EEG applied to altered states of consciousness: Oscillations, complexity, and phenomenology.” *Isience* 26: 106589 (2023). One experiment that should be done is to ask tripping volunteers to signal when they have strong “visions” while lying in the magnet or sitting in a chair with their head enveloped by EEG or MEG measurement devices and to compare their visual cortical activity to times during the trip without any active visual manifestations of the psychedelic (always with closed eyes).

[22](#). The existing behavioral read-out for the potency of psychedelics in rodents that everybody in the field relies on is the *head-twitch response*, a very rapid head rotation (a bit like a wet dog shake) that has no human counterpart. A. L. Halberstadt et al., “Correlation between the potency of hallucinogens in the mouse head-twitch response assay and their behavioral and subjective effects in other species.” *Neuropharmacology* 167: 107933 (2020).

[23](#). One monkey, upon self-administering DMT, exhibited “several signs of intoxication, including stereotyped visual scanning around the room, head twitches, bizarre body postures, hyperactivity, and ‘fly catching’ (fixating on an empty point in space and attempting to quickly grasp it).” W. E. Fantegrossi, J. H. Woods, and G. Winger, “Transient reinforcing effects of phenylisopropylamine and indolealkylamine hallucinogens in rhesus monkeys.” *Behavioural Pharmacology* 15: 149–157 (2004).

[24](#). The best evidence for enhanced neuroplasticity comes from rodents given psychedelics. These trigger growth in dendrites and synaptic spines, specialized junctions through which most of the synaptic traffic among neurons occurs, in the prefrontal cortex. Some evidence suggests that an important nerve growth factor, brain-derived neurotrophic factor (BDNF), is elevated in the peripheral blood following psychedelic consumption by users. A. E. Calder and G. Hasler, “Towards an understanding of psychedelic-induced neuroplasticity.” *Neuropsychopharmacology* 48(1):

104–112 (2022); R. Nardou et al., “Psychedelics reopen the social reward learning critical period.” *Nature* 618: 790–798 (2023).

25. The clashing views in the debate are expressed by D. E. Olson, “The subjective effects of psychedelics may not be necessary for their enduring therapeutic effects.” *ACS Pharmacology & Translational Science* 4: 563–567 (2020); D. B. Yaden and R. G. Roland, “The subjective effects of psychedelics are necessary for their enduring therapeutic effects.” *ACS Pharmacology & Translational Science* 4: 568–572 (2020). The latest shot, in mice, in this debate is R. Moliner et al., “Psychedelics promote plasticity by directly binding to BDNF receptor TrkB.” *Nature Neuroscience* 26: 1032–1041 (2023).

26. Numerous studies have found a link between mystical experience and therapeutic benefit: R. R. Griffiths et al., “Psilocybin produces substantial and sustained decreases in depression and anxiety in patients with life-threatening cancer: A randomized double-blind trial.” *Journal of Psychopharmacology* 30: 1181–1197 (2016); A. Garcia-Romeu et al., “Psilocybin-occasioned mystical experiences in the treatment of tobacco addiction.” *Current Drug Abuse Reviews* 7: 157–164 (2014); C. L. Raison et al., “Single-dose psilocybin treatment for major depressive disorder: A randomized clinical trial.” *Journal of the American Medical Association*: <https://doi.org/10.1001/jama.2023.14530> (2023).

27. Of course, the extraordinary experience will have a physical substrate, say the binding of the psychedelic to serotonergic 5-HT-2A receptors, leading to a cascade of biophysical and biochemical effects down the line. True agency, per integrated information theory, lies with intrinsic causal powers, not the extrinsic ones, that is, with consciousness, not with its underlying substrate.

28. None of these have yet been addressed by the current wave of psychedelic trials in clinical populations. For a critical look at the existing trials and their inadequate designs, see M. van Elk and E. Fried, “History repeating: A roadmap to address common problems in psychedelic science.” PsyArXiv: <https://doi.org/10.31234/osf.io/ak6gx> (March 10, 2023).

29. Chronic use can cause potentially life-threatening cardiac valvulopathy via 5-HT-2B receptors. T. D. McClure-Begley and B. L. Roth, “The promises and perils of psychedelic pharmacology for psychiatry.”

*Nature Reviews Drug Discovery* 21: 463–473 (2022).

[30](#). S. D. Muthukumaraswamy, A. Forsyth, and T. Lumley, “Blinding and expectancy confounds in psychedelic randomized controlled trials.” *Expert Review of Clinical Pharmacology* 14: 1133–1152 (2021).

[31](#). Named in homage to Carl Sagan’s famous “Pale Blue Dot” image of planet Earth from six billion kilometers away, the foundation can be found at <https://www.tinybluedotfoundation.org>. I am its chief scientist. And no, Elizabeth R. Koch and I are not related.

[32](#). Imagine if politicians, leaders, and (ex-)presidents, whose public selves are deified and demonized, were to experience oceanic boundlessness. How could this not be transformative for their attitude to life and the well-being of the citizens they are responsible for, thereby reducing tension, conflict, and strife on the planet? Of course, those who are the most ego infatuated and besotted with testosterone are also the most afraid of losing control and will vehemently avoid ego loss.

## Chapter 9: The End of Consciousness

1. This title is drawn from a story concerning the inevitability of death whose pedigree reaches all the way back to ancient Mesopotamia. Here is a pithy retelling from Somerset Maugham (1933):

A merchant in Baghdad sends his servant to the marketplace for provisions. Soon afterwards, the servant comes home white and trembling and tells him that in the marketplace, he was jostled by a woman, whom he recognized as Death, who made a threatening gesture. Borrowing the merchant's horse, the servant flees at great speed to Samarra where he believes Death will not find him. The merchant then goes to the marketplace and finds Death and asks why she made the threatening gesture to his servant. She replies, "That was not a threatening gesture, it was only a start of surprise. I was astonished to see him in Baghdad, for I have an appointment with him tonight, in Samarra."

2. Ernest Becker argued in his 1974 Pulitzer Prize-winning work *The Denial of Death* that religion, literature, music, and other arts, and much else that drives people, are simply ways of dealing with the anxiety and terror caused by the finitude of existence.

3. There are numerous examples of collective forgetting, such as the violence associated with the "settlement" of the American West, the colonial empires of European powers in Africa and Asia, the Holocaust in Germany and Nazi-occupied Europe, the killings and starvations of the Stalinist Soviet Union, and on and on. Another example of a failure to remember is the great influenza epidemic, also known as the "Spanish flu," which has left little trace in our collective imagination. Although between 1 and 5 percent of the world population fell victim to the flu virus, far deadlier than the Covid-19 pandemic, few novels, movies, book, statues, or other works of art commemorate this epidemic.

4. G. Egan, *Permutation City*. Night Shade Books: New York (1994); N. Stephenson, *Fall; or, Dodge in Hell*. William Morrow: New York (2019).

5. A. Sullivan, "I used to be a human being." *New York Magazine*

(September 2016).

[6.](#) A. Lewis et al., “Determination of death by neurologic criteria around the world.” *Neurology* 95: e299–e309 (2020).

[7.](#) E. F. Wijdicks et al., “Evidence-based guideline update: Determining brain death in adults: Report of the Quality Standards Subcommittee of the American Academy of Neurology.” *Neurology* 74: 1911–1918 (2010); J. A. Russell et al., “Brain death, the determination of brain death, and member guidance for brain death accommodation requests: AAN position statement.” *Neurology* 92: 228–232 (2019).

[8.](#) R. D. Truog, F. G. Miller, and S. D. Halpern, “The dead-donor rule and the future of organ donation.” *New England Journal of Medicine* 369:1287–1289 (2013); J. L. Verheijde, M. Y. Rady, and M. Potts, “Neuroscience and brain death controversies: The elephant in the room.” *Journal of Religion and Health* 57: 1745–1763 (2018).

[9.](#) Telling people not to worry, that their conscious mind, their “soul,” will return after death to the same condition of nonexistence it was in before birth, does not seem to make much of a dent in existential dread and excessive ego attachment!

[10.](#) About 2 percent of all hospital deaths in the United States are brain deaths. A. Seifi, J. V. Lacci, and D. A. Godoy, “Incidence of brain death in the United States.” *Clinical Neurology and Neurosurgery* 195: 105885 (2020).

[11.](#) R. Aviv, “What does it mean to die?” *New Yorker* (February 5, 2018).

[12.](#) “A code of practice for the diagnosis and confirmation of death.” Academy of Medical Royal Colleges, London (2008).

[13.](#) Z. Vrselja et al., “Restoration of brain circulation and cellular functions hours post-mortem.” *Nature* 568: 336–343 (2019). See my commentary in “Is death reversible?” *Scientific American* 321: 34–37 (October 2019). Based on current knowledge, say from near drowning of children, the chances of functional recovery after such restoration are slim, raising the question whether it is worthwhile in terms of quality of life.

[14.](#) L. Butcher, “When the line between life and death is a little bit fuzzy.” *Undark* (May 10, 2021). See also A. D. Marcus, “Doctors and lawyers debate meaning of death as families challenge practices.” *Wall Street Journal* (December 11, 2022); R. D. Truog, “The uncertain future of



the determination of brain death.” *Journal of the American Medical Association* 329(12): <https://doi.org/10.1001/jama.2023.1472> (2023).

[15.](#) L. S. Chawla et al., “Characterization of end-of-life electroencephalographic surges in critically ill patients.” *Death Studies* 41: 385–392 (2017); D. Kondziella, “The neurology of death and the dying brain: A pictorial essay.” *Frontiers in Neurology* 11: 736 (2020); N. A. Shlobin et al., “What happens in the brain when we die? Deciphering the neurophysiology of the final moments in life.” *Frontiers in Aging Neuroscience* 15: 281 (2023).

[16.](#) F. C. Crick and C. Koch, “Towards a neurobiological theory of consciousness.” *Seminars in the Neurosciences* 2: 263–275 (1990); C. Koch et al., “The neural correlates of consciousness: Progress and problems.” *Nature Reviews Neuroscience* 17: 307–321 (2016).

[17.](#) The gamma band activity in the 25–150 Hz range was coupled with the slower alpha and beta waves as discussed in G. Xu et al., “Surge of neurophysiological coupling and connectivity of gamma oscillations in the dying human brain.” *Proceedings of the National Academy of Sciences of the United States of America* 120: e2216268120 (2023). My most serious concern with these findings is that the massive increase in gamma power (up to a factor of 392) is compatible with epileptic seizures. Triggered by anoxia and ischemia, ionic gradients of individual neurons are running down, and massive neurotransmitter release occurs, all under extreme pathological conditions that perturb the fine balance between excitation and inhibition. This is a breeding ground for localized, hypersynchronized neuronal assemblies. See also the case of an eighty-seven-year-old man undergoing cardiac arrest after traumatic subdural hematoma. R. Vicente et al., “Enhanced interplay of neuronal coherence and coupling in the dying human brain.” *Frontiers in Aging Neuroscience* 14: <https://doi.org/10.3389/fnagi.2022.813531> (2022). Both clinical reports corroborate an earlier study in dying rats that likewise found a paradoxical transient increase in synchronized high-frequency activity following cardiac arrest. J. Borjigin et al., “Surge of neurophysiological coherence and connectivity in the dying brain.” *Proceedings of the National Academy of Sciences of the United States of America* 110: 14432–14437 (2013).

[18.](#) M. Nahm et al., “Terminal lucidity: A review and a case collection.” *Archives of Gerontology and Geriatrics* 55:138–142 (2012); S. Macleod,



*The Psychiatry of Palliative Medicine: The Dying Mind.* Radcliffe Publishing: Abingdon, UK (2011); A. A. Chiriboga-Oleszczak, “Terminal lucidity.” *Current Problems of Psychiatry* 18: 34–46 (2017).

## Chapter 10: The Future of Consciousness

1. N. A. Farahany, *The Battle for Your Brain: Defending the Right to Think Freely in the Age of Neurotechnology*. St. Martin's Press: New York (2023).

2. M. Leber et al., "Advances in penetrating multichannel microelectrodes based on the Utah array platform." *Neural Interface: Frontiers and Applications* 1101: 1–40 (2019); E. Musk, "An integrated brain-machine interface platform with thousands of channels." *Journal of Medical Internet Research* 21: e16194 (2019).

3. B. Hubert, "Reverse engineering the source code of the BioNTech/Pfizer SARS-CoV-2 Vaccine." [berthub: https://berthub.eu/articles/posts/reverse-engineering-source-code-of-the-biontech-pfizer-vaccine](https://berthub.eu/articles/posts/reverse-engineering-source-code-of-the-biontech-pfizer-vaccine) (December 25, 2020).

4. The DSM-5, the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders*, is the bible of psychiatry, providing a common language for an exhaustive list of mental disorders. However, it does not contain a single quantitative imaging, blood, or genetic test that can validate any of these conditions in a reliable manner. This, more than anything else, demonstrates our limited insights into psychopathologies that leave no apparent bodily traces. While some biomarkers can distinguish between a group of depressed patients and a matched group of nondepressed controls, they fail to have high enough sensitivity and specificity to be helpful in diagnosing a specific individual. This requires an actual conversation with that person. T. Insel, *Healing: Our Path from Mental Illness to Mental Health*. Penguin Press: New York (2022).

5. The worm *Caenorhabditis elegans*, a nematode lovingly referred to as *C. elegans*, is roughly one millimeter long and lives in the soil. Its body consists of about 1,000 somatic cells, of which 302 make up its nervous system. The original worm connectome is described in J. G. White et al., "The structure of the nervous system of the nematode *Caenorhabditis elegans*." *Philosophical Transactions of the Royal Society B: Biological Sciences* 314: 1–340 (1986).

6. According to M. Winding et al., "The connectome of an insect brain." *Science* 379: eadd9330 (2023), a single fly maggot includes three

thousand neurons and more than a half million synapses. Next up is the brain of the adult fruit fly with about a hundred thousand neurons. S. Dorkenwald et al., “FlyWire: Online community for whole brain connectomics.” *Nature Methods* 19: 119–128 (2022). For the mouse connectome, see L. Abbott et al., “The Mind of a Mouse.” *Cell* 182: 1372–1376 (2020); N. L. Turner et al., “Reconstruction of neocortex: Organelles, compartments, cells, circuits, and activity.” *Cell* 185: 1082–1100 (2022).

7. Y. N. Billeh et al., “Systematic integration of structural and functional data into multi-scale models of mouse primary visual cortex.” *Neuron* 106: 388–403 (2020). The model recapitulates the electrical events in the visual cortex (e.g., synaptic and action potentials) when mice run past photos and movies.

8. The number of network parameters grows with the square of the network size. That is, a simulation one thousand times larger will have a million times more parameters. The actual connectome might be helpful to constrain the parameter search, if the synaptic weights and other parameters could be inferred from the high-resolution electron microscopic images of the tissue—say, the synaptic weight by the thickness of the postsynaptic density.

9. J. H. Morrison and P. R. Hof, “Life and death of neurons in the aging brain.” *Science* 278: 412–419 (1997); S. M. Marks et al., “Tau and  $\beta$ -amyloid are associated with medial temporal lobe structure, function, and memory encoding in normal aging.” *Journal of Neuroscience* 37: 3192–3201 (2017).

10. *Emergent* implies that consciousness is a system-level property not apparent at the level of the microscopic constituents of the system. Philosophers distinguish *weak emergence*, in which macroscopic properties follow in some lawful and systematic manner from the underlying micro-physical variables, from *strong emergence* in which entirely new system-level properties appear. Wetness is an example of weak emergence, as two molecules of  $H_2O$  are not wet, but a liter of water is. Yet functionalists are at a loss to explain how phenomenal consciousness emerges out of a bunch of interconnected neurons. It is more in the nature of a hope than a specific research program.

11. S. Dehaene, H. Lau, and S. Kouider, “What is consciousness, and

could machines have it?” *Science* 358: 486–492 (2017).

[12](#). This validates the intuition that conscious systems require massive intrinsic feedback connections. M. Oizumi, L. Albantakis, and G. Tononi, “From the phenomenology to the mechanisms of consciousness: Integrated information theory 3.0.” *PLoS Computational Biology* 10: e1003588 (2014). I refer the interested reader to my last book, *The Feeling of Life Itself*, which covers this ground in more detail. The article by the philosopher Francis Fallon highlights that IIT, unique among theories of consciousness, has a precise answer to the fundamental mereological challenge of what constitutes a whole: F. Fallon, “Integrated information theory, Searle, and the arbitrariness question.” *Review of Philosophy and Psychology* 11: 629–645 (2020). See also M. Tegmark, “Consciousness as a state of matter.” *Chaos, Solitons & Fractals* 76: 238–270 (2015).

[13](#). The computer carrying out the relativistic simulations has some mass that will influence spatial curvature by an itty-bitsy amount. It has a whit of extrinsic causal power.

[14](#). Physicist and circuit designer Carver Mead originated many of the ideas and much of the early practice of designing neuromorphic chips. C. Mead, “Neuromorphic electronic systems.” *Proceedings of the IEEE* 78: 1629–1636 (1990). For a review, see A. R. Young et al., “A review of spiking neuromorphic hardware communication systems.” *IEEE Access* 7: 135606–135620 (2019).

## Chapter 11: What Computers Can Never Be

1. GPT stands for generative pretrained transformer, describing the powerful algorithm, related to neural network techniques, that these large language models use. Transformers were first described in A. Vaswani et al., “Attention is all you need.” *Advances in Neural Information Processing Systems* 30 (2017). This preprint has already been cited close to hundred thousand times in the intervening seven years, a testament to the rapidity with which transformer algorithms have been adopted by the community. The average successful scientist at a major research university will perhaps be cited ten thousand times over their professional lifetime. All tech companies and a bevy of start-ups are bringing their own models to market. The pace is torrential, fueled by open-access publications and near instantaneous feedback from millions of users.

2. One of the most popular is the Common Crawl, generated monthly by a not-for-profit that crawls the entire publicly accessible internet, a very large artifact, with its large bias toward the English-speaking world. The latest crawl from spring 2023 captured more than three billion webpages. Large language models are trained on versions of the crawl with offensive, violent, and pornographic material removed, such as Google’s Colossal Clean Common Crawl, or C4. This can be quite traumatizing to the offshored workers tasked with this scrubbing, given the cesspool that constitutes humanity’s darker desires.

3. The missing word is *access*.

4. N. N. Taleb, *The Black Swan: The Impact of the Highly Improbable*. Random House: New York (2007).

5. E. Roivainen, “I gave ChatGPT an IQ test. Here’s what I discovered.” *Scientific American* (March 28, 2023). The author, a clinical psychologist, used the standard Wechsler Adult Intelligence Scale and excluded nonverbal tests as the chatbot has neither eyes nor hands.

6. S. Bubeck et al., “Sparks of artificial general intelligence: Early experiments with GPT-4.” arXiv:2303.12712 (2023).

7. This is known as the imitation game, or the Turing test, proposed in this foundational paper: A. M. Turing, “Computing machinery and intelligence.” *Mind* 49: 433–460 (1950). It is meant as an operational test to

assess whether a computer can successfully imitate a human engaged in a conversation about any topic.

8. For the record, when I asked ChatGPT whether it was conscious, it replied, “As an artificial intelligence language model, I do not have subjective experience or consciousness in the sense of Thomas Nagel’s definition of subjective experience, also known as ‘what it’s like’ to experience something. I can perform tasks such as answering questions and generating responses based on my training data and programming, but I do not have a subjective experience of those tasks. My responses are based solely on algorithms and statistical patterns, without the subjective experience that humans possess.”

I kept on pressing, asking it to imagine a situation in which some other chatbot could be conscious, but it was clever enough to detect these subterfuges.

9. Finally, fed the same software, computers behave identically to each other. This is by design as you don’t want your software to operate differently from one laptop to the next. Contrariwise, no two biological organisms are ever the same—even when they are identical twins or cloned animals. This is a feature and not a bug as variability is what drives evolution by selection.

10. Because of this powerful biological urge to attribute theory of mind to other, there will be a large market for personalized language models mimicking specific individuals, MyGPT. Celebrities will perpetuate their brand in this way: MyGPT for Taylor Swift or Justin Bieber, doubles in all ways that matter in our social-media-saturated world in which physical encounters, which would give the game away, are becoming less frequent. MyGPT would gather as much information about me as possible, including emails, letters, social media, images, videos, papers, talks, and so on. If shown a photo it could pinpoint its location, its time, and the identity of the family or friends portrayed; it would be able to “recall” details, whether imagined or not, in my accent-flavored voice. Although disembodied, MyGPT would be Christof to any observer communicating via email, phone, or Zoom. You could ask it questions that I have never thought about, and MyGPT would generate plausible answers, based on my personal history and beliefs, in my idiosyncratic style. It would spawn the illusion of having captured the essence of me, carrying the echo of who I am down the

river of time.

[11](#). This stark distinction is well expressed in this thoughtful essay on what can be learned about machine consciousness from studying human consciousness: P. Butlin et al., “Consciousness in artificial intelligence: Insights from the science of consciousness.” arXiv:2308.08708 (2023).

[12](#). On the topic of free will, I can recommend the slim volume by Berkeley philosopher John Searle (he of the “Chinese Room”), *Mind: A Brief Introduction*. Oxford University Press: Oxford, UK (2004). What is rare, in my experience of philosophers, is Searle’s admitting to being puzzled by the discrepancy between his intellectual conviction that free will cannot exist in a determined universe and his conscious experience of freedom of choice. To quote Searle: “If... I am in a restaurant and I am confronted with a menu and the waiter asks me what I would like, I cannot say ‘I’m a determinist, I’ll just wait and see what happens,’ because even that utterance is only intelligible to me as an exercise of my free will.”

[13](#). The stronger my belief in vegetarianism, the “easier” the choice of what to eat. The debate then shifts to considering the determinants of this freely willed ethical choice.

[14](#). These are deep waters. For the technical details, see G. Tononi et al., “Only what exists can cause: An intrinsic powers view of free will.” arXiv:2206.02069 (2023). This view leaves no room for emergence or dualism of any sort. There is also no *interaction problem* as there are no dual domains of the physical and the mental as in classical Cartesian dualism. There is only causal power. Unfolded intrinsic power is identical to conscious experience. Operationally, it manifests as a causal agent of change.

[15](#). B. Libet et al., “Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): The unconscious initiation of a freely voluntary act.” *Brain* 106: 623–642 (1983). This experiment has been repeated in different ways, including using a magnetic scanner rather than an EEG. C. S. Soon et al., “Unconscious determinants of free decisions in the human brain.” *Nature Neuroscience* 11: 543–545 (2008); U. Maoz et al., “Pre-deliberation activity in prefrontal cortex and striatum and the prediction of subsequent value judgment.” *Frontiers in Neuroscience* 7: 1–16 (2013). For a review of the neuropsychology of volition, see P. Haggard, “The neurocognitive bases of human volition.” *Annual Review of*



*Psychology* 70: 9–28 (2019).

[16](#). Unlike for, say, the speed of light, there are no known upper bounds to intelligence, however defined. A useful way to think about superintelligence is to imagine a computer that has the thinking skills of an Albert Einstein or a Madame Curie but a thousand times faster. That would effectively constitute an intelligence that *Homo sapiens* could not hope to match. N. Bostrom, *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press: Oxford (2014).