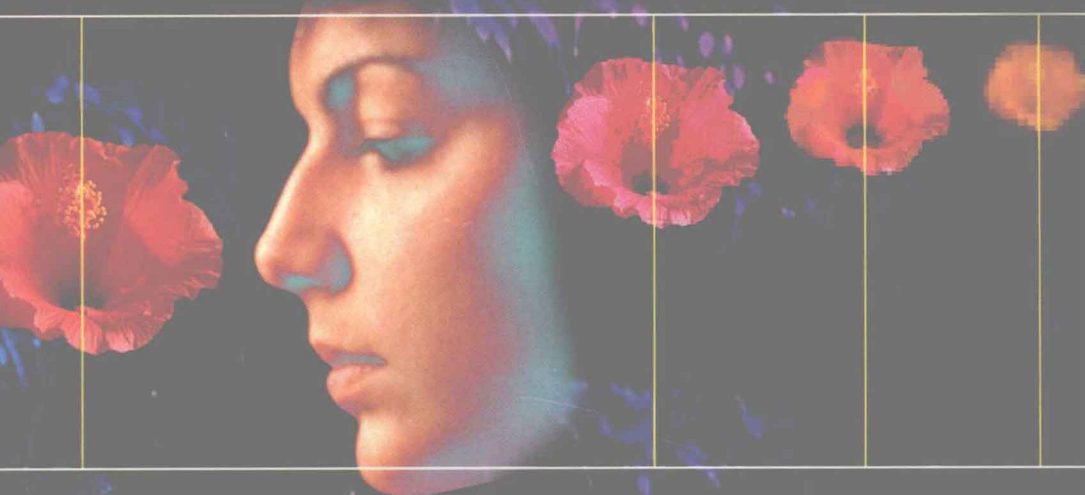
THE CREATIVE LOOP

How the Brain Makes a Mind



ERICH HARTH

Author of *Windows on the Mind*

The Creative Loop

How the Brain Makes a Mind

Erich Harth

A William Patrick Book

ADDISON-WESLEY PUBLISHING COMPANY

Reading, Massachusetts · Menlo Park, California · New York
Don Mills, Ontario · Wokingham, England · Amsterdam · Bonn
Sidney · Singapore · Tokyo · Madrid · San Juan
Paris · Seoul · Milan · Mexico City · Taipei

Excerpts from "The Universe" and "The Watch" by May Swenson reprinted by permission of The Literary Estate of May Swenson.

Excerpt from the "The Idea of Order at Key West" from *Collected Poems* by Wallace Stevens. Copyright 1936 by Wallace Stevens and renewed 1964 by Holly Stevens. Reprinted by permission of Alfred A. Knopf, Inc.

Excerpt from "Flux" by Theodore Melnechuk reprinted by permission of the author.

Illustration credits: p. xxiv—Photograph of Hindsgavl dagger reprinted by permission of Nationalmuseet, Kopenhagen; p. 69—"St. Helena," from *The Senses*, Barlow and Mollon (eds.), reprinted by permission of Cambridge University Press; p. 81—*The Innocent Eye Test* by Mark Tansey reprinted by permission of the Metropolitan Museum of Art, New York; pp. 54, 98, 138, 143, and 175—drawings by Elaine de la Mata; pp. 51, 55, 64, 65, 91, 127, 139, 152, 153, and 177—graphics by Ronald P. Bouverat.

Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book and Addison-Wesley was aware of a trademark claim, the designations have been printed in initial capital letters.

Library of Congress Cataloging-in-Publication Data

Harth, Erich.
The creative loop; how the brain makes a mind / Erich Harth.
p. cm.
Includes bibliographical references and index.
ISBN 0-201-57079-3
1. Brain. 2. Mind and body. 3. Mind-brain identity theory.
4. Biophysics. I. Title.
QP360.H34 1993
612.8'2--dc20 93-22829
CIP

Jacket design by Jean Seal
Jacket art by Lance Hidy
Text design by Wilson Graphics & Design (Kenneth J. Wilson)
Set in 10-point New Century Schoolbook by CopyRight, Inc.

1 2 3 4 5 6 7 8 9-MA-96959493
First printing, July 1993

The Creative Loop

How the Brain Makes a Mind

Also by Erich Harth

Windows on the Mind:
Reflections on the Physical Basis of Consciousness

Dawn of a Millennium: Beyond Evolution and Culture

for
Dorothy
with gratitude
and love

Preface

It was my intention to write another book on the brain. I have been intrigued with this organ for over a quarter of a century, and although I have studied its structure and function, and written about it in professional journals and in one book-length work for the general reader, I find myself as mystified as ever by the unique position it occupies for me in the universe. I am referring in particular to *my* brain, but I presume your brain occupies a similar position for you.

In thinking about this venture, I became more and more tempted to go beyond a description of how the brain helps me navigate through life, or what the underlying brain mechanisms are, and to give rein to the whole gamut of existential puzzlements that revolve around the two entities—myself on one side and the rest of the universe on the other—that interact through this three pounds of neural mass. I probably should have stifled the impulse, but the physicist in me becomes impatient when the talk is confined to neuroscience, and the neuroscientist in me grows restless when it is all about physics. In one sense, this book is to reconcile the two realms.

I believe that the questions that trouble me most are also the most puzzling to others. They become even more irksome when I try to define them for the purpose of telling the reader what is in store. To talk about "my place vis-à-vis the universe" must seem a hopelessly presumptuous, even arrogant, undertaking. But the brain that looks both in and out cannot help wondering at the strange duality of the world in which there is an *I* and an *it*, and the knowledge that in a matter of years, decades at most, there will be only *it*. We humans all share this predicament and face it in different ways.

In trying to assess the nature of my existence I came to examine a number of related puzzles. I believe that a juxtaposition of cosmic and biological evolution can help delineate our position

in the universe, and that we may gain a better fix on the human mind by comparing our brains with our computers, and by looking at our electronic robots alongside the clockwork puppets that so intrigued our ancestors two centuries ago.

Here I must confess to a profound dilemma. I wish I could unequivocally believe in a soul, because it would explain so much that is mysterious about human existence. I can't, and I am not even flirting with the idea. I also wish I could subscribe to orthodox materialism, to take the veil off the mystery of the mind-brain, and to convince you that all the wondrous peregrinations of your conscious mind not only can be understood as the workings of a complex machinery of neural switches and relays, but also can be translated, duplicated, even improved, by machinery of our design and making.

Unfortunately, I see materialism as an outdated concept, rooted in the nineteenth-century belief that all phenomena in the world could be explained as the mechanical interactions between many small indivisible and permanent material objects or *elementary particles*. Since then, the world of these supposedly indestructible units has been opened to reveal an immaterial confusion of fields, virtual states, and questionable causal relations. In a most unexpected development, this world of the ultrasmall has been linked recently to questions concerning the opposite end of the cosmic scale, the birth and evolution of the universe at large.[1] A *materialist* explanation may be appropriate for an understanding of the workings of a steam engine or even of a computer chip, but may miss the mark when we try to understand what makes us think. I would favor a *physicalist* approach, by which I mean the working assumption that the broad field of *physics* ultimately must account for most things we observe, including our own minds. This differs fundamentally from orthodox materialism because it leaves the door open to the unknown laws and relations of a yet to be explored territory.

The fields of physics and neuroscience have grown into formidable structures whose languages are all but impenetrable to the uninitiated. But when we come to questions concerning the workings of the mind, we find that the thicket of technical jargon only hides our profound ignorance, and that simple language has a better chance of cutting through some of the underbrush. We must come out into the open again and speak clearly, that is, simply. This I will try to do.

The questions raised will, I hope, be more important than the few answers I may provide. The brain is still to be the centerpiece of our discussions. How, I will ask, does it create its world of images, and how does that world compare with the corresponding *real* objects and events? What role do time and space play in the image world, the world of the mind? Is there anything at all we can say about consciousness? What is sensation? What are intentionality, creativity?

In discussing the brain I will pay particular attention to a peculiar feature of brain structure that all too often is overlooked: the existence, the ubiquity, of feedback loops. This means that what is received at any one brain level depends on what goes on at that same level, and what is sent to the next level depends on things happening at that next level. The mechanism is one of *self-reference.* I will propose ways in which these self-referent loops contribute to some of the qualities we associate with the human mind: consciousness, creativity.

I use the word *contribute* to express my belief that these mechanisms are necessary in the making of a mind. But they are not the whole story; otherwise we could build a robot tomorrow, using these same principles, and expect it to be conscious and creative.

But perhaps we are taking too much for granted. It has been suggested that consciousness may be an artifact, and that creativity, intentionality, and selfhood are illusions. Why couldn't I be replaced by a machine? Would it, *should* it, matter to me? How human can a robot be, and to what extent do humans act like puppets? Are humans disappearing from the scene, being displaced by their own more intelligent creations?

These are complex questions, and some material from various sciences will be introduced in a nontechnical way as background material, but the important discussions are of such a nature that the simplest language is as expedient as the most abstruse. By aiming for the former, I hope that a broad audience will be able to participate in this adventure.

I owe a great debt to my family for helpful criticism and support, to my students at Syracuse University who collaborated on many ideas expressed here, to Lee Smolin and K. P. Unnikrishman for many helpful comments on the manuscript, and to my editor, William Patrick, without whose guidance this would have been a lesser work.

Introduction

What
is it about
the universe,
the universe about us stretching out?
We, within our brains,
within it,
think
we must unspin it.

From *The Universe*, May Swenson

The human body is an exquisite machine, the end product of several billion years of biological evolution. We have learned to understand much of the intricate structure and function of its various parts, including those of the most complex organ of all: the human brain. But some brain functions have stubbornly eluded our attempts at reducing them to the kind of mechanistic principles that allow us to understand an organ such as the heart.

The functions we are at a loss to explain all seem to emanate from, or pertain to, a *person*, an *I*, a *subject* who not only sees and hears but *perceives* what he or she sees and hears, who recalls and projects, associates, imagines, invents, creates. He or she also *feels*—is happy or sad, hopeful or despairing, elated or depressed, angry or in love. All of these functions, and many more, are subsumed collectively under the heading the *mind of man*.[1]

They all have in common the fact that nobody has yet succeeded in explaining them as chains of logical steps or mechanistic events. Niels Bohr was aware of that when he admonished a student with "You are not thinking, you are just being logical." Unable to conceive of a way of reducing mind functions to body functions, René Descartes (1596–1650) proposed a *dualist* model of man: a mechanistic body and a mind, or *soul*, made of different stuff and subject to different laws. The two touched and interacted at a single point in the brain, the pineal gland, but otherwise carried out their functions independently of one another.

The reason Descartes's name still figures so prominently in discussions of the mind-brain dilemma is that the matter simply has not been resolved. Many schools of thought have arisen in the past 350 years. Many *isms* have been defined and written about extensively. But it is not an overstatement when I say that in all this time there has not been a true advance in the subject matter, if by advance we mean the acquisition of an understanding that is generally agreed on by the experts in the field. Certainly, no such agreement exists, not among philosophers and not among workers in the various branches of neuroscience. Dualists of more or less Cartesian persuasion still exist in all these groups. Like Descartes, they maintain that mind requires something beyond the neural machinery of our brain.

Different varieties of *materialism*, of which there are many, form the majority opinion nowadays; the philosopher Daniel Dennett calls materialism an "opinion approaching unanimity," which is an overstatement. His carefully reasoned book *Consciousness Explained*[2] has been widely quoted and praised as providing definitive answers to the most puzzling aspects of the mind-body problem. It is easy to refute the existence of a mind that is independent of a brain, and the idea that bodily actions could be influenced by a nonphysical *entity* contradicts everything we know about nature. If the body is the deterministic machine Descartes claimed it is, then only physical forces should be able to affect its actions.

Does it follow then that man is "just a machine"? This is the inevitable conclusion in what I will call *orthodox materialism*. The machine cited most often for this metaphor is the modern digital computer, which—given appropriate *software* (elaborate instructions entered into the machine memory)—can imitate any real machine or process, however complex. The brain in Dennett's phrase is such a *virtual machine*.

Still, there are thoughtful dissenting voices: *The Emperor's New Mind*, by the English physicist Roger Penrose;[3] *Bright Air, Brilliant Fire*, by the American biologist and Nobel laureate Gerald Edelman;[4] and *Consciousness Revisited*[5] and *The Rediscovery of Mind*,[6] by the American philosophers Owen J. Flanagan and John R. Searle, respectively. Why are we so reluctant to accept the materialist interpretation of mind? It is partly because such *mental* qualities as feelings and consciousness remain unexplained, and partly because we suspect that a purely physical theory of

brain function carries with it the grim implications of a clocklike determinism and predictability, or at least the absence of anything like free will. It also follows that the machinery of our brain, although intricate, can in principle be duplicated, and with it all the properties we associate with mind, including consciousness, the sensation of selfhood, imagination, and creativity. Eventually, the machine may excel in all these faculties and replace humans as the dominant species.

In the theory I will present here, I will shun the computer metaphor (which is flawed), and replace materialism (an outdated concept left over from nineteenth-century physics) with a more up-to-date *physicalism*. By this I mean the working assumption that physical processes ultimately must account for mental phenomena. For this the arsenal of contemporary physics already offers a richness of possibilities undreamed of only a few decades ago. I will propose specific mechanisms that are conceptually simple yet open the door to the unpredictable, to the flow of thought and the vagaries of creativity.

The mind-body problem is not the only dualist dilemma we face. Chasms separate the individual from the rest of the cosmos, life from the inanimate, humans from animals, man from machine, image from reality. We make the distinctions between members of each pair by painstaking definitions, only to seek unity again by building elaborate bridges across the chasms. We consider unification to be an intellectual triumph. We succeed only rarely, and every one of these dualisms is an open question in our attempt to define ourselves.

In Judeo-Christian culture, humans, until the mid-nineteenth century, viewed themselves as unique creatures, the only ones in creation made in God's image, the only ones endowed with souls. They were able to overlook the startling resemblance they bore to the rest of the animal kingdom. Man was man and beast was beast. Denying animals a soul was an expression of man's feeling that his position in the universe was both singular and solitary.

Darwin's theory of evolution changed all that. Although animals had been worshiped in antiquity, it was only after the realization of evolutionary kinship that we thought much about animal pain and emotions. Except for the rearguard action of a few creationists, we now see man firmly embedded in, and related to, all other terrestrial life.

What makes the theory of evolution so remarkable is the enormous simplification and unification it imposes on the panorama of life on earth. But a gap remains. We are not just another species. Our ability to reason, our ingenuity, and our linguistic skills place us so far above any competing animal species that many of us feel that something beyond the mechanisms of evolution must have occurred to produce *homo sapiens.* Or perhaps a unique mutation freed our brains from the constraints of instinct and gave us a *mind.* The sociobiologist Edward O. Wilson dubbed it the *Promethean gene.*[7]

What distinguishes us more than anything else, however, is our acute awareness of a *self,* and a mental preoccupation with our own being that goes far beyond the kind of self-preserving behavior that all animals exhibit. Our strong sensation of selfhood often gives rise to a feeling of ineffable solitude, an existential angst engendered by an *outside world,* the *it* that surrounds the lone *I.*

But what is the nature of the *I,* of this subjective existence? How does it come about? By what mechanisms does it arise in our brains?

Bordering the outside world are our own bodies. We call them *ours,* but they are also part of the physical world, the world of objects. We know our bodies through the senses of pain and pleasure; we are concerned about them and depend on their well-being. We could not exist without them. But if *they* are part of the world around us, who are *we,* around whom this world is displayed?

We try to escape in different ways from this painful dualism, this stark cosmic solitude. The pious find solace in the belief in eternally caring deities and in a universe designed specifically to become the home for human beings. The latter is also the assertion of the *anthropic principle,* a theory that has come out of physics and astronomy. It holds that the evolving universe, long before life appeared on earth, was a benign system from the start, its laws and the so-called constants of nature delicately balanced to make possible the emergence of life and to facilitate the evolution of man.

This almost theistic conception of a benign universe that had man in mind almost from its violent start is to be contrasted with a world that often appears brutal and uncaring. We may destroy ourselves in a nuclear holocaust, Camille Paglia points out, but "nature will absorb it all. After the bomb, nature will pick up the cards, shuffle them, and begin the game again. Nature is forever playing solitaire with herself."[8]

Nature as friend, or nature the implacable? We court nature in many ways. In natural science we seek a bond with the cosmos through knowledge. We try to strip away some of the strangeness in nature by listening carefully to her pulse, and we gain some measure of oneness through understanding. It seems to be a never-ending process. Mysticism and magic try to achieve the same thing in a different way. The mystic lays claim to a hidden, private connectedness with the universe. But the question "Where do we fit in?" persists.

Accident or providence, at some time during the physical evolution of the planet earth, life appeared as though a seed had been dropped into a sterile but fertile ground. This is indeed one of the many theories of the origin of life on earth.[9] Most scientists, however, favor the opinion that this *animation* of earth happened gradually and spontaneously, starting with very primitive prebiotic forms.

In the course of long geologic epochs amid a profusion of species, a creature evolved that was to become very different from the rest of the animal kingdom. *Homo sapiens* left his evolutionary niche some hundred thousand years ago and embarked on a course on which imagination and creativity became more valuable than swiftness and strength. Our biological similarity notwithstanding, the difference between humans and our nearest evolutionary cousins is profound. No animal ever carved the face of a human into a cave wall or gained control over fire.

Our environment has undergone the most profound changes, mostly due to our own intervention, but we remain biologically unchanged. We have every reason to believe that, if one of our forebears from the paleolithic past were brought to life, he or she would have no difficulty competing in all the skills that our technological society requires. We ascribe this adaptability, this seeming independence from biological constraints, to a unique possession, evolution's last gift to man: his *mind*.

But in what form do we possess this gift? Calling it the *Promethean gene*, as Wilson did, suggests that it is a physical, inherited characteristic, most likely residing in the brain. We therefore must examine this organ of mind.

A formidable array of disciplines, the *neurosciences*, has evolved, mostly in the last few decades, with the aim of understanding brain structure and function. The questions most fre-

quently asked—and in some instances answered—are "How does neural circuitry distinguish between different inputs?" and "How is that information used to produce different reactions?" The brain is viewed and studied as an input-output device, appropriately called the *sensorimotor brain*. Can we explain all of the manifestations of mind in this manner? Is mind a function of the brain, as digestion is a function of the stomach, or is it an excretion of the brain, as bile is an excretion of the liver?

We will seek the mind-producing powers of the brain by examining how our senses generate neural messages about the world, and how these messages are transformed on their way to higher brain centers.

In vision, man's dominant sense, *images* are picked up at the retina of each eye and transported along the visual pathway to a succession of brain centers, where they are transformed, sifted, and mixed with other information. We shall see how the pattern of neural activity starts out as a replica of the pattern of light projected onto the retina. But after a few transformations, it bears no more resemblance to the original scene, and probably could not be interpreted by anyone even if it were known in its entirety. This raises the question of how the brain is able to refer these garbled messages back to the reality outside.

How, in other words, do *I* sense, perceive, and interpret this neural cryptogram in my head? If I could observe it the way an operator of a complex machine or system observes an instrument panel of gauges and indicator lights, it would be as meaningless to me as it would be to any other observer. And yet the information becomes transformed in my head from the seemingly chaotic firing patterns involving millions of neurons into something we can talk about again: perception, recognition of familiar forms, associations, emotions. As we proceed along the sensory pathways toward higher brain centers, it becomes more meaningful to talk not about neural firing patterns but about images and thoughts. We switch from brain talk to mind talk.

This raises a dual question: How are the mental images related to the corresponding objects and events in the outside world, and how are these mental functions related to strictly physical processes in the brain? What does my thought or mental image of a giraffe have to do with a real giraffe, and what is the physical state of the brain when I think of a giraffe?

In answer to the first question we can say that mental images are not just replicas of corresponding objects in the real world. They are compounded by meanings and associations derived from a lifetime of experiences. The image does not behave, therefore, like the corresponding physical reality, and its dynamics are not constrained by the laws that govern the behavior of physical objects.

But the fact that the image of a stone does not always behave like a stone—we can make it *fall* upward if we want to—does not exclude the possibility that thoughts have a solid physical basis. If the image is not simply a mental replica of reality, can we at least identify it with another physical reality, namely, the physical state of the brain itself? It has been argued, in another grand attempt at unification, that there is a strict relationship between our thoughts and the activity of the neurons in the brain. A thought, it is argued in this *psychoneural identity theory*, is just another way of talking about a particular sequence of physical events in the brain. A mental state *is* a brain state. On this assumption, the recall of a particular event in memory also could be described as the simultaneous activity of a specified, or at least specifiable, set of neurons.

This kind of description should please the physicist. In classical physics, physical states lend themselves to precise specification and hold out hope of valid predictions once the dynamics of the system are understood. The computational task may be daunting, but, in principle, it is argued, we should be able to predict the progression of our thoughts, just as we are able to predict the trajectory of a spacecraft tumbling through a complicated gravitational field.

This is the view I call *orthodox materialism*. It is based on classical—that is, nineteenth-century—physics, which, even today, is wrongly considered by some philosophers to be *the* scientific approach. Modern physics, which began in the early part of this century with the revolutionary concepts of relativity and quantum mechanics, has fundamentally altered the physicist's outlook. The machine that runs with deterministic precision like a perfect clock is no longer an adequate description for most processes in nature. There is no reason why it should apply to the brain. Paul Davies, a noted Australian physicist and writer, put it succinctly: "Materialism is dead."[10]

This does not mean that we should not look for mechanisms operating in the vast network of neurons for clues and explanations of mental phenomena. We will do just that, and we will talk at length about one mechanism that my students and I have written about extensively. We called it the *creative loop*.

A central problem in trying to devise a physical model of the mind has to do with the question of unification of our cognitive functions. It has been felt that—because a single self appears to be doing all the seeing, hearing, and thinking—there must be a place in the brain where everything comes together.

But first everything is scattered. Our sensory systems collect messages about the outside world in the form of images or neural codes. Hierarchies of neural analyzers are tuned to pick up the presence of specific features. In the visual system alone we find a special brain center concerned with color discrimination, another with motion, and apparently many shape-specific centers. Some neurons, or neuron groups, signal the presence of a single line segment of a particular orientation and location in the field of vision. Others appear to be tuned to patterns as complex as a human face.

The information picked up by the eye as a complete replica of the physical scene outside is dispersed among many different centers by the time it reaches the highest neural levels, the *cerebral cortex*.

Although I have the distinct impression that a single *I* views and is aware of all these features, there appears to be no place in the brain where it is all reassembled into a complete image. This dilemma has given rise to the myth of an immaterial presence, a spooky homunculus, who observes the state of the physical brain. The many feature-specific analyzers are then regarded as an instrument panel with flashing lights, gauges, and other indicators from which an intelligent operator draws valid conclusions. Another metaphor philosophers use is that of a theater, a stage, populated by many actors presenting many scripts, all of it unified through the eyes of the observer. For Descartes the pineal gland was such a stage.

But if there is a stage, the single observer of it all is missing. There now is a strong consensus that the putative *unification* of sensory perceptions and their elaborations into associations, thoughts, and so on, is nothing but a figment of our imagination. The notion has become popular among philosophers and neuro-