ANTONIO DAMASIO

Looking *for* Spinoza

Joy, Sorrow, and the Feeling Brain

A HARVEST BOOK
HARCOURT, INC.
Orlando Austin New York San Diego Toronto London

Contents

hmhco.com

Library of Congress Cataloging-in-Publication Data
Damasio, Antonio R.
Looking for Spinoza: joy, sorrow, and the feeling brain/
Antonio Damasio.—1st ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-15-100557-5 ISBN 0-15-602871-9 (pbk.)
1. Emotions. 2. Neuropsychology. 3. Spinoza, Benedictus de, 1632–1677.
I. Title.
QP401.D203 2003
152.4—d21 2002011347

All figures, diagrams, and drawings are by Hanna Damasio except for the portrait on page 263. Her drawings in Chapters 1, 5, and 6 depict Spinoza's house on 72-74 Paviljoensgracht (page 9), a statue of Spinoza (page 16), the back of the New Church and Spinoza's tomb in The Hague (page 19), the Portuguese Synagogue in Amsterdam (page 185), the house where Spinoza lived in Rijnsburg (page 223), a bust of Spinoza (page 225), and the old synagogue in Amsterdam (inspired by a 17th century engraving by Jan Veenhuysen). The portrait on page 263 is by Jean Charles François and was published by A. Savérien, *Histoire des Philosophes Modernes,* Paris, 1761.

eISBN 978-0-547-54171-6

v3.0818

To Hanna

Chapter 1

Enter Feelings

Enter Feelings

Feelings of pain or pleasure or some quality in between are the bedrock of our minds. We often fail to notice this simple reality because the mental images of the objects and events that surround us, along with the images of the words and sentences that describe them, use up so much of our overburdened attention. But there they are, feelings of myriad emotions and related states, the continuous musical line of our minds, the unstoppable humming of the most universal of melodies that only dies down when we go to sleep, a humming that turns into all-out singing when we are occupied by joy, or a mournful requiem when sorrow takes over.*

Given the ubiquity of feelings, one would have thought that their science would have been elucidated long ago—what feelings are, how they work, what they mean—but that is hardly the case. Of all the mental phenomena we can describe, feelings and their essential ingredients—pain and pleasure—are the least understood in biological and specifically neurobiological terms. This is all the more puzzling considering that advanced societies cultivate feelings shamelessly and dedicate so many resources and efforts to manipulating those feelings with alcohol, drugs of abuse, medical drugs, food, real sex, virtual sex, all manner of feel-good consumption, and all manner of feel-good social and religious practices. We doctor our feelings with pills, drinks, health spas, workouts, and spiritual exercises, but

neither the public nor science have yet come to grips with what feelings are, biologically speaking.

I am not really surprised at this state of affairs, considering what I grew up believing about feelings. Most of it simply was not true. For example, I thought that feelings were impossible to define with specificity, unlike objects you could see, hear, or touch. Unlike those concrete entities, feelings were intangible. When I started musing about how the brain managed to create the mind, I accepted the established advice that feelings were out of the scientific picture. One could study how the brain makes us move. One could study sensory processes, visual and otherwise, and understand how thoughts are put together. One could study how the brain learns and memorizes thoughts. One could even study the emotional reactions with which we respond to varied objects and events. But feelings—which can be distinguished from emotions, as we shall see in the next chapter—remained elusive. Feelings were to stay forever mysterious. They were private and inaccessible. It was not possible to explain how feelings happened or where they happened. One simply could not get "behind" feelings.

As was the case with consciousness, feelings were beyond the bounds of science, thrown outside the door not just by the naysayers who worry that anything mental might actually be explained by neuroscience, but by card-carrying neuroscientists themselves, proclaiming allegedly insurmountable limitations. My own willingness to accept this belief as fact is evidenced by the many years I spent studying anything but feelings. It took me awhile to see the degree to which the injunction was unjustified and to realize that the neurobiology of feelings was no less viable than the neurobiology of vision or memory. But eventually I did,

mostly, as it turns out, because I was confronted by the reality of neurological patients whose symptoms literally forced me to investigate their conditions.

Imagine, for example, meeting someone who, as a result of damage to a certain location of his brain, became unable to feel compassion or embarrassment—when compassion or embarrassment were due—yet could feel happy, or sad, or fearful just as normally as before brain disease had set in. Would that not give you pause? Or picture a person who, as a result of damage located elsewhere in the brain, became unable to experience fear when fear was the appropriate reaction to the situation and yet still could feel compassion. The cruelty of neurological disease may be a bottomless pit for its victims—the patients and those of us who are called to watch. But the scalpel of disease also is responsible for its single redeeming feature: By teasing apart the normal operations of the human brain, often with uncanny precision, neurological disease provides a unique entry into the fortified citadel of the human brain and mind.

Reflection on the situation of these patients and of others with comparable conditions raised intriguing hypotheses. First, individual feelings could be prevented through damage to a discrete part of the brain; the loss of a specific sector of brain circuitry brought with it the loss of a specific kind of mental event. Second, it seemed clear that different brain systems controlled different feelings; damage to one area of the brain anatomy did not cause all types of feelings to disappear at once. Third, and most surprising, when patients lost the ability to express a certain emotion, they also lost the ability to experience the corresponding feeling. But the opposite was not true: Some patients who lost their ability to experience certain feelings still could express the

corresponding emotions. Could it be that while emotion and feeling were twins, emotion was born first and feeling second, with feeling forever following emotion like a shadow? In spite of their close kinship and seeming simultaneity, it seemed that emotion preceded feeling. Knowledge of this specific relationship, as we shall see, provided a window into the investigation of feelings.

Such hypotheses could be tested with the help of scanning techniques that allow us to create images of the anatomy and activity of the human brain. Step by step, initially in patients and then in both patients and people without neurological disease, my colleagues and I began to map the geography of the feeling brain. We aimed at elucidating the web of mechanisms that allow our thoughts to trigger emotional states and engender feelings.[1]

Emotion and feeling played an important but very different part in two of my previous books. *Descartes' Error* addressed the role of emotion and feeling in decision-making. *The Feeling of What Happens* outlined the role of emotion and feeling in the construction of the self. In the present book, however, the focus is on feelings themselves, what they are and what they provide. Most of the evidence I discuss was not available when I wrote the previous books, and a more solid platform for the understanding of feelings has now emerged. The main purpose of this book, then, is to present a progress report on the nature and human significance of feelings and related phenomena, as I see them now, as neurologist, neuroscientist, and regular user.

The gist of my current view is that feelings are the expression of human flourishing or human distress, as they occur in mind and body. Feelings are not a mere decoration added on to the emotions, something one might keep or

discard. Feelings can be and often are *revelations* of the state of life within the entire organism—a lifting of the veil in the literal sense of the term. Life being a high-wire act, most feelings are expressions of the struggle for balance, ideas of the exquisite adjustments and corrections without which, one mistake too many, the whole act collapses. If anything in our existence can be revelatory of our simultaneous smallness and greatness, feelings are.

How that revelation comes to mind is itself beginning to be revealed. The brain uses a number of dedicated regions working in concert to portray myriad aspects of the body's activities in the form of neural maps. This portrait is a composite, an ever-changing picture of life on the fly. The chemical and neural channels that bring into the brain the signals with which this life portrait can be painted are just as dedicated as the canvas that receives them. The mystery of how we feel is a little less mysterious now.

It is reasonable to wonder if the attempt to understand feelings is of any value beyond the satisfaction of one's curiosity. For a number of reasons, I believe it is. Elucidating the neurobiology of feelings and their antecedent emotions contributes to our views on the mind-body problem, a problem central to the understanding of who we are. Emotion and related reactions are aligned with the body, feelings with the mind. The investigation of how thoughts trigger emotions and of how bodily emotions become the kind of thoughts we call feelings provides a privileged view into mind and body, the overtly disparate manifestations of a single and seamlessly interwoven human organism.

The effort has more practical payoffs, however. Explaining the biology of feelings and their closely related emotions is likely to contribute to the effective treatment of some major

causes of human suffering, among them depression, pain, and drug addiction. Moreover, understanding what feelings are, how they work, and what they mean is indispensable to the future construction of a view of human beings more accurate than the one currently available, a view that would take into account advances in the social sciences, cognitive science, and biology. Why is such a construction of any practical use? Because the success or failure of humanity depends in large measure on how the public and the institutions charged with the governance of public life incorporate that revised view of human beings in principles and policies. An understanding of the neurobiology of emotion and feelings is a key to the formulation of principles and policies capable of reducing human distress and enhancing human flourishing. In effect, the new knowledge even speaks to the manner in which humans deal with unresolved tensions between sacred and secular interpretations of their own existence.

Now that I have sketched my main purpose, it is time to explain why a book dedicated to new ideas on the nature and significance of human feeling should invoke Spinoza in the title. Since I am not a philosopher and this book is not about Spinoza's philosophy, it is sensible to ask: why Spinoza? The short explanation is that Spinoza is thoroughly relevant to any discussion of human emotion and feeling. Spinoza saw drives, motivations, emotions, and feelings—an ensemble Spinoza called *affects*—as a central aspect of humanity. Joy and sorrow were two prominent concepts in his attempt to comprehend human beings and suggest ways in which their lives could be lived better.

The long explanation is more personal.

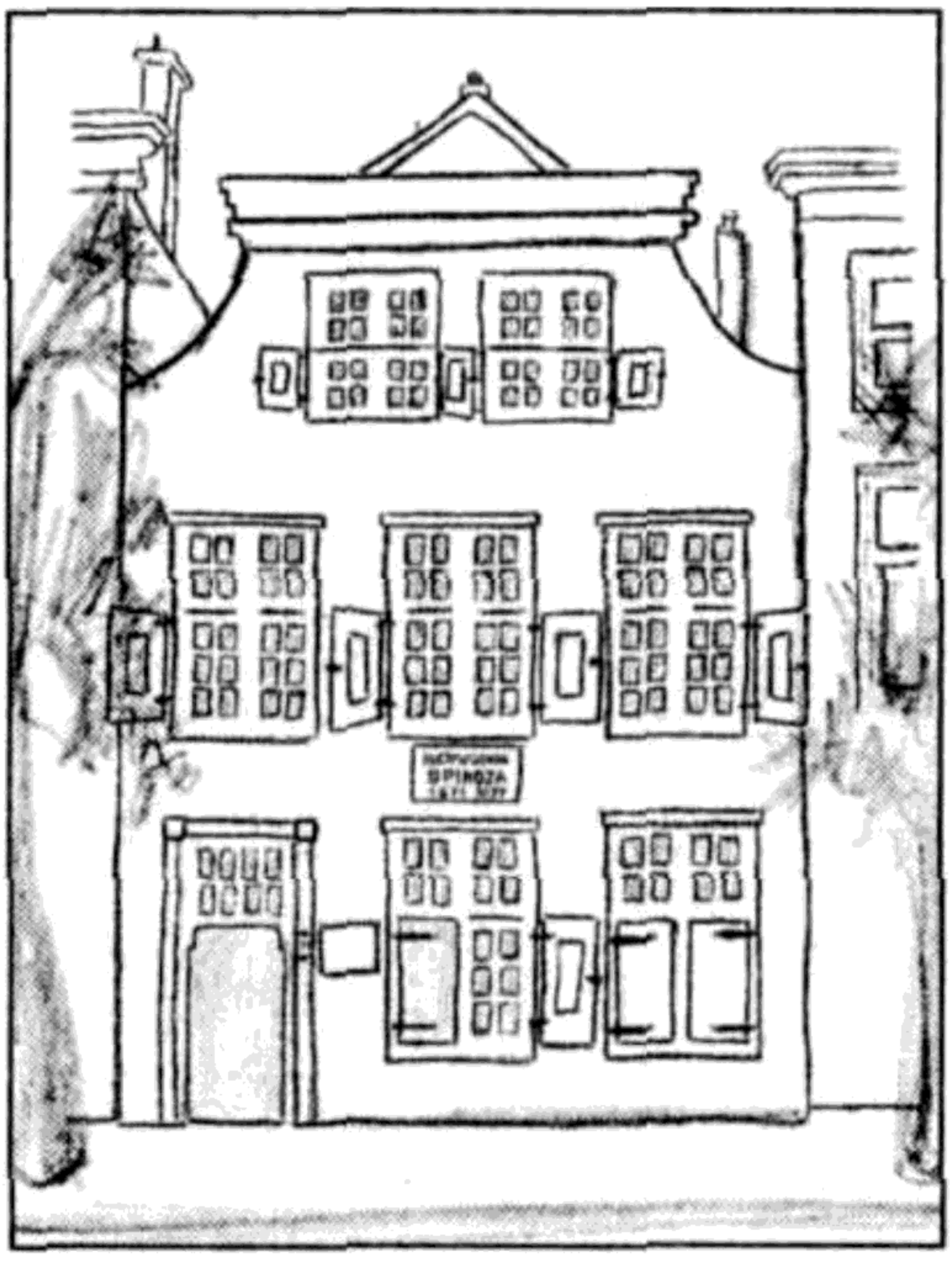
The Hague

December 1, 1999. The friendly doorman of the Hotel des Indes insists: "You should not walk in this weather, sir, let me get a car for you. The wind is bad. It is almost a hurricane, sir. Look at the flags." True, the flags have taken wing, and the fast-moving clouds are racing toward the east. Although The Hague's Embassy Row seems about to lift off, I decline the offer. I prefer to walk, I say. I will be all right. Besides, see how beautiful the sky looks in between the clouds? My doorman has no idea where I am going, and I am not going to tell him. What would he have thought?

The rain has almost stopped and with some determination it is easy to overcome the wind. I actually can walk fast and follow my mental map of the place. At the end of the promenade in front of the Hotel des Indes, to my right, I can see the old palace and the Mauritshuis, festooned with Rembrandt's face—they are showing a retrospective of his self-portraits. Past the museum square the streets are almost deserted, although this is the center of town and it is a regular working day. There must be warnings telling people to stay indoors. So much the better. I arrive at the Spui without having to brave a crowd. After I get to the New Church, the route is entirely unfamiliar and I hesitate for a second, but the choice becomes clear: I turn right on Jacobstraat, then left on Wagenstraat, then right again on Stilleverkade. Five minutes later I am on the Paviljoensgracht. I stop in front of number 72–74.

The front of the house is much as I imagined it, a small building with three floors, three windows wide, a version of the average canal townhouse, more modest than rich. It is well kept and not very different from what it must have looked like in the seventeenth century. All the windows are

closed, and there is no sign of activity. The door is well kept and well painted, and next to it there is a shiny brass bell, set in the frame. The word SPINOZAHUIS is etched in the rim. I press the button resolutely but without much hope. There is no sound from inside and no movement in any curtain. No one had answered the phone when I tried to call earlier. Spinoza is closed for business.



This is where Spinoza lived the last seven years of his brief life and where he died in 1677. The *Theologico-Political Treatise,* which he carried when he arrived, was published from here, anonymously. The *Ethics* was completed here and published after his death, almost as anonymously.

I have no hope of seeing the house today but all is not lost. In the landscaped middle section that separates the two lanes of the street, an unexpected urban garden, I discover Spinoza himself, semiobscured by the windswept foliage, sitting quietly and pensively, in sturdy bronze perpetuity. He looks pleased and entirely undisturbed by the meteorological commotion, as well he should, having survived stronger forces in his day.

For the past few years I have been looking for Spinoza, sometimes in books, sometimes in places, and that is why I am here today. A curious pastime, as you can see, and one that I had never planned to adopt. The reason why I did has a lot to do with coincidence. I first read Spinoza as an adolescent—there is no better age to read Spinoza on religion and politics—but it is fair to say that while some ideas made lasting impressions, the reverence I developed for Spinoza was rather abstract. He was both fascinating and forbidding. Later, I never thought of Spinoza as especially relevant to my work, and my acquaintance with his ideas was sparse. And yet there was a quote of his that I had long treasured—it came from the *Ethics* and pertained to the notion of self—and it was when I thought of citing it and needed to check its accuracy and context that Spinoza returned to my life. I found the quote, all right, and it did match the contents of the yellowed paper I had once pinned to a wall. But then I started reading backward and forward from the particular passage where I had landed, and I simply could not stop. Spinoza was still the same, but I was not. Much of what once seemed impenetrable now seemed familiar, strangely familiar, in fact, and quite relevant to several aspects of my recent work. I was not about to endorse all of Spinoza. For one thing, some passages were

still opaque, and there were conflicts and inconsistencies of ideas unresolved after multiple readings. I still was puzzled and even exasperated. Mostly, however, for better or worse, I found myself in a pleasant resonance with the ideas, a bit like the character in Bernard Malamud's *The Fixer,* who read a few pages of Spinoza and who kept on going as though there were a whirlwind on his back: ". . .I didn't understand every word but when you're dealing with such ideas you feel as though you were taking a witch's ride."[2] Spinoza dealt with the subjects that preoccupy me most as a scientist—the nature of emotions and feelings and the relation of mind to body—and those same subjects have preoccupied many other thinkers of the past. To my eyes, however, he seemed to have prefigured solutions that researchers are now offering on a number of these issues. That was surprising.

For example, when Spinoza said that *love is nothing but a pleasurable state, joy, accompanied by the idea of an external cause,* he was separating with great clarity the process of feeling from the process of having an idea about an object that can cause an emotion.[3] Joy was one thing; the object that caused joy was another. Joy or sorrow, along with the idea of the objects that caused either, eventually came together in the mind, of course, but they were distinct processes to begin with, within our organisms. Spinoza had described a functional arrangement that modern science is revealing as fact: Living organisms are designed with an ability to react emotionally to different objects and events. The reaction is followed by some pattern of feeling and a variation of pleasure or pain is a necessary component of feeling.

Spinoza also proposed that the power of affects is such that the only hope of overcoming a detrimental affect—an irrational passion—is by overpowering it with a stronger

positive affect, one triggered by reason. *An affect cannot be restrained or neutralized except by a contrary affect that is stronger than the affect to be restrained.*[4] In other words, Spinoza recommended that we fight a negative emotion with an even stronger but positive emotion brought about by reasoning and intellectual effort. Central to his thinking was the notion that the subduing of the passions should be accomplished by reason-induced emotion and not by pure reason alone. This is by no means easy to achieve, but Spinoza saw little merit in anything easy.

Of great importance for what I shall be discussing was his notion that both the mind and the body were parallel attributes (call them manifestations) of the very same substance.[5] At the very least, by refusing to ground mind and body on different substances, Spinoza was serving notice of his opposition to the view of the mind-body problem that prevailed in his time. His dissent stood out in a sea of conformity. More intriguing, however, was his notion that *the human mind is the idea of the human body.*[6] This raised an arresting possibility. Spinoza might have intuited the principles behind the natural mechanisms responsible for the parallel manifestations of mind and body. As I shall discuss later, I am convinced that mental processes are grounded in the brain's mappings of the body, collections of neural patterns that portray responses to events that cause emotions and feelings. Nothing could have been more comforting than coming across this statement of Spinoza's and wondering about its possible meaning.

This would have been more than enough to fuel my curiosity about Spinoza, but there was more to sustain my interest. For Spinoza, organisms naturally endeavor, of necessity, to persevere in their own being; that necessary

endeavor constitutes their actual essence. organisms come to being with the capacity to regulate life and thereby permit survival. Just as naturally, organisms strive to achieve a "greater perfection" of function, which Spinoza equates with joy. All of these endeavors and tendencies are engaged unconsciously.

Darkly, through the glass of his unsentimental and unvarnished sentences, Spinoza apparently had gleaned an architecture of life regulation along the lines that William James, Claude Bernard, and Sigmund Freud would pursue two centuries later. Moreover, by refusing to recognize a purposeful design in nature, and by conceiving of bodies and minds as made up of components that could be combined in varied patterns across different species, Spinoza was compatible with Charles Darwin's evolutionary thinking.

Armed with this revised conception of human nature, Spinoza proceeded to connect the notions of good and evil, of freedom and salvation, to the affects and to the regulation of life. Spinoza suggested that the norms that govern our social and personal conduct should be shaped by a deeper knowledge of humanity, one that made contact with the God or Nature *within* ourselves.

Some of Spinoza's ideas are part and parcel of our culture, but to the best of my knowledge Spinoza is absent as a reference from the modern efforts to understand the biology of the mind.[7] This absence is interesting in itself. Spinoza is a thinker far more famous than known. Sometimes Spinoza appears to rise out of nothing, in solitary and unexplained splendor, although the impression is false—in spite of his originality he is very much a part of his intellectual times. And he appears to dissolve as abruptly, without succession—

another false impression given that the essence of some of his forbidden proposals can be found behind the Enlightenment and well beyond in the century that followed his death.[8]

• • •

One explanation for Spinoza's status as unknown celebrity is the scandal he caused in his own time. As we shall see (in Chapter Six), his words were deemed heretical and banned for decades and with rare exceptions were quoted only as part of the assault on his work. The attacks paralyzed most attempts by Spinoza admirers to discuss his ideas publicly. The natural continuity of intellectual acknowledgment that follows a thinker's work was thus interrupted, even as some of his ideas were used uncredited. This state of affairs, however, hardly explains why Spinoza continued to gain fame but remained unknown once the likes of Goethe and Wordsworth began to champion him. Perhaps a better explanation is that Spinoza is not easy to know.

The difficulty begins with the problem that there are several Spinozas with which to reckon, at least four by my count. The first is the accessible Spinoza, the radical religious scholar who disagrees with the churches of his time, presents a new conception of God, and proposes a new road to human salvation. Next comes Spinoza as political architect, the thinker who describes the traits of an ideal democratic state populated by responsible, happy citizens. The third Spinoza is the least accessible of the set: the philosopher who uses scientific facts, a method of geometric demonstration and intuition to formulate a conception of the universe and the human beings in it.

Recognizing these three Spinozas and their web of dependencies is enough to suggest how convoluted Spinoza can be. But there is a fourth Spinoza: the protobiologist. This is the biological thinker concealed behind countless propositions, axioms, proofs, lemmas, and scholia. Given that many of the advances on the science of emotions and feeling are consonant with proposals that Spinoza began to articulate, my second purpose in this book is to connect this least-known Spinoza to some of the corresponding neurobiology of today. But I note, again, that this book is not about Spinoza's philosophy. I do not address Spinoza's thinking outside of the aspects I regard as pertinent to biology. The goal is more modest. One of the values of philosophy is that throughout its history it has prefigured science. In turn, I believe, science is well served by recognizing that historical effort.

Looking for Spinoza

Spinoza is relevant to neurobiology in spite of the fact that his reflections on the human mind came out of a larger concern for the condition of human beings. Spinoza's ultimate preoccupation was the relation of human beings to nature. He attempted to clarify that relationship so he could propose realistic means for human salvation. Some of those means were personal, under the sole control of the individual, and some relied on the help that certain forms of social and political organization provided the individual. His thinking descends from Aristotle's, but the biological grounding, not surprisingly, is firmer. Spinoza seems to have gleaned a relation between personal and collective happiness, on the one hand, and human salvation and the structure of

the state, on the other, long before John Stuart Mill. At least regarding the social consequences of his thinking there seems to be considerable recognition.[9]

Spinoza prescribed an ideal democratic state, where the hallmarks were freedom of speech—*let every man think what he wants and say what he thinks,* he wrote[10]—separation of church and state, and a generous social contract that promoted the well-being of citizens and the harmony of government. Spinoza offered this prescription more than a century ahead of the Declaration of Independence and First Amendment. That Spinoza, as a part of his revolutionary endeavors, also anticipated some aspects of modern biology is all the more intriguing.

Who was this man, then, who could think about mind and body in ways that were not only profoundly opposed to the thinking of most of his contemporaries, but remarkably current three hundred and some years later? What circumstances produced such a contrary spirit? To attempt an answer to these questions, we must consider yet another Spinoza, the man behind three distinct first names—Bento, Baruch, Benedictus—a person at once courageous and cautious, uncompromising and accommodating, arrogant and modest, detached and gentle, admirable and irritating, close to the observable and the concrete and yet unabashedly spiritual. His personal feelings are never revealed directly in his writings, not even in his style, and he must be pieced together from a thousand indirections.

Almost without noticing, I began looking for the person behind the strangeness of the work. I simply wanted to meet the man in my imagination and chat a little, have him sign *The Ethics* for me. Reporting on my search for Spinoza and the story of his life became the third purpose of this book.

Spinoza was born in the prosperous city of Amsterdam in 1632, literally in the middle of Holland's Golden Age. That same year, a brief walk from the Spinoza household, a twenty-three-year-old Rembrandt van Rijn was painting *The Anatomy Lesson of Dr. Tulp*, the picture that began his fame. Rembrandt's patron, Constantijn Huygens, statesman and poet, secretary to the Prince of Orange, and friend of John Donne, had recently become the father of Christiaan Huygens, who was to be one of the most celebrated astronomers and physicists of all time. Descartes, the leading philosopher of the day, then thirty-two, also was living in Amsterdam, on the Prinsengraacht, and worrying about how his new ideas on human nature would be received in Holland and abroad. Soon he would come to teach algebra to young Christiaan Huygens. Spinoza came into the world amid embarrassing riches, intellectual and financial, to draw on Simon Schama's apt descriptor of the place in this age.[11]

Bento was the name Spinoza received at his birth from his parents, Miguel and Hana Debora, Portuguese Sephardic Jews who had resettled in Amsterdam. He was known as Baruch in the synagogue and among friends while he was growing up in Amsterdam's affluent community of Jewish merchants and scholars. He adopted the name Benedictus at age twenty-four after he was banished by the synagogue. Spinoza abandoned the comfort of his Amsterdam family home and began the calm and deliberate errancy whose last stop was here in the Paviljoensgracht. The Portuguese name Bento, the Hebrew name Baruch, and the Latin name Benedictus, all mean the same: blessed. So, what's in a name? Quite a lot, I would say.

The words may be superficially equivalent, but the concept behind each of them was dramatically different.

Beware

I need to get inside the house, I think, but for now the door is closed. All I can do is imagine someone emerging from a barge moored close to it, walking into the house, and inquiring after Spinoza (the Paviljoensgracht was a wide canal, in those days; later it was filled in and turned into a street, as were so many canals in Amsterdam and Venice). The wonderful Van der Spijk, the owner and a painter, would open the door. He would amiably usher the visitor into his studio, behind the two windows next to the main door, invite him to wait, and go tell Spinoza, his lodger, that a caller had arrived.

Spinoza's rooms were on the third floor, and he would come down the spiral staircase, one of those tightly curled, horrifying stairs for which Dutch architecture is infamous. Spinoza would be elegantly dressed in his *fidalgo* garb—nothing new, nothing very worn, all well kept, a white starched collar, black breeches, a black leather vest, a black camel-hair jacket nicely balanced on his shoulders, shiny black leather shoes with large silver buckles, and a wood cane, perhaps, to help negotiate the stairs. Spinoza had a thing for black leather shoes. Spinoza's harmonious and cleanshaven face, his large black eyes shining brilliantly, would dominate his appearance. His hair was black too, as were the long eyebrows; the skin was olive; the stature medium; the frame light.

With politeness, even affability, but with economic directness, the visitor would be prompted to come to the

matter at hand. This generous teacher could entertain discussions of optics, politics, and religious faith during his office hours. Tea would be served. Van der Spijk would continue painting, mostly silently, but with salubrious democratic dignity. His seven ebullient children would stay out of the way in the back of the house. Mrs. Van der Spijk sewed. The help toiled away in the kitchen. You see the picture.

Spinoza would be smoking his pipe. The aroma would do battle with the fragrance of turpentine as questions were pondered, answers given, and daylight waned. Spinoza received countless visitors, from neighbors and relatives of the Van der Spijks to eager young male students and impressionable young women, from Gottfried Leibniz and Christiaan Huygens to Henry Oldenburg, president of the newly created Royal Society of Britain. Judging from the tone of his correspondence he was most charitable with the simple folk and least patient with his peers. Apparently he could suffer modest fools easily but not the other kind.

I also can imagine a funeral cortege, on another gray day, February 25, 1677, Spinoza's simple coffin, followed by the Van der Spijk family, and "many illustrious men, six carriages in all," marching slowly to the New Church, just minutes away. I walk back to the New Church retracing their likely route. I know Spinoza's grave is in the churchyard, and from the house of the living I may as well go to the house of the dead.

Gates surround the churchyard but they are wide open. There is no cemetery to speak of, only shrubs and grass and moss and muddy lanes amid the tall trees. I find the grave much where I thought it would be, in the back part of the yard,

behind the church, to the south and east, a flat stone at ground level and a vertical tombstone, weathered and unadorned. Besides announcing whose grave it is, the inscription reads *Caute!* which is Latin for "Be careful!" This is a chilling bit of advice considering Spinoza's remains are not really inside the tomb, and that his body was stolen, no one knows by whom, sometime after the burial when the corpse lay inside the church. Spinoza had told us that every man should think what he wants and say what he thinks, but not so fast, not quite yet. Be careful. Watch out for what you say (and write) or not even your bones will escape.

Spinoza used *caute* in his correspondence, printed just beneath the drawing of a rose. For the last decade of his life his written words were indeed *sub-rosa.* He listed a fictitious printer for the *Tractatus,* along with an incorrect city of

publication (Hamburg). The author's page was blank. Even so, and even though the book was written in Latin rather than Dutch, authorities in Holland prohibited it in 1674. Predictably, it also was placed in the Vatican's Index of dangerous books. The church considered the book an all-out assault on organized religion and the political power structure. After that Spinoza refrained from publishing altogether. No surprise. His last writings still were in the drawer of his desk on the day of his death, but Van der Spijk knew what to do: He shipped the entire desk aboard a barge to Amsterdam where it was delivered to Spinoza's real publisher, John Rieuwertz. The collection of posthumous manuscripts—the much-revised *Ethics,* a *Hebrew Grammar,* the second and unfinished *Political Treatise,* and the *Essay on the Improvement of the Understanding*—was published later that same year, anonymously. We should keep this situation in mind when we describe the Dutch provinces as the haven of intellectual tolerance. Without a doubt they were, but the tolerance had its limits.

For most of Spinoza's life Holland was a republic, and during Spinoza's mature years the Grand Pensionary Jan De Witt dominated political life. De Witt was ambitious and autocratic but also was enlightened. It is not clear how well he knew Spinoza, but he certainly knew of Spinoza and probably helped contain the ire of the more conservative Calvinist politicians when the *Tractatus* began to cause scandal. De Witt owned a copy of the book since 1670. He is rumored to have sought the philosopher's opinion on political and religious matters, and Spinoza is rumored to have been pleased by the esteem De Witt showed him. Even if the rumors are untrue, there is little question De Witt was interested in Spinoza's political thinking and at least

sympathetic to his religious views. Spinoza felt justifiably protected by De Witt's presence.

Spinoza's sense of relative safety came to an abrupt close in 1672 during one of the darkest hours of Holland's golden age. In a sudden turn of events, of the sort that define this politically volatile era, De Witt and his brother were assassinated by a mob, on the false suspicion that they were traitors to the Dutch cause in the ongoing war with France. Assailants clubbed and knifed both De Witts as they dragged them on the way to the gallows, and by the time they arrived there was no need to hang them anymore. They proceeded to undress the corpses, suspend them upside down, butcher-shop style, and quarter them. The fragments were sold as souvenirs, eaten raw, or eaten cooked, amid the most sickening merriment. All this took place not far from where I am standing now, literally around the corner from Spinoza's home, and it was probably Spinoza's darkest hour as well. The attacks shocked many thinkers and politicians of the time. Leibniz was horrified and so was the unflappable Huygens, in the safety of Paris. But Spinoza was undone. The savagery revealed human nature at its shameful worst and jolted him out of the equanimity he had worked so hard to maintain. He prepared a placard that read ULTIMI BARBORORUM (Ultimate barbarians) and wanted to post it near the remains. Fortunately Van der Spijk's dependable wisdom prevailed. He simply locked the door and kept the key, and Spinoza was thus prevented from leaving the house and facing a certain death. Spinoza cried publicly—the only time, it is said, that others saw him in the throes of uncontrolled emotion. The intellectual safe harbor, such as it was, had come to an end.

I look at Spinoza's grave one more time and am reminded of the inscription Descartes prepared for his own tombstone:

"He who hid well, lived well."[12] Only twenty-seven years separate the death of these two part-time contemporaries (Descartes died in 1650). Both spent most of their lives in the Dutch paradise, Spinoza by birthright, the other by choice—Descartes had decided early in his career that his ideas were likely to clash with the Catholic Church and monarchy in his native France and left quietly for Holland. Yet both had to hide and pretend, and in the case of Descartes, perhaps distort his own thinking. The reason should be clear. In 1633, one year after Spinoza's birth, Galileo was questioned by the Roman Inquisition and placed under house arrest. That same year Descartes withheld publication of his *Treatise of Man* and, even so, had to respond to vehement attacks on his views of human nature. By 1642, in contradiction with his earlier thinking, Descartes was postulating an immortal soul separate from the perishable body, perhaps as a preemptive measure to forestall further attacks. If that was his intent, the strategy eventually worked, but not quite in his lifetime. Later he made his way to Sweden to mentor the spectacularly irreverent Queen Christina. He died midway through his first winter in Stockholm, at age fifty-four. Amid the thanks we must give for living in different times, even today one shudders to think of the threats against such hard-won freedoms. Perhaps *caute* still is in order.

As I leave the churchyard, my thoughts turn to the bizarre significance of this burial site. Why is Spinoza, who was born a Jew, buried next to this powerful Protestant church? The answer is as complicated as anything else having to do with Spinoza. He is buried here, perhaps, because having been expelled by his fellow Jews he could be seen as Christian by default; he certainly could not have been buried in the Jewish cemetery at Ouderkerk. But he is *not* really here, perhaps,

because he never became a proper Christian, Protestant or Catholic, and in the eyes of many he was an atheist. And how fitting it all is. Spinoza's God was neither Jewish nor Christian. Spinoza's God was everywhere, could not be spoken to, did not respond if prayed to, was very much in every particle of the universe, without beginning and without end. Buried and unburied, Jewish and not, Portuguese but not really, Dutch but not quite, Spinoza belonged nowhere and everywhere.

Back at the Hotel des Indes the doorman is glad to see me in one piece. I can't resist. I do tell him that I am looking for Spinoza, that I have been to his house. The solid Dutchman is taken aback. He stops in bewilderment and utters, after a pause, "You mean . . . the philosopher?" Well, he does know who Spinoza was, after all, Holland being one of the best educated places on earth. But he has no idea that Spinoza lived the last part of his life in The Hague, finished his most important work here, died here, is buried here—well, sort of—and has a house and a statue and a tomb to his credit here, a mere twelve blocks away. To be fair, few people have any idea of this either. "They don't speak much of him, these days," says my friendly doorman.

In the Paviljoensgracht

Two days later I return to 72 Paviljoensgracht, and this time my gracious hosts have arranged for me to visit the house. The weather is even worse today and something like a hurricane has been blowing in from the North Sea.

Van der Spijk's studio is only marginally warmer and certainly darker than outside. A mush of gray and green

remains in my mind. It is a small space, easy to commit to memory, and easy to play with in one's imagination. Mentally, I rearrange the furniture, relight the room, and warm it up. I sit long enough to imagine the movements of Spinoza and Van der Spijk on this confined stage, and conclude that no amount of redecoration will turn the room into the comfortable salon that Spinoza deserved. It is a lesson in modesty. In this small space Spinoza received his countless visitors, Leibniz and Huygens included. In this small space Spinoza took his meals—when he was not too distracted with his work and forgot all about eating—and talked to Van der Spijk's wife and to their noisy children. In this small space he sat crushed by the news of the De Witts' assassination.

How could Spinoza have survived this confinement? No doubt by freeing himself in the infinite expanse of his mind, a place larger and no less refined than Versailles and its gardens, where, on those very same days, Louis XIV, barely six years younger than Spinoza and destined to survive him by another thirty, would be strolling with his large retinue in tow.

It must be that Emily Dickinson was right, that one single brain, being wider than the sky, can comfortably accommodate a good man's intellect and the whole world besides.

CHAPTER 2

Of Appetites and Emotions

Trust Shakespeare

Trust Shakespeare to have been there before. Toward the end of *Richard II*, the crown now lost and the prospect of jail looming close, Richard unwittingly tells Bolingbroke about a possible distinction between the notion of emotion and that of feeling.[1] He asks for a looking glass, confronts his face, and studies the spectacle of ravage. Then he notes that the "external manner of laments" expressed in his face is merely "shadows of the unseen grief," a grief that "swells with silence in the tortured soul." His grief, as he says, "lies all within." In just four lines of verse, Shakespeare announces that the unified and apparently singular process of affect, which we often designate casually and indifferently as emotion *or* feeling, can be analyzed in parts.

My strategy for elucidating feelings capitalizes on this distinction. It is true that the common usage of the word emotion tends to encompass the notion of feeling. But in our attempt to understand the complex chain of events that begins with emotion and ends up in feeling, we can be helped by a principled separation between the part of the process that is made public and the part that remains private. For the purposes of my work I call the former part *emotion* and the latter part *feeling* in keeping with the meaning of the term feeling I outlined earlier. I ask the reader to accompany me in this choice of words and concepts for the good reason that it may permit us to uncover something about the biology that lies beneath. By the end of chapter 3, I promise to put emotion and feeling back together again.[2]

In the context of this book then, emotions are actions or movements, many of them public, visible to others as they occur in the face, in the voice, in specific behaviors. To be sure, some components of the emotion process are not visible to the naked eye but can be made "visible" with current scientific probes such as hormonal assays and electrophysiological wave patterns. Feelings, on the other hand, are always hidden, like all mental images necessarily are, unseen to anyone other than their rightful owner, the most private property of the organism in whose brain they occur.

Emotions play out in the theater of the body. Feelings play out in the theater of the mind.[3] As we shall see, emotions and the host of related reactions that underlie them are part of the basic mechanisms of life regulation; feelings also contribute to life regulation, but at a higher level. Emotions and related reactions seem to precede feelings in the history of life. Emotions and related phenomena are the foundation for feelings, the mental events that form the bedrock of our minds and whose nature we wish to elucidate.

Emotions and feelings are so intimately related along a continuous process that we tend to think of them, understandably, as one single thing. In the normal situation, however, we can glean different segments along the continuous process and, under the microscope of cognitive neuroscience, it is legitimate to dissociate one segment from the other. With naked eyes and a slew of scientific probes, an observer may objectively examine the behaviors that make up an emotion. In effect, the prelude to the process of feeling can be studied. Turning emotion and feeling into separate research objects helps us discover how it is that we feel.

The goal of this chapter is to explain the brain and body mechanisms responsible for triggering and executing an emotion. The focus here is on the intrinsic "machinery of emotion" rather than the circumstances leading to emotion. I expect the elucidation of emotions to tell us how feelings come about.

Emotions Precede Feelings

In discussing the precedence of emotion over feeling let me begin by calling attention to something Shakespeare left ambiguous in his lines for Richard. It has to do with the word shadow and with the possibility that while emotion and feeling are distinct, the latter comes before the former. The external laments are a shadow of the unseen grief, says Richard, some sort of mirror reflection of the principal object —the feeling of grief—just as Richard's face in the mirror is a reflection of the play's principal object, Richard. This ambiguity resonates well with one's untutored intuitions. We tend to believe that the hidden is the source of the expressed. Besides, we know that as far as the mind is concerned, feeling is what really counts. "There lies the substance," says Richard, speaking of his hidden grief, and we agree. We suffer or delight from actual feelings. In the narrow sense, emotions are externalities. But "principal" does not mean "first" and does not mean "causative." The centrality of feeling obscures the matter of how feelings arise and favors the view that somehow feelings occur first and are expressed subsequently in emotions. That view is incorrect, and it is to blame, at least in part, for the delay in finding a plausible neurobiological account for feelings.

It turns out that it is feelings that are mostly shadows of the external manner of emotions. Here is what Richard should have said, in effect, with due apologies to Shakespeare: "Oh, how this external manner of laments casts an intolerable and unseen shadow of grief in the silence of my tortured soul." (Which reminds me of James Joyce when he says in *Ulysses,* "Shakespeare is the happy hunting ground of all minds that have lost their balance."[4])

It is legitimate to ask at this point why emotions precede feelings. My answer is simple: We have emotions first and feelings after because evolution came up with emotions first and feelings later. Emotions are built from simple reactions that easily promote the survival of an organism and thus could easily prevail in evolution.

In brief, those whom the gods wanted to save they first made smart, or so it would seem. Long before living beings had anything like a creative intelligence, even before they had brains, it is as if nature decided that life was both very precious and very precarious. We know that nature does not operate by design and does not decide in the way artists and engineers do, but this image gets the point across. All living organisms from the humble amoeba to the human are born with devices designed to solve *automatically,* no proper reasoning required, the basic problems of life. Those problems are: finding sources of energy; incorporating and transforming energy; maintaining a chemical balance of the interior compatible with the life process; maintaining the organism's structure by repairing its wear and tear; and fending off external agents of disease and physical injury. The single word homeostasis is convenient shorthand for the ensemble of regulations and the resulting state of regulated life.[5]

In the course of evolution the innate and automated equipment of life governance—the homeostasis machine—became quite sophisticated. At the bottom of the organization of homeostasis we find simple responses such as *approaching* or *withdrawing* of an entire organism relative to some object; or increases in activity *(arousal) or* decreases in activity *(calm* or *quiescence).* Higher up in the organization we find *competitive* or *cooperative* responses.[6] We can picture the homeostasis machine as a large multibranched tree of phenomena charged with the automated regulation of life. In multicellular organisms, working our way from the ground up, here is what we will find in the tree.

In the lowest branches

- The process of metabolism. This includes chemical and mechanical components (e.g., endocrine/hormonal secretions; muscular contractions related to digestion, and so forth) aimed at maintaining the balance of internal chemistries. These reactions govern, for example, heart rate and blood pressure (which help the proper distribution of blood flow in the body); adjustments of acidity and alkalinity in the internal milieu (the fluids in the bloodstream and in the spaces between cells); and the storage and deployment of proteins, lipids, and carbohydrates required to supply the organism with energy (necessary for motion, manufacture of chemical enzymes, and maintenance and renewal of its structure).
- Basic reflexes. This includes the startle reflex, which organisms deploy in reaction to a noise or touch or as the tropisms or taxes that guide organisms away from extreme heat or extreme cold, away from dark and into

light.

- The immune system. It is prepared to ward off viruses, bacteria, parasites, and toxic chemical molecules invading from outside the organism. Curiously, it also is prepared to deal with chemical molecules normally contained in healthy cells in the body that can become dangerous to the organism when released from dying cells into the internal milieu (e.g., breakdown of hyaluron; glutamate). In brief, the immune system is a first line of defense of the organism when its integrity is menaced from outside or from within.

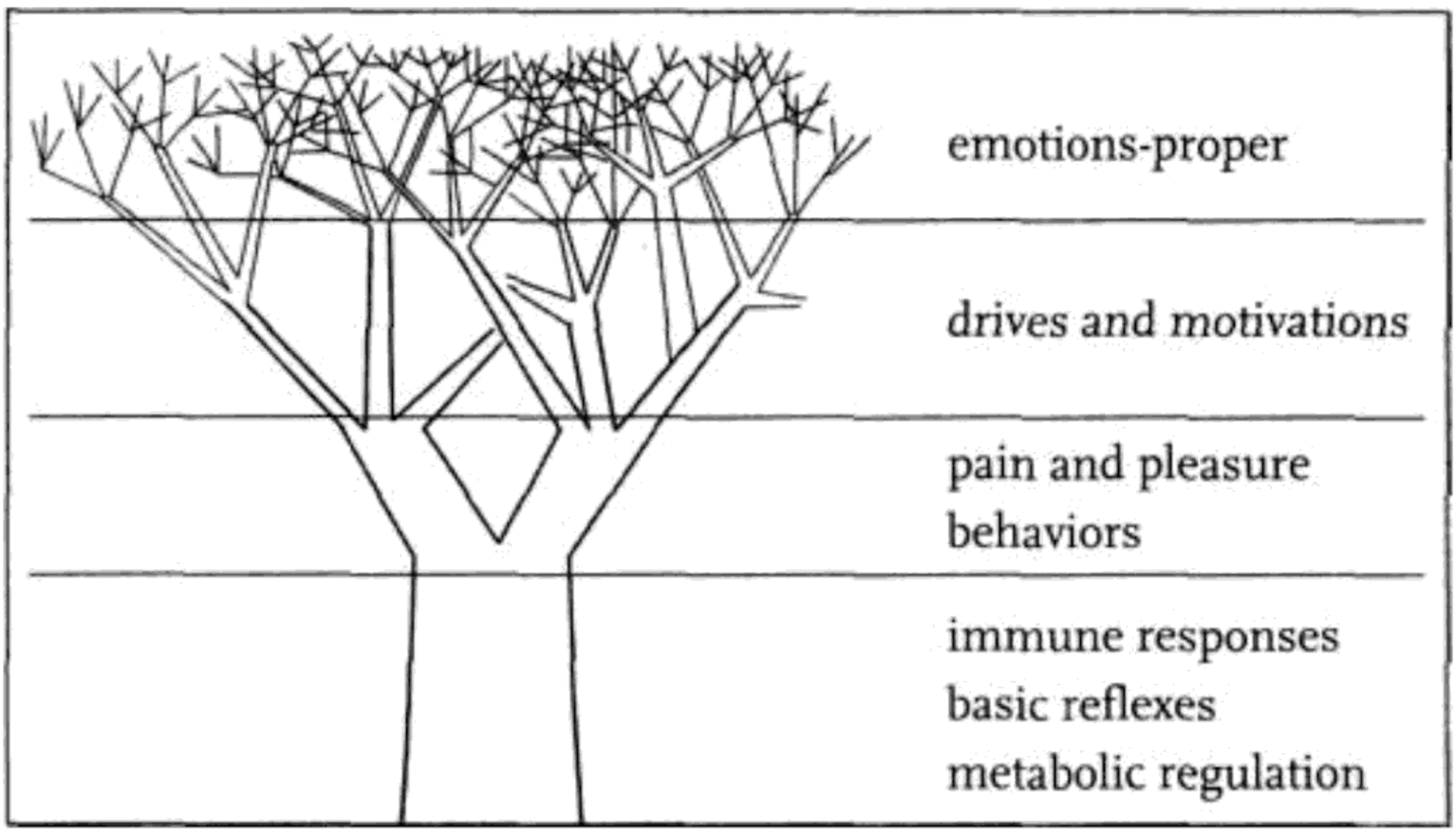


Figure 2.1: Levels of automated homeostatic regulation, from simple to complex.

In the middle-level branches

- Behaviors normally associated with the notion of pleasure (and reward) or pain (and punishment). These include reactions of approach or withdrawal of the whole organism relative to a specific object or situation. In humans, who can both feel and report what is felt,

such reactions are described as painful or pleasurable, rewarding or punishing. For example, when there is malfunction and impending damage to tissues in the body—as happens in a local burn or infection—cells in the affected region emit chemical signals that are called nociceptive (this means "indicative of pain"). In response, the organism automatically reacts with *pain behaviors* or *sickness behaviors.* These are packages of actions, clearly visible or subtle, with which nature automatically counters the insult. Such actions include withdrawal of the whole body or a part thereof from the source of trouble if that source is external and identifiable; protection of the affected body part (holding a hand that has been wounded; hugging the chest or abdomen); and facial expressions of alarm and suffering. There also is a host of responses invisible to the naked eye and organized by the immune system. Those include increasing certain classes of white blood cells, dispatching those cells to the body areas in danger, and producing chemicals such as cytokines that help solve the problem the body is facing (fight off an invading microbe, repair damaged tissue). The ensemble of these actions and the chemical signals involved in their production form the basis for what we experience as *pain.*

In the same way the brain reacts to a problem in the body, it also reacts to the good function of that body. When the body operates smoothly, without hitch and with ease of transformation and utilization of energy, it behaves with a particular style. The approach to others is facilitated. There is relaxation and opening of the body frame, facial expressions of confidence and well-being, and production of certain

classes of chemicals, such as endorphins, which are as invisible to the naked eye as some of the reactions in pain and sickness behaviors. The ensemble of these actions and the chemical signals associated with them form the basis for the experience of *pleasure.*

Pain or pleasure are prompted by many causes—glitches in some body function, optimal operation of metabolic regulation, or from external events that damage the organism or protect it. But the *experience* of pain or pleasure is *not the cause of the pain or pleasure behaviors,* and is by no means necessary for those behaviors to occur. As we will see in the next section, very simple creatures can carry out some of these emotive behaviors even if the likelihood of feeling those behaviors is low or nil.

In the next level up

- A number of *drives and motivations.* Major examples include hunger, thirst, curiosity and exploration, play and sex. Spinoza lumped them together under a very apt word, *appetites,* and with great refinement used another word, *desires,* for the situation in which conscious individuals become cognizant of those *appetites.* The word appetite designates the behavioral state of an organism engaged by a particular drive; the word desire refers to the conscious feelings of having an appetite and the eventual consummation or thwarting of the appetite. This Spinozian distinction is a nice counterpart for the distinction between emotion and feeling with which we started this chapter. Obviously humans have both appetites and desires just as seamlessly connected as emotions and feelings.

Near the top but not quite

- Emotions-proper. This is where we find the crown jewel of automated life regulation: emotions in the narrow sense of the term—from joy and sorrow and fear, to pride and shame and sympathy. And in case you wonder what we find at the very top, the answer is simple: feelings, which we will address in the next chapter.

The genome makes certain that all of these devices are active at birth, or shortly thereafter, with little or no dependence on learning, although as life continues learning will play an important role in determining *when* the devices are deployed. The more complex the reaction, the more this holds true. The package of reactions that constitutes crying and sobbing is ready and active at birth; what we cry *for*, across a lifetime, changes with our experience. All of these reactions are automatic and largely stereotyped, and are engaged under specific circumstances. (Learning, however, can modulate the execution of the stereotyped pattern. Our laughter or crying *plays* differently in different circumstances, just as the musical notes that constitute a movement of a sonata can be played in very different ways.) All of these reactions are aimed, in one way or another, directly or indirectly, at regulating the life process and promoting survival. Pleasure and pain behaviors, drives and motivations, and emotions-proper are sometimes referred to as emotions in the broad sense, which is both understandable and reasonable given their shared form and regulatory goal.[7]

Not content with the blessings of mere survival, nature seems to have had a nice afterthought: The innate equipment

of life regulation does not aim for a neither-here-nor-there neutral state midway between life and death. Rather, the goal of the homeostasis endeavor is to provide a better than neutral life state, what we as thinking and affluent creatures identify as wellness and *well-being.*

The entire collection of homeostatic processes governs life moment by moment in every cell of our bodies. This governance is achieved by means of a simple arrangement: First, something changes in the environment of an individual organism, internally or externally. Second, the changes have the potential to alter the course of the life of the organism (they can constitute a threat to its integrity, or an opportunity for its improvement). Third, the organism detects the change and acts accordingly, in a manner designed to create the most beneficial situation for its own self-preservation and efficient functioning. All reactions operate under this arrangement and are thus a means to *appraise* the internal and external circumstances of an organism and act accordingly. They detect trouble or detect opportunity and solve, by means of action, the problem of getting rid of the trouble or reaching out for the opportunity. Later, we shall see that even in "emotions-proper"—emotions such as sadness, or love, or guilt—the arrangement remains, except that the complexity of the appraisal and response are far greater than with the simple reactions from which such emotions were pieced together in evolution.

It is apparent that the continuous attempt at achieving a state of positively regulated life is a deep and defining part of our existence—the first reality of our existence as Spinoza intuited when he described the relentless endeavor *(conatus)* of each being to preserve itself. Striving, endeavor, and tendency are three words that come close to rendering the

Latin term *conatus*, as used by Spinoza in Propositions 6, 7, and 8 of the *Ethics*, Part III. In Spinoza's own words: "Each thing, as far as it can by its own power, strives to persevere in its being" and "The striving by which each thing strives to persevere in its being is nothing but the actual essence of the thing." Interpreted with the advantages of current hindsight, Spinoza's notion implies that the living organism is constructed so as to maintain the coherence of its structures and functions against numerous life-threatening odds.

The conatus subsumes both the impetus for self-preservation in the face of danger and opportunities and the myriad actions of self-preservation that hold the parts of a body together. In spite of the transformations the body must undergo as it develops, renews its constituent parts, and ages, the conatus continues to form the *same* individual and respect the *same* structural design.

What is Spinoza's *conatus* in current biological terms? It is the aggregate of dispositions laid down in brain circuitry that, once engaged by internal or environmental conditions, seeks both survival and well-being. In the next chapter, we shall see how the large compass of activities of the *conatus* is conveyed to the brain, chemically and neurally. This is accomplished by chemical molecules transported in the bloodstream, as well as by electrochemical signals transmitted along nerve pathways. Numerous aspects of the life process can be so signaled to the brain and represented there in numerous maps made of circuits of nerve cells located in specific brain sites. By that point we have reached the treetops of life regulation, the level at which feelings begin to coalesce.

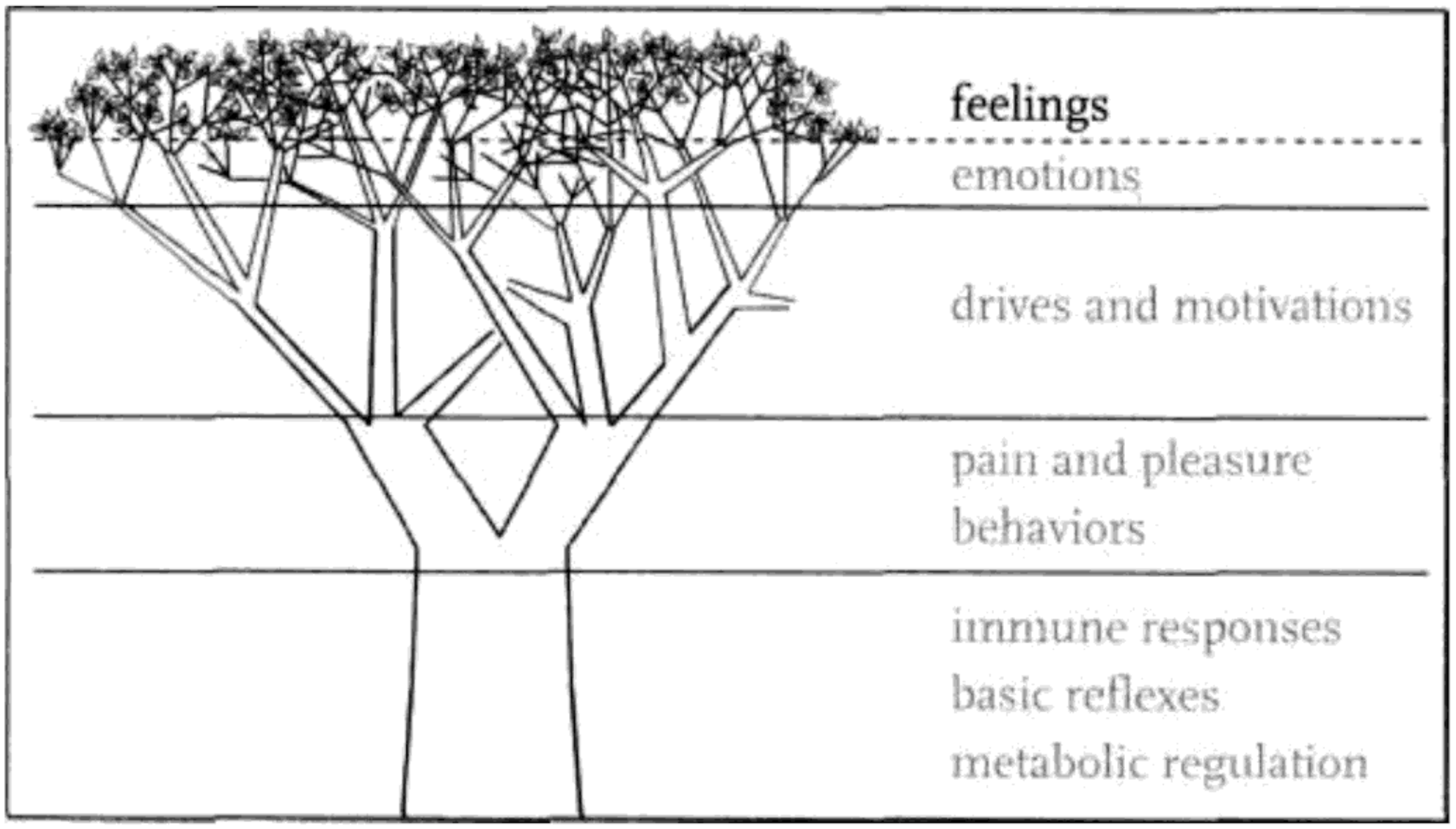


Figure 2.2: Feelings support yet another level of homeostatic regulation. Feelings are a mental expression of all other levels of homeostatic regulation.

A Nesting Principle

When we survey the list of regulatory reactions that ensure our homeostasis we glean a curious construction plan. It consists of having parts of simpler reactions incorporated as components of more elaborate ones, a nesting of the simple within the complex. *Some* of the machinery of the immune system and of metabolic regulation is incorporated in the machinery of pain and pleasure behaviors. *Some* of the latter is incorporated in the machinery of drives and motivations (most of which revolve around metabolic corrections and all of which involve pain or pleasure). *Some* of the machinery from all the prior levels—reflexes, immune responses, metabolic balancing, pain or pleasure behaviors, drives—is incorporated in the machinery of the emotions-proper. As we shall see, the different tiers of emotions-proper are assembled on the very same principle. The ensemble does not look exactly like a neat Russian doll because the bigger part is not merely an enlargement of the smaller part nested in it.

Nature is never that tidy. But the "nesting" principle holds. Each of the different regulatory reactions we have been considering is not a radically different process, built from scratch for a specific purpose. Rather, each reaction consists of tinkered rearrangements of bits and parts of the simpler processes below. They are all aimed at the same overall goal—survival with well-being—but each of the tinkered rearrangements is secondarily aimed at a new problem whose solution is necessary for survival with well-being. The solution of each new problem is required for the overall goal to be achieved.

The image for the ensemble of these reactions is not that of a simple linear hierarchy. That is why the metaphor of a tall building with many floors only captures some of the biological reality. The image of the great chain of being is not good either. A better image is that of a tall, messy tree with progressively higher and more elaborate branches coming off the main trunks and thus maintaining a two-way communication with their roots. The history of evolution is written all over that tree.

More on the Emotion-Related Reactions: From Simple Homeostatic Regulation to Emotions-Proper

Some of the regulatory reactions we have been considering respond to an object or situation in the environment—a potentially dangerous situation; or an opportunity for feeding or mating. But some of the reactions respond to an object or situation *within* the organism. This can be a drop in the amount of available nutrients for the production of energy, causing the appetitive behaviors known as hunger and

including the search for food. Or it could be a hormonal change that prompts the searching for a mate, or a wound that causes the reactions we call pain. The range of reactions encompasses not only highly visible emotions such as fear or anger, but also drives, motivations, and behaviors associated with pain or pleasure. They all occur within an organism, a body limited by a boundary, within which life ticks away. All of the reactions, directly or indirectly, exhibit an apparent aim: making the internal economy of life run smoothly. The amount of certain chemical molecules must be maintained within certain ranges, not higher and not lower, because outside those ranges life is in peril. Temperature also must be maintained within narrow parameters. Sources of energy must be procured—and curiosity and exploration strategies help locate those sources. Once found, those sources of energy must be incorporated—literally, placed inside the body—and modified for immediate consumption or storage; waste products resulting from all the modifications must be eliminated; and repair of the tissue wear and tear must be carried out so that the integrity of the organism is maintained.

Even the emotions-proper—disgust, fear, happiness, sadness, sympathy, and shame—aim directly at life regulation by staving off dangers or helping the organism take advantage of an opportunity, or indirectly by facilitating social relations. I am not suggesting every time we engage an emotion we are promoting survival and well-being. Not all emotions are alike in their potential to promote survival and well-being, and both the context in which an emotion is engaged and the intensity of the emotion are important factors in the potential value of an emotion on a specific occasion. But the fact that the deployment of some emotions

in current human circumstances may be maladaptive does not deny their evolutionary role in advantageous life regulation. Anger is mostly counterproductive in modern societies, and so is sadness. Phobias are a major hindrance. And yet think of how many lives have been saved by fear or anger in the right circumstances. These reactions are likely to have prevailed in evolution because they automatically supported survival. They still do, and that is probably why they remain part and parcel of the daily existence of human as well as nonhuman species.

On a practical note, understanding the biology of emotions and the fact that the value of each emotion differs so much in our current human environment, offers considerable opportunities for understanding human behavior. We can learn, for example, that some emotions are terrible advisors and consider how we can either suppress them or reduce the consequences of their advice. I am thinking, for example, that reactions that lead to racial and cultural prejudices are based in part on the automatic deployment of social emotions evolutionarily meant to detect *difference* in others because difference may signal risk or danger, and promote withdrawal or aggression. That sort of reaction probably achieved useful goals in a tribal society but is no longer useful, let alone appropriate, to ours. We can be wise to the fact that our brain still carries the machinery to react in the way it did in a very different context ages ago. And we can learn to disregard such reactions and persuade others to do the same.

The Emotions of Simple Organisms

There is abundant evidence of "emotional" reactions in simple organisms. Think of a lone paramecium, a simple unicellular organism, all body, no brain, no mind, swimming speedily away from a possible danger in a certain sector of its bath—maybe a poking needle, or too many vibrations, or too much heat, or too little. Or the paramecium may be swimming speedily along a chemical gradient of nutrients toward the sector of the bath where it can have lunch. This simple organism is designed to detect certain signs of danger —steep variations in temperature, excessive vibrations, or the contact of a piercing object that might rupture its membrane —and react by proceeding to a safer, more temperate, quieter place. Likewise, it will swim in the trail toward greener water pastures after detecting the presence of chemical molecules it needs for energy supply and chemical balance. The events I am describing in a brainless creature already contain the essence of the process of emotion that we humans have—detection of the presence of an object or event that recommends avoidance and evasion or endorsement and approach. The ability to react in this manner was not taught —there is not much pedagogy going on in paramecium school. It is contained in the apparently simple and yet so complicated gene-given machinery inside the unbrained paramecium. This shows that nature has long been concerned with providing living organisms with the means to regulate and maintain their lives automatically, no questions asked, no thoughts needed.

Having a brain, even a modest brain, is helpful for survival, of course, and indispensable if the environment is more challenging than the paramecium's. Think of a tiny fly—a small creature with a small nervous system but no spine. You can make the fly quite angry if you swat it repeatedly and

unsuccessfully. It will buzz around you in daredevil supersonic dives and avoid the fatal swat. But you also can make the fly happy if you feed it sugar. You can see how its movements slow down and round themselves in response to comfort food. And you can make the fly giddily happy if you give it alcohol. I am not inventing: The experiment has been carried out on a fly species known as *Drosophila Melanogaster.*[8] After exposure to ethanol vapor the flies are as uncoordinated as we would be, given a comparable dose. They walk with the abandon of contented inebriation, and fall down an experimental tube like drunks staggering to a lamppost. Flies have emotions, although I am not suggesting that they *feel* emotions, let alone that they would reflect on such feelings. And if anyone is skeptical about the sophistication of the life-regulation mechanisms in such small creatures, consider the sleep mechanisms of the fly described by Ralph Greenspan and his colleagues.[9] Tiny *Drosophila* has the equivalent of our day-night cycles, periods of intense activity and restorative sleep, and even the sort of response to sleep deprivation that we show when we are jet-lagged. They need more sleep, as do we.

Or think of the marine snail *Aplysia Californica*—again no spine, little brain, and much sloth. Touch it in the gill and it will fold into itself, increase its blood pressure, and jump up its heart rate. The snail produces a number of concerted reactions that, transposed to you or me, probably would be recognized as important components of the emotion fear. Emotion? Yes. Feeling? Probably not.[10]

None of these organisms produce these reactions as a result of deliberation. Nor do they *construct* the reaction either, bit by bit, with some original flair for each instance in which the reaction is displayed. The organisms react

reflexively, automatically, in a stereotypical fashion. Like the distracted shopper selecting from a ready-to-wear display, they "select" ready-to-use responses and move on. It would be incorrect to call these reactions reflexes because classical reflexes are simple responses, whereas these reactions are complex packages of responses. The multiplicity of components and the coordination of the components distinguish emotion-related reactions from reflexes. Better to say that they are collections of reflex responses, some quite elaborate and all quite well coordinated. They allow an organism to respond to certain problems with an effective solution.

The Emotions-Proper

There is a venerable tradition of classifying emotions in varied categories. Although the classifications and labels are manifestly inadequate, there is no alternative at this point given the provisional stage of our knowledge. As knowledge accrues, the labels and the classifications are likely to change. In the meantime, we must remember that the borders between categories are porous. For the time being, I find it helpful to classify the emotions-proper in three tiers: background emotions, primary emotions, and social emotions.

As the term suggests, background emotions are not especially prominent in one's behavior, although they are remarkably important. You may never have paid much attention to it, but you probably are a good reader of background emotions if you accurately detect energy or enthusiasm in someone you have just met; or if you are capable of diagnosing subtle malaise or excitement, edginess

or tranquillity, in your friends and colleagues. If you are really good, you can do the diagnostic job without a single word being uttered by your victim. You assess the contour of movements in the limbs and the entire body. How strong? How precise? How ample? How frequent? You observe facial expressions. If words do get uttered you do not just listen to the words and picture their dictionary meanings, you listen to the music in the voice, to the prosody.

Background emotions can be distinguished from moods, which refer to the sustaining of a given emotion over long periods of time, measured over many hours or days, such as when "Peter has been in a foul mood." Mood also can be applied to the frequently repeated engagement of the same emotion, such as when Jane, who is such a steady girl, "has been flying off the handle for no reason."

When I developed this notion,[11] I began seeing background emotions as the consequence of deploying certain combinations of the simpler regulatory reactions (e.g., basic homeostatic processes, pain and pleasure behaviors, and appetites), according to the nesting principle noted earlier. Background emotions are composite expressions of those regulatory actions as they unfold and intersect moment by moment in our lives. I imagine background emotions as the largely unpredictable result of several concurrent regulatory processes engaged within the vast playground that our organisms resemble. These include metabolic adjustments associated with whatever internal need is arising or has just been satisfied; and with whatever external situation is now being appraised and handled by other emotions, appetites, or intellectual calculation. The ever-changing result of this cauldron of interactions is our "state of being," good, bad, or

somewhere in-between. When asked "how we feel," we consult this "state of being" and answer accordingly.

It is appropriate to ask if there are any regulatory reactions that do *not* contribute to background emotions; or which regulatory reactions are most frequently encountered in the makeup of background emotions such as discouragement or enthusiasm; or how do temperament and state of health interact with background emotions. The simple answer is that we do not know yet; the necessary investigations have not been done.

The *primary* (or basic) *emotions* are easier to define because there is an established tradition of lumping certain prominent emotions in this group. The frequent listing includes fear, anger, disgust, surprise, sadness, and happiness —the emotions that first come to mind whenever the term emotion is invoked. There are good reasons for this centrality. These emotions are easily identifiable in human beings across several cultures and in nonhuman species as well.[12] The circumstances that cause the emotions and pattern of behaviors that define the emotions also are quite consistent across cultures and species. Not surprisingly, most of what we know about the neurobiology of emotion comes from studying the primary emotions.[13] Fear leads the way, as Alfred Hitchcock would have no doubt predicted, but notable strides are being made regarding disgust,[14] sadness and happiness.[15]

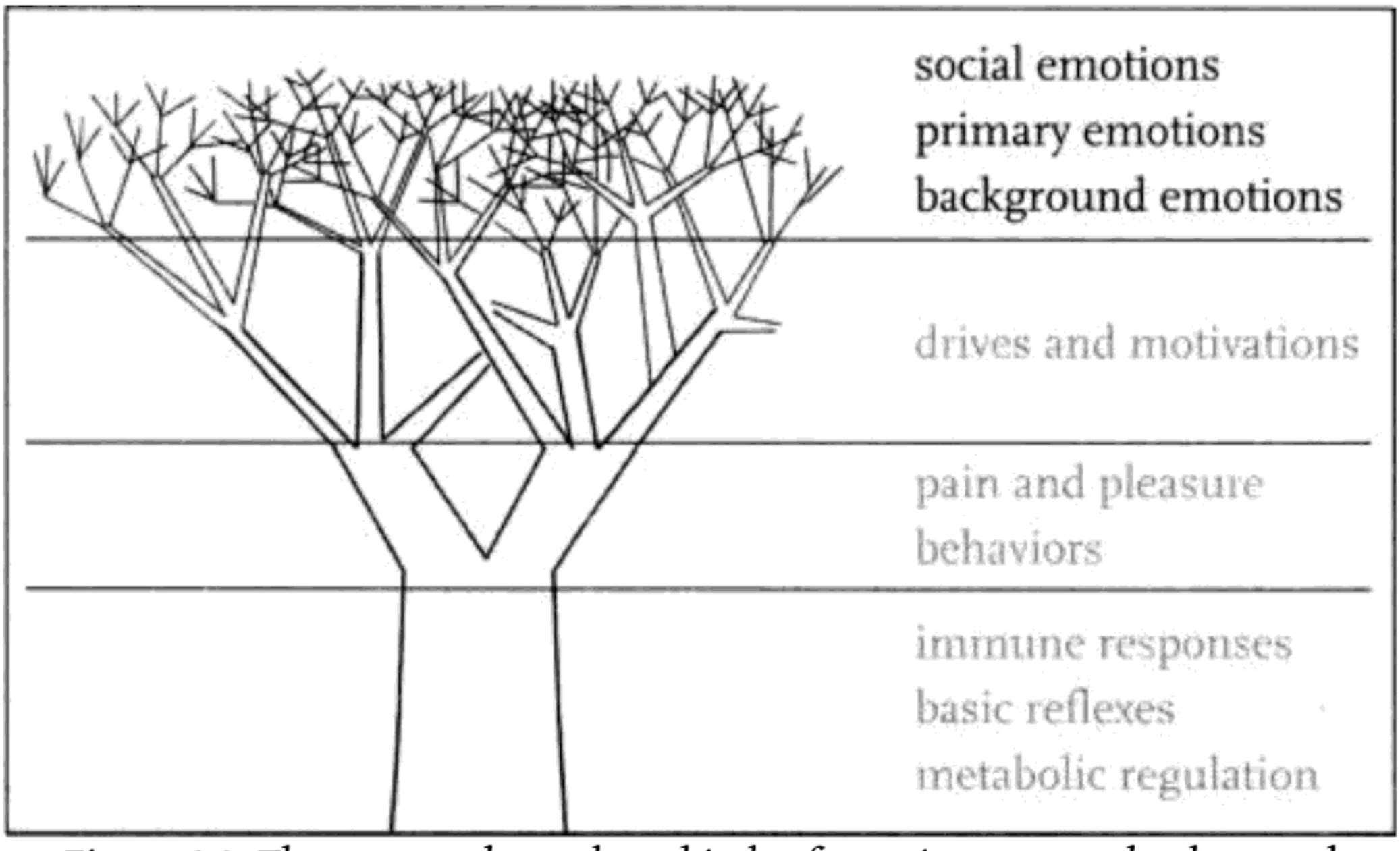


Figure 2.3: There are at least three kinds of emotion-proper: background emotions, primary emotions, and social emotions. The nesting principle applies here, too. For example, social emotions incorporate responses that are part of primary and background emotions.

The *social emotions* include sympathy, embarrassment, shame, guilt, pride, jealousy, envy, gratitude, admiration, indignation, and contempt. The nesting principle applies to social emotions as well. A whole retinue of regulatory reactions along with elements present in primary emotions can be identified as subcomponents of social emotions in varied combinations. The nested incorporation of components from lower tiers is apparent. Think of how the social emotion "contempt" borrows the facial expressions of "disgust," a primary emotion that evolved in association with the automatic and beneficial rejection of potentially toxic foods. Even the words we use to describe situations of contempt, and moral outrage—we profess to be *disgusted*—revolve around the nesting. Pain and pleasure ingredients also are evident under the surface of social emotions, albeit subtler than in the primary emotions.

We are just beginning to understand how the brain triggers and executes the social emotions. Because the term "social" inevitably conjures up the notion of human society and of culture, it is important to note that social emotions are by no means confined to humans. Look around and you will find examples of social emotions in chimpanzees, baboons, and plain monkeys; in dolphins and lions; in wolves; and, of course, in your dog and cat. The examples abound—the proud ambulations of a dominant monkey; the literally regal deportment of a dominant great ape or wolf that commands the respect of the group; the humiliated behavior of the animal that does not dominate and must yield space and precedence at mealtimes; the sympathy an elephant shows toward another that is injured and ailing; or the embarrassment the dog shows after doing what he should not.[16]

Since none of these animals is likely to have been taught to emote, it appears that the disposition to exhibit a social emotion is ingrained deep in the organism's brain, ready to be deployed when the appropriate situation manages to trigger it. There is no doubt that the general brain arrangement that permits such sophisticated behaviors in the absence of language and instruments of culture is a gift of the genome of certain species. It is part of the roster of their largely innate and automated life-regulation devices, no less so than the others we have just discussed.

Does this mean these emotions are innate in the strict sense of the term and ready to be deployed immediately after birth in the same manner that metabolic regulation clearly is, after our first breath? The answer is likely to be different for different emotions. In some instances, emotional responses may be strictly innate; in others they may require minimal

help from an appropriate exposure to the environment. Robert Hinde's work on fear is perhaps a good pointer as to what may happen in the social emotions. Hinde showed that the monkey's innate fear of snakes requires an exposure not just to a snake but to the mother's expression of fear of the snake. Once is enough for the behavior to kick into gear, but without that "once" the "innate" behavior is not engaged.[17] Something of this sort applies to the social emotions. An example is the establishment of patterns of dominance and submission in very young primates during play.

It remains difficult to accept, for anyone raised on the conviction that social behaviors are the necessary products of education, that simple animal species not known for their culture can exhibit intelligent social behaviors. But they do, and once again, they do not require that much brain to dazzle us. The modest worms *C. elegans* have exactly 302 neurons and about 5,000 interneuron connections. (For the sake of comparison, humans have several billion neurons and several trillion connections.) When these sexy little beasts (they are hermaphrodites!) are up and about in an environment with enough food and without stressors, they keep to themselves and feed in isolation. But if food is scarce or if a pestilent odor is present in the environment—by which I mean a threat if you lead a worm's existence and connect with the world through your nose—the worms congregate in single regions and feed together in groups. Just in case.[18] A number of curious social concepts are foreshadowed in this necessarily embryonic and yet far-reaching behavior: safety in numbers, strength through cooperation, belt-tightening, altruism, and the original labor union. Did you ever think humans invented such behavioral solutions? Just consider the honeybee, small

and very social in its hive society. A honeybee has 95,000 neurons. Now, that's a brain.

It is highly probable that the availability of such social emotions has played a role in the development of complex cultural mechanisms of social regulation (see chapter 4). It also is apparent that some social emotional reactions are elicited in human social situations without the stimulus for the reaction being immediately apparent to the reactor and to observers. Displays of social dominance and dependence are an example—think of all the strange antics of human behavior in sports, politics, and the workplace. One of the many reasons why some people become leaders and others followers, why some command respect and others cower, has little to do with knowledge or skills and a lot to do with how certain physical traits and the manner of a given individual promote certain emotional responses in others. To observers of such responses and to the individuals exhibiting them, some of the displays appear unmotivated because they have their origin in the innate, nonconscious apparatus of social emotion and self-preservation. We should credit Darwin for leading us to the evolutionary trail of these phenomena.

These are not the only emotional reactions of mysterious origin. There is another class of reactions with a nonconscious origin shaped by learning during one's individual development. I am referring to the affinities and detestations we acquire discreetly in the course of a lifetime of perceiving and emoting in relation to people, groups, objects, activities, and places to which Freud called our attention. Curiously, these two sets of nondeliberate, nonconscious reactions—those innate and those learned—may well be interrelated in the bottomless pit of our unconscious. One is tempted to say that their possible

nonconscious interplay signals the intersection of two intellectual legacies, that of Darwin and that of Freud, two thinkers who dedicated their work to studying the diverse influences of the innate and the acquired from below stairs.[19]

From chemical homeostatic processes to emotions-proper, life-regulation phenomena, without exception, have to do, directly or indirectly, with the integrity and health of the organism. Without exception, all of these phenomena are related to adaptive adjustments in body state and eventually lead to the changes in the brain mapping of body states, which form the basis for feelings. The nesting of the simple within the complex ensures that the regulatory purpose remains present in the higher echelons of the chain. While the purpose remains constant, complexity varies. Emotions-proper are certainly more complex than reflexes; and the triggering stimuli and target of the responses varies as well. The precise situations that initiate the process and their specific aim differ.

Hunger and thirst, for example, are simple appetites. The causative object is usually internal—a diminution in the availability of something vital for survival, namely, energy from food and water. But the ensuing behaviors are aimed at the environment and involve the search for the missing something, a search that involves exploratory motion of the surroundings and sensory detection of the thing being searched. This is not that different from what happens in emotions-proper, say, fear or anger. There, too, a competent object triggers the routine of adaptive behaviors. But the competent objects for fear and anger are almost always external (even when they are conjured up from memory and imagination in our brains they tend to stand for external

objects), and are quite varied in design (many kinds of physical stimulus, evolutionarily set or associatively learned, can cause fear). The most frequent competent triggers for hunger or thirst tend to be internal (although we can become hungry or thirsty from watching one more French movie in which the characters eat and drink and are merry). Also some drives, at least in relation to non-humans, are periodic and limited to seasons and physiological cycles, e.g., sex, while emotions occur anytime and can be sustained over time.

We also discover curious interactions across classes of regulatory reactions. Emotions-proper influence appetites, and vice versa. For example, the emotion fear inhibits hunger and sexual drives, and so do sadness and disgust. On the contrary, happiness promotes both hunger and sexual drives. The satisfaction of drives—hunger, thirst, and sex for example—causes happiness; but thwarting the satisfaction of those drives can cause anger, or despair, or sadness. Also, as noted earlier, the composite of the daily unfolding of adaptive reactions, e.g., homeostatic adjustments and drives, constitutes the ongoing background emotions and helps define mood over extended periods of time. Nonetheless, when you consider these different levels of regulatory reaction at some distance, one is struck by their remarkable formal similarity.[20]

To the best of our knowledge, most of the living creatures equipped to emote for the sake of their lives have no more brain equipment to feel those emotions than they do to think of having such emotions in the first place. They detect the presence of certain stimuli in the environment and respond to them with an emotion. All they require is a simple perceptual apparatus—a filter to detect the emotionally competent stimulus and the capacity to emote. Most living

creatures act. They probably do not feel like we do, let alone think like we do. This is a presumption, of course, but it is justified by our idea of what it takes to feel as explained in the next chapter. The simpler creatures lack the brain structures necessary to portray in the form of sensory maps the transformations that occur in the body when emotive reactions take place and that result in feeling. They also lack the brain necessary to represent the anticipated simulation of such body transformations, which would constitute the basis for desire or anxiety.

It is apparent that the regulatory reactions discussed above are advantageous to the organism that exhibits them, and that the causes of those reactions—the objects or situations that trigger them—could be judged "good" or "bad" depending on their impact on survival or well-being. But it should be apparent that the paramecium or the fly or the squirrel do not know the good or evil qualities of these situations let alone consider acting for the "good" and against the "bad." Nor are we humans striving for goodness when we balance the pH in our internal milieu or react with happiness or fear to certain objects around us. Our organisms gravitate toward a "good" result of their own accord, sometimes directly as in a response of happiness, sometimes indirectly as in a response of fear that begins by avoiding "evil" and then results in "good." I am suggesting, and I will return to this point in chapter 4, that organisms can produce advantageous reactions that lead to good results without *deciding* to produce those reactions, and even without *feeling* the unfolding of those reactions. And it is apparent from the makeup of those reactions that, as they take place, the organism moves for a certain period toward states of greater or lesser physiological balance.

I offer qualified congratulations to us humans for two reasons. First, in comparable circumstances, these automated reactions certainly create conditions in the human organism that, once mapped in the nervous system, can be represented as pleasurable or painful and eventually known as feelings. Let us say that this is the real source of human glory and human tragedy. Now for the second reason. We humans, conscious of the relation between certain objectives and certain emotions, can *willfully* strive to control our emotions, to some extent at least. We can decide which objects and situations we allow in our environment and on which objects and situations we lavish time and attention. We can, for example, decide not to watch commercial television, and advocate its eternal banishment from the households of intelligent citizens. By controlling our interaction with objects that cause emotions we are in effect exerting some control over the life process and leading the organism into greater or lesser harmony, as Spinoza would wish. We are in effect overriding the tyrannical automaticity and mindlessness of the emotional machinery. Curiously, humans long ago discovered this possibility without quite knowing the physiological basis for the strategies they use. This is what we do when we make choices regarding what we read or whom we befriend. This is what humans have done for centuries when they follow social and religious percepts that in effect modify the environment and our relation with it. This is what we try to do when we flirt with all the healthy living programs that make us exercise and diet.

It is not accurate to say that regulatory reactions including the emotions-proper are fatally and inevitably stereotyped. Some "low branch" reactions are and should be stereotyped—one does not want to interfere with nature's wisdom when it

comes to regulating cardiac function or running away from danger. But the "high branch" reactions can be modified to some extent. We can control our exposure to the stimuli that bring on the reactions. We can learn over a lifetime to engage modulating "brakes" on those reactions. We can simply use sheer willpower and just say no. Sometimes.

A Hypothesis in the Form of a Definition

Taking the varied kinds of emotion in consideration, I can now offer a working hypothesis of emotion-proper in the form of a definition.

1. An emotion-proper, such as happiness, sadness, embarrassment, or sympathy, is a complex collection of chemical and neural responses forming a distinctive pattern.
2. The responses are produced by the normal brain when it detects an emotionally competent stimulus (an ECS), the object or event whose presence, actual or in mental recall, triggers the emotion. The responses are automatic.
3. The brain is prepared by evolution to respond to certain ECSs with specific repertoires of action. However, the list of ECSs is not confined to those prescribed by evolution. It includes many others learned in a lifetime of experience.
4. The immediate result of these responses is a temporary change in the state of the body proper, and in the state of the brain structures that map the body and support thinking.

> 5. The ultimate result of the responses, directly or indirectly, is the placement of the organism in circumstances conducive to survival and well-being.[21]

The classic components of an emotional reaction are encompassed by this definition, although the separation of the phases of the process and the weight accorded to those phases may appear unconventional. The process begins with an appraisal-evaluation phase, starting with the detection of an emotionally competent stimulus. My inquiry is focused on what happens after the stimulus is detected in the mind's process—the tail end of the appraisal phase. For obvious reasons, I also leave feelings, the next phase of the emotion-to-feeling cycle, out of the definition of emotion itself.

It might be argued, for the sake of functional purity, that the appraisal phase should be left out as well—appraisal being the process leading to emotion rather than emotion itself. But the radical excision of the appraisal phase would obscure rather than illuminate the real value of emotions: their largely intelligent connection between the emotionally competent stimulus and the set of reactions that can alter our body function and our thinking so profoundly. Leaving out appraisal also would render the biological description of the phenomena of emotion vulnerable to the caricature that emotions without an appraisal phase are meaningless events. It would be more difficult to see how beautiful and amazingly intelligent emotions can be, and how powerfully they can solve problems for us.[22]

The Brain Machinery of Emotion

Emotions provide a natural means for the brain and mind to evaluate the environment within and around the organism, and respond accordingly and adaptively. Indeed, in many circumstances, we actually evaluate consciously the objects that cause emotions, in the proper sense of the term "evaluate." We process not only the presence of an object but its relation to others and its connection to the past. In those circumstances the apparatus of emotions naturally evaluates, and the apparatus of the conscious mind thinkingly coevaluates. We even can modulate our emotional response. In effect, one of the key purposes of our educational development is to interpose a nonautomatic evaluative step between causative objects and emotional responses. We attempt by doing so to shape our natural emotional responses and bring them in line with the requirements of a given culture. All that is very true, but the point I wish to make here, however, is that in order for emotions to occur there is no *need* to analyze the causative object consciously let alone evaluate the situation in which it appears. Emotions can operate in different settings.

Even when the emotional reaction occurs without conscious knowledge of the emotionally competent stimulus the emotion signifies nonetheless the result of the organism's appraisal of the situation. Never mind that the appraisal is not made clearly known to the self. Somehow the notion of appraisal has been taken too literally to signify conscious evaluation, as if the splendid job of assessing a situation and responding to it automatically would be a minor biological achievement.

One of the main aspects of the history of human development pertains to how most objects that surround our brains become capable of triggering some form of emotion or

another, weak or strong, good or bad, and can do so consciously or unconsciously. Some of these triggers are set by evolution, but some are not, instead becoming associated by our brains with emotionally competent objects by virtue of our individual experiences. Think of the house where once, as a child, you may have had an experience of intense fear. When you visit that house today you may feel uncomfortable without any cause for the discomfort other than the fact that, long ago, you had a powerful negative emotion in those same surroundings. It may even happen that in a different but somewhat similar house you experience the same discomfort, again for no reason other than you can detect the brain's record of a comparable object and situation.

There is nothing in your brain's basic makeup prepared to respond with displeasure to houses of a certain kind. But your life experience has made your brain associate such houses with the displeasure you once had. Never mind that the cause of the displeasure had nothing to do with the house itself. Call it guilt by association. The house is an innocent bystander. You have been *conditioned* to feel uncomfortable in certain houses, perhaps even to dislike certain houses without really knowing why. Or to feel well in certain houses, by precisely the same mechanism. Many of our perfectly normal and banal likes and dislikes arise this way. But note that phobias, which are neither normal nor banal, can be acquired by the same mechanism. At any rate, by the time we are old enough to write books, few if any objects in the world are emotionally neutral. The emotional distinction among objects is a distinction of grades: Some objects evoke weak, barely perceptible emotional reactions, some objects evoke strong emotional reactions, and there is every other grade in between. We even are beginning to uncover the molecular

and cellular mechanisms necessary for emotional learning to occur.[23]

Complex organisms also learn to modulate the execution of emotions in harmony with the individual circumstances—and here the terms appraisal and evaluation are most apt. The emotional modulation devices can adjust the magnitude of emotional expression without an organism's conscious deliberation. One simple example: After being told the same amusing story for the second time you will smile or laugh quite differently depending on the social context of the moment—a diplomatic dinner, a casual hallway encounter, Thanksgiving dinner with close friends, and so on. If your parents have done a good job you will not need to *think* about the context. The adjustment is automatic. Some of the adjuster devices, however, do reflect a judgment on the part of the organism's self and may result in an attempt to modify or even suppress emotions. For a number of reasons that range from the honorable to the despicable, you may elect to conceal your disgust or mirth regarding some statement that a colleague or the person you are negotiating with just made. Conscious knowledge of the context and awareness of the future consequences of every aspect of your own behavior help you decide to suppress the natural expression of emotion. But try to avoid it as you get older. It is very energy consuming.

Emotionally competent objects can be actual or recalled from memory. We have seen how a nonconscious conditioned memory can lead to a current emotion. But memory can play the same trick out in the open. For example, the actual near-accident that frightened you years ago can be recalled from memory and cause you to be frightened anew. Whether actually present, as a freshly minted image, or as a

reconstructed image recalled from memory, the kind of effect is the same. If the stimulus is emotionally competent an emotion ensues, and only the intensity varies. Actors of every sort of schooling rely on this so-called emotional memory for their trade. In some cases they let memory overtly lead them to emote. In other cases they let memory infiltrate their performance subtly, setting themselves up to behave in a certain way. Our ever-observant Spinoza did not leave this one alone either: *A man is as much affected pleasurably or painfully by the image of a thing past or future, as by the image of a thing present [Ethics,* Part III, Proposition 28].

Triggering and Executing Emotions

The appearance of an emotion depends on a complicated chain of events. Here is how I see it. The chain begins with the appearance of the emotionally competent stimulus. The stimulus, a certain object or situation actually present or recalled from memory, comes to mind. Think of the bear you came across on your trip to Alaska (this in homage to William James who wove his discussion of fear on the sighting of one such bear). Or think of a forthcoming meeting with someone you miss.

In neural terms, images related to the emotionally competent object must be represented in one or more of the brain's sensory processing systems, such as the visual or auditory regions. Let us call this the presentation stage of the process. Regardless of how fleeting the presentation, signals related to the presence of that stimulus are made available to a number of emotion-triggering sites elsewhere in the brain. You can conceive of those sites as locks that open only if the appropriate keys fit. The emotionally competent stimuli are

the keys, of course. Note that they select a preexisting lock, rather than instruct the brain on how to create one. The emotion-triggering sites subsequently activate a number of emotion-execution sites elsewhere in the brain. The latter sites are the immediate cause of the emotional state that occurs in the body and in brain regions that support the emotion-feeling process. Eventually, the process can reverberate and amplify itself, or shrivel away and close down. In the language of neuroanatomy and neurophysiology, this process begins when neural signals of a certain configuration (that originate in visual cortices that are holding neural patterns corresponding to the fast approach of a threatening object) are relayed in parallel along several pathways to several brain structures. Some of the recipient structures, for example, the amygdala, will become active when they "detect" a certain configuration—when the key fits the lock—and initiate signals toward other brain regions, thus giving rise to a cascade of events that will *become* an emotion.

These descriptions sound a lot like that of an antigen (e.g., a virus) entering the bloodstream and leading to an immune response (consisting of a large number of antibodies capable of neutralizing the antigen). And well they should because the processes are formally similar. In the case of emotion the "antigen" is presented through the sensory system and the "antibody" is the emotional response. The "selection" is made at one of several brain sites equipped to trigger an emotion. The conditions in which the process occurs are comparable, the contour of the process is the same, and the results just as beneficial. Nature is not that inventive when it comes to successful solutions. Once it works, it tries it again and again.

If only things would work as well for Hollywood producers, sequels would always make money.

Some of the brain regions now identified as emotion-triggering sites are the amygdala, deep in the temporal lobe; a part of the frontal lobe known as the ventromedial prefrontal cortex; and yet another frontal region in the supplementary motor area and cingulate. They are not the only triggering sites, but so far they are the best understood. These "triggering" sites are responsive to both natural stimuli, the electrochemical patterns that support the images in our minds, and to very unnatural stimuli, such as an electric current applied to the brain. But the sites should not be seen as rigid, delivering the same stereotyped performance time after time, because a number of influences can modulate their activity. Again, simple images in the mind as well as direct stimulation of brain structures can do the trick.

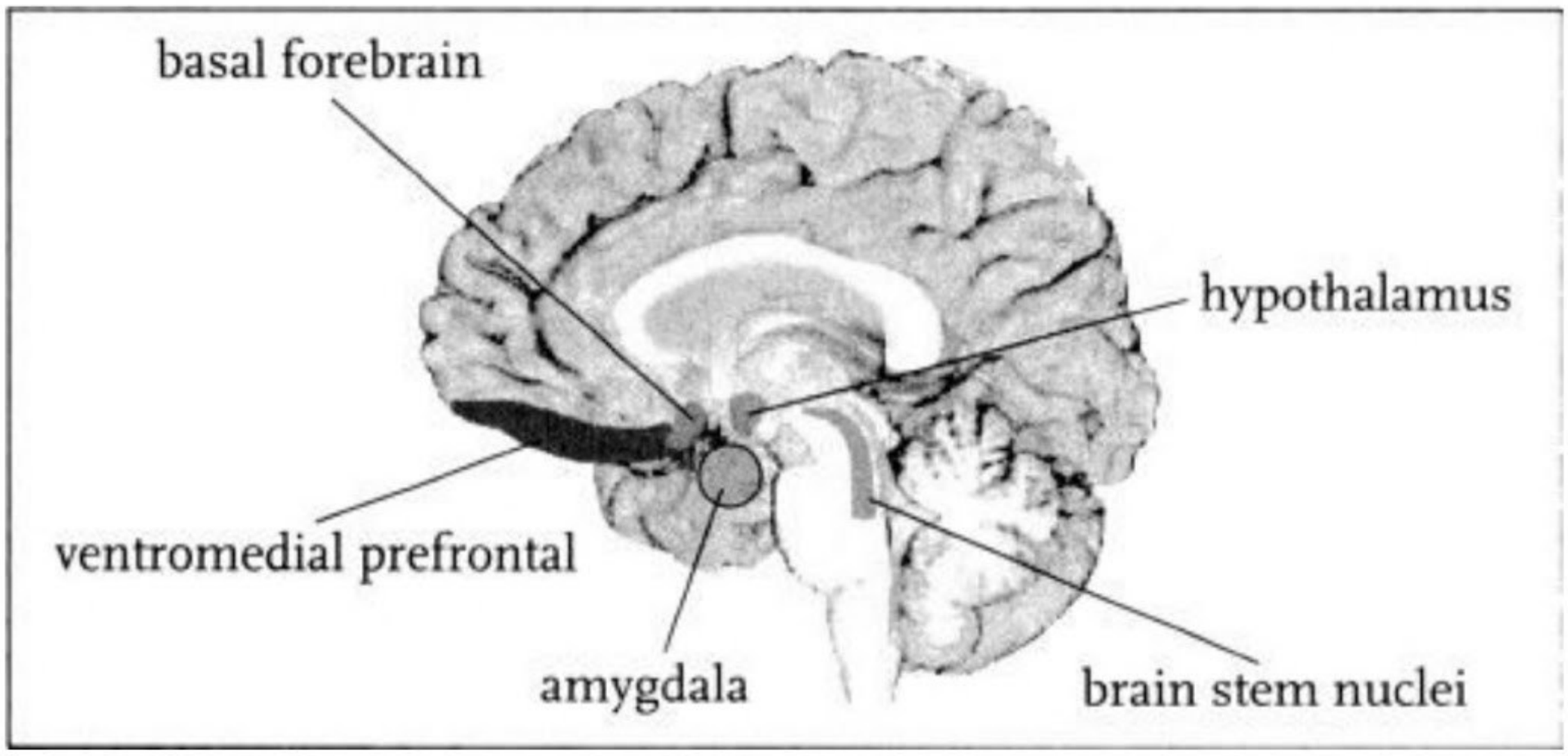


Figure 2.4: A minimalist view of the brain's triggering and execution sites for emotion. A large variety of emotions can be triggered when activity elsewhere in the brain induces activity in one of these sites, e.g., *in parts of the amygdala or the ventromedial prefrontal cortex. None of these triggering sites produces an emotion by itself. For an emotion to occur the site must cause subsequent activity in other sites,* e.g., *in basal forebrain, hypothalamus, or nuclei of the brainstem. As with any other form of complex behavior, emotion results from the concerted participation of several sites within a brain system.*

The study of the amygdala in animals has yielded important new information, most notably in the work of Joseph LeDoux, and modern brain imaging techniques have made studies of the human amygdala possible too, as exemplified by the studies of Ralph Adolphs and those of Raymond Dolan.[24] Those studies suggest that the amygdala is an important interface between visual and auditory emotionally competent stimuli and the triggering of emotions, in particular, though not exclusively, fear and anger. Neurological patients with damage to the amygdala cannot trigger those emotions and as a result do not have the corresponding feelings either. The locks for fear and anger seem to be missing, at least for visual and auditory triggers operating under regular circumstances. Recent studies also show that when recordings are made directly from single

neurons in the human amygdala, a larger proportion of neurons are tuned to unpleasant stimuli than to pleasant.[25]

Curiously, the normal amygdala serves some of its triggering functions whether we are aware of the presence of an emotionally competent stimulus. Evidence for the amygdala's ability to detect emotionally competent stimuli nonconsciously first came from the work of Paul Whalen. When he showed such stimuli very rapidly to normal people who were entirely unaware of what they were seeing, brain scans revealed that the amygdala became active.[26] Recent work from Arnie Ohman and Raymond Dolan has shown that normal subjects can learn, covertly, that a certain stimulus but not another (e.g., a particular angry face but not another angry face) is associated with an unpleasant event. The covert representation of the face associated with the bad event prompts the activation of the *right* amygdala; but the covert representation of the other face does not.[27]

Emotionally competent stimuli are detected very fast, ahead of selective attention, as shown by an impressive finding: after lesions of the occipital lobe or parietal lobe cause a blind field of vision (or a field of vision in which stimuli are *not* detected due to neglect), emotionally competent stimuli (e.g., angry or happy faces) nevertheless "break through" the barrier of blindness or neglect and are indeed detected.[28] The triggering emotional machinery captures these stimuli because they bypass the normal processing channels—channels that might have led to cognitive appraisal but simply could not do so because of blindness or neglect. The value of this "bypass" biological arrangement is apparent: whether one is paying attention, emotionally competent stimuli *can be detected.* Subsequently, attention and proper thought *can be diverted* to those stimuli.

Another important triggering site is in the frontal lobe, especially in the ventromedial prefrontal region. This region is tuned to detecting the emotional significance of more complex stimuli, for example objects and situations, natural as well as learned, competent to trigger social emotions. The sympathy evoked by witnessing someone else's accident, as well as the sadness evoked by one's personal loss, require the mediation of this region. Many of the stimuli that acquire their emotional significance in one's life experience—as in the example of the house that becomes a source of unpleasantness—trigger the respective emotions via this region.

My colleagues Antoine Bechara, Hanna Damasio, and Daniel Tranel and I have shown that damage to the frontal lobe alters the ability to emote when the emotionally competent stimulus is social in nature, and when the appropriate response is a social emotion such as embarrassment, guilt, or despair. Impairments of this sort compromise normal social behavior.[29]

In a recent series of studies from our group, Ralph Adolphs has shown that neurons in the ventromedial prefrontal regions respond rapidly and differently to the pleasant or unpleasant emotional content of pictures. Single-cell recordings from the ventromedial prefrontal region of neurological patients being assessed for the surgical treatment of seizures reveal that numerous neurons in this region, and more so in the right frontal region than in the left, respond dramatically to pictures capable of inducing *unpleasant* emotions. They begin to react as early as 120 milliseconds after the stimulus is presented. First they suspend their spontaneous firing pattern; then, after a silent interval, they fire more intensely and more frequently. Fewer

neurons respond to pictures capable of inducing *pleasant* emotions, and do so without the stop-and-go pattern noted for the unpleasantly tuned neurons.[30] The right-left brain asymmetry is more extreme than I would have predicted, but it is in keeping with a proposal made by Richard Davidson several years ago. Based on electroencephalographic studies conducted in normal individuals, Davidson suggested that the right frontal cortices were more associated with negative emotions than the left.

In order to create an emotional state, the activity in triggering sites must be propagated to execution sites by means of neural connections. The emotion-execution sites identified to date include the hypothalamus, the basal forebrain, and some nuclei in the brain stem tegmentum. The hypothalamus is the master executor of many chemical responses that are part and parcel of emotions. Directly or via the pituitary gland it releases into the bloodstream chemical molecules that alter the internal milieu, the function of viscera, and the function of the central nervous system itself. Oxytocin and vasopressin, both peptides, are examples of molecules released under the control of hypothalamic nuclei with the help of the posterior pituitary gland. A host of emotional behaviors (such as attachment and nurturing) depends on the timely availability of these hormones within the brain structures that command the execution of those behaviors. Likewise, the local brain availability of molecules such as dopamine and serotonin, which modulate neural activity, causes certain behaviors to occur. For example, the sort of behaviors experienced as rewarding and pleasurable appears to depend on the release of dopamine from one particular area (the ventrotegmental area in the brain stem),

and its availability in yet another area (the nucleus accumbens in the basal forebrain). In short, the basal forebrain and hypothalamic nuclei, some nuclei in the brain stem tegmentum, and the brain stem nuclei that control the movement of the face, tongue, pharynx, and larynx are the ultimate executors of many behaviors, simple as well as complex, that define the emotions, from courting or fleeing to laughing and crying. The complex repertoires of actions we observe are the result of the exquisite coordination of the activities of those nuclei that contribute parts of the execution in a well-concerted order and concurrence. Jaak Panksepp has dedicated a lifetime of research to this execution process.[31]

In all emotions, multiple volleys of neural and chemical responses change the internal milieu, the viscera, and the musculoskeletal system for a certain period and in a particular pattern. Facial expressions, vocalizations, body postures, and specific patterns of behavior (running, freezing, courting, or parenting) are thus enacted. The body chemistries as well as viscera such as the heart and lungs help along. Emotion is all about transition and commotion, sometimes real bodily upheaval. In a parallel set of commands the brain structures that support image-production and attention change as well; as a result, some areas of the cerebral cortex appear to be less active, while others become especially so.

In the simplest of diagrams, here is how a visually presented threatening stimulus triggers the emotion fear and leads to its execution.

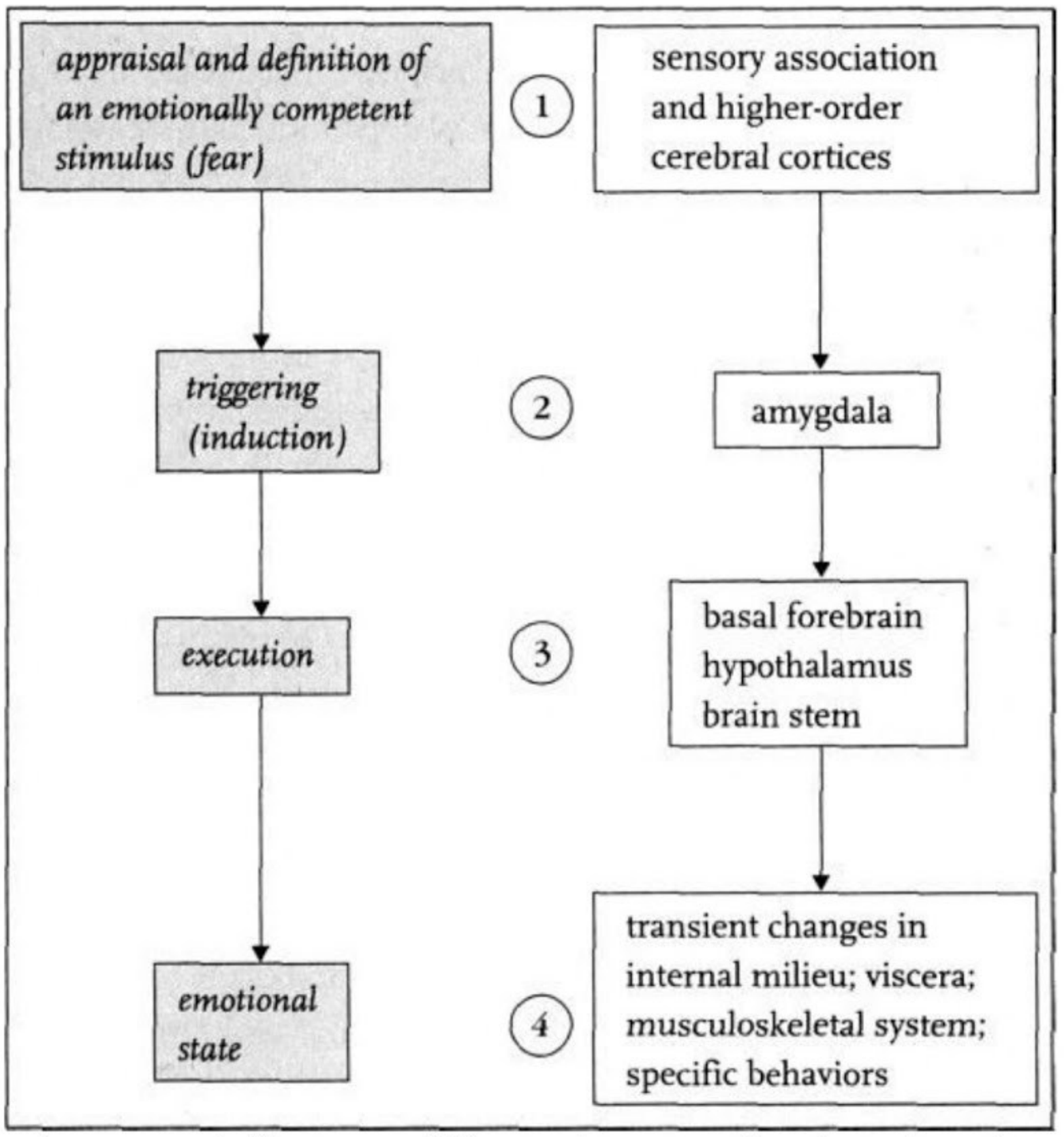


Figure 2.5: A diagram of the main stages for the triggering and execution of an emotion, using fear as an example. The shadow boxes on the left vertical column identify the stages of the process (1 to 3), from the appraisal and definition of the emotionally competent stimulus to the full-blown emotional state of fear (4). The boxes on the right vertical column identify the brain structures that are most necessary for each stage to unfold, (1 to 3) and the physiological consequences of this chain of events (4).

For the purposes of providing a manageable description of the processes of emotion and feeling, I have simplified them to fit into a single chain of events beginning with a single stimulus and terminating with the establishment of the substrates of the feeling related to the stimulus. In reality, as might be expected, the process spreads laterally into parallel chains of events and amplifies itself. This is because the presence of the initial emotionally competent stimulus often leads to the recall of other related stimuli that are also

emotionally competent. As time unfolds, the additional competent stimuli may sustain the triggering of the same emotion, trigger modifications of it, or even induce conflicting emotions. Relative to that initial stimulus, the continuation and intensity of the emotional state is thus at the mercy of the ongoing cognitive process. The contents of the mind either provide further triggers for the emotional reactions or remove those triggers, and the consequence is either the sustaining or even amplification of the emotion, or else its abatement.

The processing of emotions involves this dual track: the flowing of mental contents that bring along the triggers for the emotional responses, and the executed responses themselves, those that constitute emotions, which eventually lead to feelings. The chain that begins with the triggering of emotion and continues with the execution of emotion continues with the establishment of the substrates for feeling in the appropriate body-sensing brain regions.

Curiously, by the time the process reaches the stage of assembling feelings, we are back in the mental realm—back in the flow of thoughts where, in normal circumstances, the entire emotional detour began. Feelings are just as mental as the objects or events that trigger the emotions. What makes feelings distinctive as mental phenomena is their particular origin and content, the state of the organism's body, actual or as mapped in body-sensing brain regions.

Out of the Blue

Recently, a number of neurological studies have given us a closer look at the machinery that controls the execution of emotions. One of the most telling observations was made in a

woman undergoing treatment for Parkinson's disease. Nothing had suggested that in the course of attempting to relieve her symptoms we would be given a glimpse of how emotions come into being and of how they relate to feeling.

Parkinson's disease is a common neurological disorder that compromises the ability to move normally. Rather than causing paralysis, the condition causes rigidity of the muscles, tremors, and, perhaps most importantly, akinesia, a difficulty in initiating movements. Movements often are slow, a symptom known as bradykinesia. The disease used to be incurable, but for the past three decades it has been possible to alleviate the symptoms with the use of a medication containing levodopa, a chemical precursor of the neurotransmitter dopamine. Dopamine is missing in certain brain circuits of Parkinson's patients, much as insulin is missing in the bloodstream of patients with diabetes. (The neurons that produce dopamine in the pars compacta of the substantia nigra die away and dopamine is no longer made available at yet another brain region, the basal ganglia.) Unfortunately, the medications designed to increase dopamine in the brain circuits where it is missing do not help all patients. Also, in those that are helped, the medications may lose their effectiveness over time or cause other alterations of movement that are no less disabling than the disease. For this reason, several other modalities of treatment are being developed, one of which appears especially promising. It involves implanting tiny electrodes in the brain stem of Parkinson's patients so that the passage of a low-intensity, high-frequency electrical current can change the way in which some of the motor nuclei operate. The results usually are stunning. As the current passes, the symptoms vanish magically. The patients can move their hands with

precision and walk so normally that a stranger might not be able to tell that something had previously been wrong.

The precise placement of the array of electrode contacts is a key to the success of the treatment. To achieve this, the surgeon uses a stereotaxic device (an apparatus that permits the localization of a brain structure in three-dimensional space) and carefully navigates the electrodes into the part of the brain stem known as the mesencephalon. There are two long, vertically oriented electrodes, one for the left side of the brain stem, another for the right, and each electrode has four contacts. The contacts are located about two millimeters from each other and each contact can be independently stimulated by the passage of an electrical current. By attempting stimulation at each contact site, it is possible to determine which contact produces the greatest degree of improvement without unwanted symptoms.

The intriguing story I am about to tell you involved a patient studied by my colleague Yves Agid and his team at the Salpêtrière Hospital in Paris. The patient was a sixty-five-year-old woman with a long history of parkinsonian symptoms that no longer responded to levodopa. She had no history of depression before or after the onset of the disease, and she had not even experienced mood changes, a common side effect of levodopa. She had no history of psychiatric disorder, personally or in her family.

Once the electrodes were in place, the procedure initially went the same way it had for nineteen other patients treated by the same group. The doctors found one electrode contact that greatly relieved the woman's symptoms. But the unexpected happened when the electric current passed through one of the four contact sites on the patient's left side, precisely two millimeters below the contact that improved

her condition. The patient stopped her ongoing conversation quite abruptly, cast her eyes down and to her right side, then leaned slightly to the right and her emotional expression became one of sadness. After a few seconds she suddenly began to cry. Tears flowed and her entire demeanor was one of profound misery. Soon she was sobbing. As this display continued she began talking about how deeply sad she felt, how she had no energies left to go on living in this manner, how hopeless and exhausted she was. Asked about what was happening, her words were quite telling:

> *I'm falling down in my head, I no longer wish to live, to see anything, hear anything, feel anything . . .*
>
> *I'm fed up with life, I've had enough . . . I don't want to live anymore, I'm disgusted with life . . .*
>
> *Everything is useless . . . I feel worthless.*
>
> *I'm scared in this world.*
>
> *I want to hide in a corner . . . I'm crying over myself, of course . . . I'm hopeless, why am I bothering you?*

The physician in charge of the treatment realized that this unusual event was due to the current and aborted the procedure. About ninety seconds after the current was interrupted the patient's behavior returned to normal. The sobbing stopped as abruptly as it had begun. The sadness vanished from the patient's face. The verbal reports of sadness also terminated. Very rapidly, she smiled, appeared relaxed, and for the next five minutes was quite playful, even jocular. What was that all about? she asked. She had felt awful but did not know why. What had provoked her uncontrollable despair? She was as puzzled as the observers were.

Yet the answer to her questions was clear enough. The electrical current had not passed into the general motor control structures as intended, but had flowed instead into one of the brain stem nuclei that control particular types of action. Those actions, as an ensemble, produce the emotion sadness. This repertoire included movements of the facial musculature; movements of the mouth, pharynx, larynx, and diaphragm, which are necessary for crying and sobbing; and the varied actions that result in the production and elimination of tears.

Remarkably, it appeared as if a switch had been turned on inside the brain in response to the switch that had been turned on outside of it. This entire repertoire of actions was engaged in a well-rehearsed instrumental concert, every step in its own time and place so that the effect appeared to manifest, for all intents and purposes, the presence of thoughts capable of causing sadness—the presence of emotionally competent stimuli. Except, of course, that no such thoughts had been present prior to the unexpected incident, nor was the patient even prone to having such thoughts spontaneously. Emotion-related thoughts only came *after* the emotion began.

Hamlet may wonder at the player's capability of conjuring up emotion in spite of having no personal cause for it. "Is it not monstrous that this player here, but in a fiction, in a dream of passion, could force his soul so to his own conceit, that from her working all his visage waned, tears in his eyes, distraction in his aspect, a broken voice, and his whole form suiting with forms to his own conceit?" The player has no personal cause whatever to be emotional—he is talking about the fate of a character called Hecuba, and, as Hamlet says, "What's Hecuba to him, or he to Hecuba." However, the

player does begin by conjuring up some sad thoughts in his mind, which subsequently trigger the emotion and help him enact it with his artistry. Not so, in the strange case of this patient. There was no "conceit" prior to her emotion. There were no thoughts whatsoever to induce her behavior, no troubling ideas that came to her mind spontaneously, and no troubling ideas that she was asked to conjure up. The display of sadness, in all its spectacular complexity, came truly out of nowhere. No less importantly, sometime *after* the display of sadness was fully organized and in progress, the patient began to have a *feeling* of sadness. And, just as importantly, after she reported feeling sad she began having thoughts consonant with sadness—concern for her medical condition, fatigue, disappointment with her life, despair, and a wish to die.

The sequence of events in this patient reveals that the emotion sadness came first. The feeling of sadness followed, accompanied by thoughts of the type that usually can cause and then accompany the emotion sadness, thoughts that are characteristic of the states of mind we colloquially describe as "feeling sad." Once the stimulation ceased these manifestations waned and then vanished. The emotion disappeared and so did the feeling. The troubling thoughts were gone as well.

The importance of this rare neurological incident is apparent. In normal conditions the speed with which emotions arise and give way to feelings and related thoughts makes it difficult to analyze the proper sequence of phenomena. As thoughts normally causative of emotions appear in the mind, they cause emotions, which give rise to feelings, which conjure up other thoughts that are thematically related and likely to amplify the emotional state.

The thoughts that are conjured up may even function as independent triggers for additional emotions and thus potentiate the ongoing affective state. More emotion gives rise to more feeling, and the cycle continues until distraction or reason put an end to it. By the time all these sets of phenomena are in full swing—the thoughts that can cause emotion; the behaviors of emotion; the mind phenomena we call feelings; and the thoughts that are consequent to feelings—it is difficult to tell by introspection what came first. This woman's case helps us see through the conflation. She had no thoughts causative of sadness or any feelings of sadness prior to having an emotion called sadness. The evidence speaks both to the relative autonomy of the neural triggering mechanism of emotion and to the dependence of feeling on emotion.

At this point one should ask: Why would this patient's brain evoke the kind of thoughts that normally cause sadness considering that the emotion and feeling were unmotivated by the appropriate stimuli? The answer has to do with the dependence of feeling on emotion and the intriguing ways of one's memory. When the emotion sadness is deployed, feelings of sadness instantly follow. In short order, the brain also brings forth the *kind* of thoughts that normally cause the emotion sadness *and* feelings of sadness. This is because associative learning has linked emotions with thoughts in a rich two-way network. Certain thoughts evoke certain emotions and vice versa. Cognitive and emotional levels of processing are continuously linked in this manner. This effect can be demonstrated experimentally as shown in a study by Paul Ekman and his colleagues. He asked subjects to move certain muscles of the face in a certain sequence, such that,

unbeknownst to the subjects, the expression became one of happiness or sadness, or fear. The subjects did not know which expression was being portrayed on their faces. In their minds there was no thought capable of causing the portrayed emotion. And yet the subjects came to feel the feeling appropriate to the emotion displayed.[32] Without a doubt, parts of the emotion pattern came first. They were under the control of the experimenter and were not motivated by the subject. Some feeling followed thereafter. All of which conforms to the wisdom of Rodgers and Hammerstein. Remember that they have Anna (she who came to Siam to teach the King's children) telling her frightened self and her frightened son that whistling a happy tune will turn fear into confidence: "The results of this deception are very strange to tell. For when I fool the people I fear, I fool myself as well." Psychologically unmotivated and "acted" emotional expressions have the power to cause feeling. The expressions conjure up the feelings and the kinds of thoughts that have been learned as consonant with those emotional expressions.

From a subjective standpoint, the state of this patient after the activation of electrode "zero left" somewhat resembles those situations in which we find ourselves aware of moods and feelings, but unable to find the cause. How many times do we note, at a certain moment of a given day, that we are feeling especially well and filled with energy and hope, but don't know the reason; or, on the contrary, that we are feeling blue and edgy? In those instances, it is likely that troubling thoughts or hopeful thoughts are being processed outside of our field of consciousness. They are, nonetheless, capable of triggering the machinery of emotion and hence that of feeling. Sometimes we come to realize the origin of those affective states, sometimes we do not. For a good part

of the twentieth century, many rushed to the psychoanalyst's couch to find out more about unconscious thoughts and about the equally unconscious conflicts that were giving rise to them. These days many people just accept that there are more unknown thoughts in the heaven and earth of our minds than Hamlet's friend Horatio could ever conceive in his philosophy. And when we cannot identify the emotion-causing thought, we are visited by unexplained emotions and feelings. Fortunately those emotions and feelings are less intense and less abrupt.

The group of physicians and investigators responsible for the patient's care further investigated her unusual case.[33] Stimulation at any of the other electrode contacts implanted in this same patient caused nothing unexpected, and as noted, this reaction did not occur in any other of the nineteen patients treated the same way. On two other occasions, and with the patient's appropriate consent, the doctors established the following facts. First, when they told the patient they were stimulating the problematic electrode contact, but actually were only clicking the switch for another electrode, no behavior whatsoever ensued. They observed nothing unusual and the patient reported nothing unusual. Second, when the problematic contact was switched on again, without warning, they reproduced the same set of events as in the original, unexpected observation. Electrode placement and electrode activation clearly were linked to the appearance of the phenomenon.

The investigators also carried out a functional imaging study (using positron-emission tomography) following stimulation of contact zero left. An important finding in the latter study was the marked activation of structures in the right parietal lobe, a region involved in the mapping of the

body state and particularly of the mapping of the body in space. This activation probably related to the patient's consistent report during stimulation of marked changes in her body state, including the sensation of falling through a hole.

The scientific value of single-subject studies is always limited. The evidence usually is a starting point for new hypotheses and explorations rather than the endpoint of an investigation. Nonetheless, the evidence in this case is quite valuable. It supports the notion that the processes of emotion and feeling can be analyzed by component. It also reinforces a fundamental notion of cognitive neuroscience: Any complex mental function results from concerted contributions by *many* brain regions at varied levels of the central nervous system rather than from the work of a single brain region conceived in a phrenological manner.

The Brain Stem Switch

It is by no means clear which particular brain stem nucleus started the emotional reaction of this patient. The problematic contact appears to have been directly over the substantia nigra, but the current itself may have passed elsewhere in the vicinity. The brain stem is a very small region of the central nervous system and is jam-packed with nuclei and circuitry involved in different functions. Some of these nuclei are tiny and a minimal variation in the standard anatomy could have led to a significant rerouting of the current. But it is not in question that the event began in the mesencephalon and gradually recruited the nuclei required to produce several components of the emotion. It is even possible, judging from what has been gathered in animal

experiments, that nuclei in the region known as the periaqueductal gray (PAG) were involved in the well-coordinated production of the emotion. We know, for example, that different columns of the PAG are involved in producing different kinds of fear reaction—the kind that ends up in fight-and-flight behaviors or, instead, in freezing behaviors. The PAG may be involved in sadness reactions as well. At any rate, within one of the emotion-related mesencephalic nuclei, a chain began that, quite rapidly, engaged extensive regions of the body—face, vocal apparatus, chest cavity, not to mention the chemical systems whose activities could not be observed directly. The changes led to a specific feeling state. Moreover, as the emotion sadness and the feelings of sadness unfolded, the patient recalled thoughts consonant with sadness. Instead of beginning in the cerebral cortex, the chain of events began in a subcortical region. But the effects were similar to those that would have been produced by thinking of a tragic event or witnessing it. Anyone who would have come on the scene at that point would not have been able to tell whether this was a perfectly natural emotion-feeling state, an emotion-feeling state created by the skills of a consummate actress, or an emotion-feeling state started by an electrical switch.

Out-of-the-Blue Laughter

Lest one would think that there is something unique about crying and sadness, I must add that a phenomenon equivalent to the one we have just analyzed can be produced for laughter, as shown in a study led by Itzhak Fried.[34] The circumstances also involved a patient undergoing electrical brain stimulation. The purpose was only slightly different:

the mapping of cerebral cortex functions. In order to help patients whose epileptic seizures do not respond to medications, it is possible to surgically remove the brain region that causes the seizures. In advance of surgery, however, the surgeon not only must localize with precision the area of brain that should be removed, but also must identify brain areas that cannot be removed because of their function, such as speech-related areas. This is achieved by stimulating the brain with electricity and observing the results.

In the particular case of patient A. K. when surgeons began stimulation in a region of the left frontal lobe known as the supplementary motor area (SMA), they noted that electrical stimulation at a number of closely located sites consistently and exclusively evoked laughter. The laughter was quite genuine, so much so that the observers described it as contagious. It came entirely out of the blue—the patient was not being shown or told anything funny, and was not entertaining any thought that might lead to laughter. And yet, there it was, entirely unmotivated but realistic laughter. Remarkably, and precisely as noted in the crying patient, laughter was followed "by a sensation of merriment or mirth" in spite of its unmotivated nature. Just as interestingly, the cause of the laughter was attributed to whichever object the patient was concentrating on at the time of the stimulation. For example, if the patient was being shown a picture of a horse, she would say, "The horse is funny." On occasion the investigators themselves were deemed to be an emotionally competent stimulus as when she concluded: "You guys are just so funny . . . standing around."

The laughter-producing brain patch was small, measuring about two centimeters by two centimeters. At nearby points,