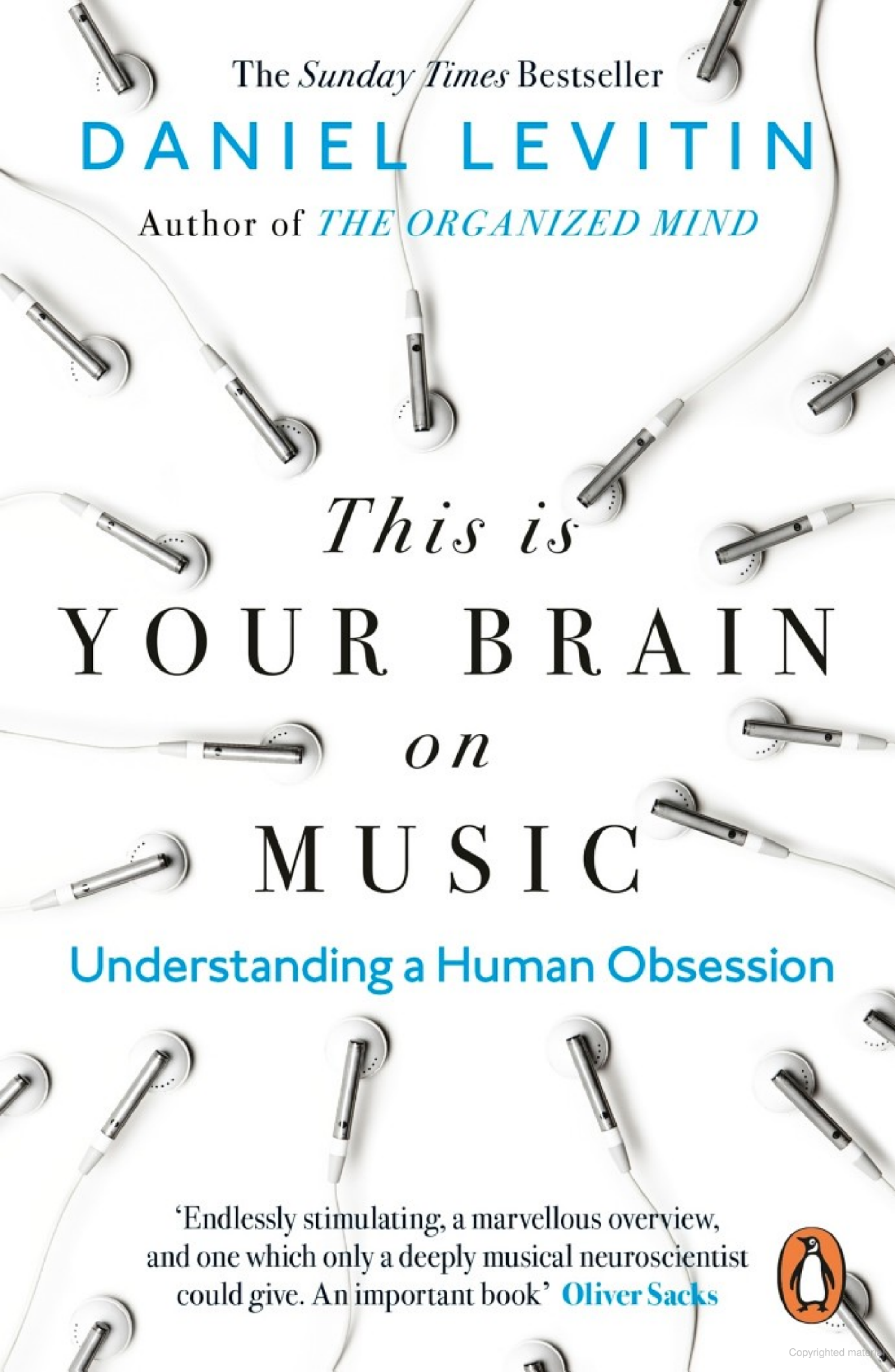


The *Sunday Times* Bestseller

DANIEL LEVITIN

Author of *THE ORGANIZED MIND*

This is YOUR BRAIN *on* MUSIC

Understanding a Human Obsession

‘Endlessly stimulating, a marvellous overview, and one which only a deeply musical neuroscientist could give. An important book’ **Oliver Sacks**

Daniel Levitin

THIS IS YOUR BRAIN ON MUSIC

Understanding a Human Obsession

Contents

About the Author

Dr. Daniel J. Levitin has a PhD in Psychology, training at Stanford University Medical School and the University of California, Berkeley. He is the author of the No. 1 bestseller *This Is Your Brain on Music* (Dutton, 2006), published in nineteen languages, and the bestsellers *The World in Six Songs* (Dutton, 2008) and *The Organized Mind* (Viking, 2014). Currently he is a James McGill Professor of Psychology, Behavioral Neuroscience and Music at McGill University in Montreal, Canada.

Introduction

I Love Music and I Love Science—Why Would I Want to Mix the Two?

> I love science, and it pains me to think that so many are terrified of the subject or feel that choosing science means you cannot also choose compassion, or the arts, or be awed by nature. Science is not meant to cure us of mystery, but to reinvent and reinvigorate it.
>
> —Robert Sapolsky, *Why Zebras Don't Get Ulcers,* p. xii

In the summer of 1969, when I was eleven, I bought a stereo system at the local hi-fi shop. It cost all of the hundred dollars I had earned weeding neighbors' gardens that spring at seventy-five cents an hour. I spent long afternoons in my room, listening to records: Cream, the Rolling Stones, Chicago, Simon and Garfunkel, Bizet, Tchaikovsky, George Shearing, and the saxophonist Boots Randolph. I didn't listen particularly loud, at least not compared to my college days when I actually set my loudspeakers on fire by cranking up the volume too high, but the noise was evidently too much for my parents. My mother is a novelist; she wrote every day in the den just down the hall and played the piano for an hour every night before dinner. My father was a businessman; he worked eighty-hour weeks, forty of those hours in his office at home on evenings and weekends. Being the businessman that he was, my father made me a proposition: He would buy me a pair of headphones if I would promise to use them when he was home. Those headphones forever changed the way I listened to music.

The new artists that I was listening to were all exploring stereo mixing for the first time. Because the speakers that came with my hundred-dollar all-in-one stereo system weren't very good, I had never before heard the depth that I could hear in the headphones—the placement of instruments both in the left-right field and in the front-back (reverberant) space. To me, records were no longer just about the songs anymore, but about the sound. Headphones opened up a world of sonic colors, a palette of nuances and details that went far beyond the chords and melody, the lyrics, or a particular singer's voice. The swampy Deep South ambience of "Green River" by Creedence, or the pastoral, open-space beauty of the Beatles' "Mother Nature's Son"; the oboes in Beethoven's Sixth (conducted by Karajan), faint and drenched in the atmosphere of a large wood-and-stone church; the sound was an enveloping experience. Headphones also made the music more personal for me; it was suddenly coming from inside my head, not out there in the world. This personal connection is ultimately what drove me to become a recording engineer and producer.

Many years later, Paul Simon told me that the sound is always what he was after too. "The way that I listen to my own records is for the sound of them; not the chords or the lyrics—my first impression is of the overall sound."

I dropped out of college after the incident with the speakers in my dorm room, and I joined a rock band. We got good enough to record at a twenty-four-track studio in California with a talented engineer, Mark Needham, who went on to record hit records by Chris Isaak, Cake, and Fleetwood Mac. Mark took a liking to me, probably because I was the only one interested in going into the control room to hear back what we sounded like, while the others were more interested in getting high in between takes. Mark treated me like a producer, although I didn't know what one was at the time, asking me what the band wanted to sound like. He taught me how much of a difference to the sound a microphone could make, or even the influence of how a microphone was placed. At first, I didn't hear some of the differences he pointed out, but he taught me what to listen for. "Notice that when I put this microphone closer to the guitar amp, the sound becomes fuller, rounder, and more even; but when I put it farther back, it picks up some of the sound of the room, giving it a more spacious sound, although you lose some of the midrange if I do that."

Our band became moderately well known in San Francisco, and our tapes played on local rock radio stations. When the band broke up—due to the guitarist's frequent suicide attempts and the vocalist's nasty habit of taking nitrous oxide and cutting himself with razor blades—I found work as a producer of other bands. I learned to hear things I had never heard before: the difference between one microphone and another, even between one brand of recording tape and another (Ampex 456 tape had a characteristic "bump" in the low-frequency range, Scotch 250 had a characteristic crispness in the high frequencies, and Agfa 467 a luster in the midrange). Once I knew what to listen for, I could tell Ampex from Scotch or Agfa tape as easily as I could tell an apple from a pear or an orange. I progressed to work with other great engineers, like Leslie Ann Jones (who had worked with Frank Sinatra and Bobby McFerrin), Fred Catero (Chicago, Janis Joplin), and Jeffrey Norman (John Fogerty, the Grateful Dead). Even though I was the producer—the person in charge of the sessions—I was intimidated by them all. Some of the engineers let me sit in on their sessions with other artists, such as Heart, Journey, Santana, Whitney Houston, and Aretha Franklin. I got a lifetime of education watching them interact with the artists, talking about subtle nuances in how a guitar part was articulated or how a vocal performance had been delivered. They would talk about syllables in a lyric, and choose among ten different performances. They could hear so well; how did they train their ears to hear things that mere mortals couldn't?

While working with small, unknown bands, I got to know the studio managers and engineers, and they steered me toward better and better work. One day an engineer didn't show up and I spliced some tape edits for Carlos Santana. Another time, the great producer Sandy Pearlman went out for lunch during a Blue Öyster Cult session and left me in charge to finish the vocals. One thing led to another, and I spent over a decade producing records in California; I was eventually lucky enough to be able to work with many well-known musicians. But I also worked with dozens of musical no-names, people who are extremely talented but never made it. I began to wonder why some musicians become household names while others languish in obscurity. I also wondered why music seemed to come so easily to some and not others. Where does creativity come from? Why do some songs move us so and others leave us cold? And what about the role of perception in all of this, the uncanny ability of great musicians and engineers to hear nuances that most of us don't?

These questions led me back to school for some answers. While still working as a record producer, I drove down to Stanford University twice a week with Sandy Pearlman to sit in on neuropsychology lectures by Karl Pribram. I found that psychology was the field that held the answers to some of my questions—questions about memory, perception, creativity, and the common instrument underlying all of these: the human brain. But instead of finding answers, I came away with more questions—as is often the case in science. Each new question opened my mind to an appreciation for the complexity of music, of the world, and of the human experience. As the philosopher Paul Churchland notes, humans have been trying to understand the world throughout most of recorded history; in just the past two hundred years, our curiosity has revealed much of what Nature had kept hidden from us: the fabric of space-time, the constitution of matter, the many forms of energy, the origins of the universe, the nature of life itself with the discovery of DNA, and the completion of the mapping of the human genome just five years ago. But one mystery has not been solved: the mystery of the human brain and how it gives rise to thoughts and feelings, hopes and desires, love, and the experience of beauty, not to mention dance, visual art, literature, and music.

What is music? Where does it come from? Why do some sequences of sounds move us so, while others—such as dogs barking or cars screeching—make many people uncomfortable? For some of us, these questions occupy a large part of our life's work. For others, the idea of picking music apart in this way seems tantamount to studying the chemical structure in a Goya canvas, at the expense of seeing the art that the painter was trying to produce. The Oxford historian Martin Kemp points out a similarity between artists and scientists. Most artists describe their work as experiments—part of a series of efforts designed to explore a common concern or to establish a viewpoint. My good friend and colleague William Forde Thompson (a music cognition scientist and composer at the University of Toronto) adds that the work of both scientists and artists involves similar stages of development: a creative and exploratory "brainstorming" stage, followed by testing and refining stages that typically involve the application of set procedures, but are often informed by additional creative problem-solving. Artists' studios and scientists' laboratories share similarities as well, with a large number of projects going at once, in various stages of incompletion. Both require specialized tools, and the results are—unlike the final plans for a suspension bridge, or the tallying of money in a bank account at the end of the business day—open to interpretation. What artists and scientists have in common is the ability to live in an open-ended state of interpretation and reinterpretation of the products of our work. The work of artists and scientists is ultimately the pursuit of truth, but members of both camps understand that truth in its very nature is contextual and changeable, dependent on point of view, and that today's truths become tomorrow's disproven hypotheses or forgotten objets d'art. One need look no further than Piaget, Freud, and Skinner to find theories that once held widespread currency and were later overturned (or at least dramatically reevaluated). In music, a number of groups were prematurely held up as of lasting importance: Cheap Trick were hailed as the new Beatles, and at one time the *Rolling Stone Encyclopedia of Rock* devoted as much space to Adam and the Ants as they did to U2. There were times when people couldn't imagine a day when most of the world would not know the names Paul Stookey, Christopher Cross, or Mary Ford. For the artist, the goal of the painting or musical composition is not to convey literal truth, but an aspect of a universal truth that if successful, will continue to move and to touch people even as contexts, societies, and cultures change. For the scientist, the goal of a theory is to convey "truth for now"—to replace an old truth, while accepting that someday this theory, too, will be replaced by a new "truth," because that is the way science advances.

Music is unusual among all human activities for both its *ubiquity* and its *antiquity*. No known human culture now or anytime in the recorded past lacked music. Some of the oldest physical artifacts found in human and protohuman excavation sites are musical instruments: bone flutes and animal skins stretched over tree stumps to make drums. Whenever humans come together for any reason, music is there: weddings, funerals, graduation from college, men marching off to war, stadium sporting events, a night on the town, prayer, a romantic dinner, mothers rocking their infants to sleep, and college students studying with music as a background. Even more so in nonindustrialized cultures than in modern Western societies, music is and was part of the fabric of everyday life. Only relatively recently in our own culture, five hundred years or so ago, did a distinction arise that cut society in two, forming separate classes of music performers and music listeners. Throughout most of the world and for most of human history, music making was as natural an activity as breathing and walking, and everyone participated. Concert halls, dedicated to the performance of music, arose only in the last several centuries.

Jim Ferguson, whom I have known since high school, is now a professor of anthropology. Jim is one of the funniest and most fiercely intelligent people I know, but he is shy—I don't know how he manages to teach his lecture courses. For his doctoral degree at Harvard, he performed fieldwork in Lesotho, a small nation completely surrounded by South Africa. There, studying and interacting with local villagers, Jim patiently earned their trust until one day he was asked to join in one of their songs. So, typically, when asked to sing with these Sotho villagers, Jim said in a soft voice, "I don't sing," and it was true: We had been in high school band together and although he was an excellent oboe player, he couldn't carry a tune in a bucket. The villagers found his objection puzzling and inexplicable. The Sotho consider singing an ordinary, everyday activity performed by everyone, young and old, men and women, not an activity reserved for a special few.

Our culture, and indeed our very language, makes a distinction between a class of expert performers—the Arthur Rubinsteins, Ella Fitzgeralds, Paul McCartneys—and the rest of us. The rest of us pay money to hear the experts entertain us. Jim knew that he wasn't much of a singer or dancer, and to him, a public display of singing and dancing implied he thought himself an expert. The villagers just stared at Jim and said, "What do you mean you don't sing?! You talk!" Jim told me later, "It was as odd to them as if I told them that I couldn't walk or dance, even though I have both my legs." Singing and dancing were a natural activity in everybody's lives, seamlessly integrated and involving everyone. The Sesotho verb for singing *(ho bina)*, as in many of the world's languages, also means to dance; there is no distinction, since it is assumed that singing involves bodily movement.

A couple of generations ago, before television, many families would sit around and play music together for entertainment. Nowadays there is a great emphasis on technique and skill, and whether a musician is "good enough" to play for others. Music making has become a somewhat reserved activity in our culture, and the rest of us listen. The music industry is one of the largest in the United States, employing hundreds of thousands of people. Album sales alone bring in $30 billion a year, and this figure doesn't even account for concert ticket sales, the thousands of bands playing Friday nights at saloons all over North America, or the thirty billion songs that were downloaded free through peer-to-peer file sharing in 2005. Americans spend more money on music than on sex or prescription drugs. Given this voracious consumption, I would say that most Americans qualify as expert music listeners. We have the cognitive capacity to detect wrong notes, to find music we enjoy, to remember hundreds of melodies, and to tap our feet in time with the music—an activity that involves a process of meter extraction so complicated that most computers cannot do it. Why do we listen to music, and why are we willing to spend so much money on music listening? Two concert tickets can easily cost as much as a week's food allowance for a family of four, and one CD costs about the same as a work shirt, eight loaves of bread, or basic phone service for a month. Understanding why we like music and what draws us to it is a window on the essence of human nature.

To ask questions about a basic, and omnipresent human ability is to implicitly ask questions about evolution. Animals evolved certain physical forms as a response to their environment, and the characteristics that conferred an advantage for mating were passed down to the next generation through the genes.

A subtle point in Darwinian theory is that living organisms—whether plants, viruses, insects, or animals—coevolved with the physical world. In other words, while all living things are changing in response to the world, the world is also changing in response to them. If one species develops a mechanism to keep away a particular predator, that predator's species is then under evolutionary pressure either to develop a means to overcome that defense or to find another food source. Natural selection is an arms race of physical morphologies changing to catch up with one another.

A relatively new scientific field, evolutionary psychology, extends the notion of evolution from the physical to the realm of the mental. My mentor when I was a student at Stanford University, the cognitive psychologist Roger Shepard, notes that not just our bodies but our minds are the product of millions of years of evolution. Our thought patterns, our predispositions to solve problems in certain ways, our sensory systems—such as the ability to see color (and the particular colors we see)—are all products of evolution. Shepard pushes the point still further: Our minds coevolved with the physical world, changing in response to ever-changing conditions. Three of Shepard's students, Leda Cosmides and John Tooby of the University of California at Santa Barbara, and Geoffrey Miller of the University of New Mexico, are among those at the forefront of this new field. Researchers in this field believe that they can learn a lot about human behavior by considering the evolution of the mind. What function did music serve humankind as we were evolving and developing? Certainly the music of fifty thousand and one hundred thousand years ago is very different from Beethoven, Van Halen, or Eminem. As our brains have evolved, so has the music we make with them, and the music that we want to hear. Did particular regions and pathways evolve in our brains specifically for making and listening to music?

Contrary to the old, simplistic notion that art and music are processed in the right hemisphere of our brains, with language and mathematics in the left, recent findings from my laboratory and those of my colleagues are showing us that music is distributed throughout the brain. Through studies of people with brain damage, we've seen patients who have lost the ability to read a newspaper but can still read music, or individuals who can play the piano but lack the motor coordination to button their own sweater. Music listening, performance, and composition engage nearly every area of the brain that we have so far identified, and involve nearly every neural subsystem. Could this fact account for claims that music listening exercises other parts of our minds; that listening to Mozart twenty minutes a day will make us smarter?

The power of music to evoke emotions is harnessed by advertising executives, filmmakers, military commanders, and mothers. Advertisers use music to make a soft drink, beer, running shoe, or car seem more hip than their competitors'. Film directors use music to tell us how to feel about scenes that otherwise might be ambiguous, or to augment our feelings at particularly dramatic moments. Think of a typical chase scene in an action film, or the music that might accompany a lone woman climbing a staircase in a dark old mansion: Music is being used to manipulate our emotions, and we tend to accept, if not outright enjoy, the power of music to make us experience these different feelings. Mothers throughout the world, and as far back in time as we can imagine, have used soft singing to soothe their babies to sleep, or to distract them from something that has made them cry.

Many people who love music profess to know nothing about it. I've found that many of my colleagues who study difficult, intricate topics such as neurochemistry or psychopharmacology feel unprepared to deal with research in the neuroscience of music. And who can blame them? Music theorists have an arcane, rarified set of terms and rules that are as obscure as some of the most esoteric domains of mathematics. To the nonmusician, the blobs of ink on a page that we call music notation might just as well be the notations of mathematical set theory. Talk of keys, cadences, modulation, and transposition can be baffling.

Yet every one of my colleagues who feels intimidated by such jargon can tell me the music that he or she likes. My friend Norman White is a world authority on the hippocampus in rats, and how they remember different places they've visited. He is a huge jazz fan, and can talk expertly about his favorite artists. He can instantly tell the difference between Duke Ellington and Count Basie by the sound of the music, and can even tell early Louis Armstrong from late. Norm doesn't have any knowledge about music in the technical sense—he can tell me that he likes a certain song, but he can't tell me what the names of the chords are. He is, however, an expert in knowing what he likes. This is not at all unusual, of course. Many of us have a practical knowledge of things we like, and can communicate our preferences without possessing the technical knowledge of the true expert. I know that I prefer the chocolate cake at one restaurant I often go to, over the chocolate cake at my neighborhood coffee shop. But only a chef would be able to analyze the cake—to decompose the taste experience into its elements—by describing the differences in the kind of flour, or the shortening, or the type of chocolate used.

It's a shame that many people are intimidated by the jargon musicians, music theorists, and cognitive scientists throw around. There is specialized vocabulary in every field of inquiry (try to make sense of a full blood-analysis report from your doctor). But in the case of music, music experts and scientists could do a better job of making their work accessible. That is something I tried to accomplish in this book. The unnatural gap that has grown between musical performance and music listening has been paralleled by a gap between those who love music (and love to talk about it) and those who are discovering new things about how it works.

A feeling my students often confide to me is that they love life and its mysteries, and they're afraid that too much education will steal away many of life's simple pleasures. Robert Sapolsky's students have probably confided much the same to him, and I myself felt the same anxiety in 1979, when I moved to Boston to attend the Berklee College of Music. What if I took a scholarly approach to studying music and, in analyzing it, stripped it of its mysteries? What if I became so knowledgeable about music that I no longer took pleasure from it?

I still take as much pleasure from music as I did from that cheap hi-fi through those headphones. The more I learned about music and about science the more fascinating they became, and the more I was able to appreciate people who are really good at them. Like science, music over the years has proved to be an adventure, never experienced exactly the same way twice. It has been a source of continual surprise and satisfaction for me. It turns out science and music aren't such a bad mix.

This book is about the science of music, from the perspective of cognitive neuroscience—the field that is at the intersection of psychology and neurology. I'll discuss some of the latest studies I and other researchers in our field have conducted on music, musical meaning, and musical pleasure. They offer new insights into profound questions. If all of us hear music differently, how can we account for pieces that seem to move so many people—Handel's *Messiah* or Don McLean's "Vincent (Starry Starry Night)" for example? On the other hand, if we all hear music in the same way, how can we account for wide differences in musical preference—why is it that one man's Mozart is another man's Madonna?

The mind has been opened up in the last few years by the exploding field of neuroscience and the new approaches in psychology due to new brain-imaging technologies, drugs able to manipulate neurotransmitters such as dopamine and serotonin, and plain old scientific pursuit. Less well known are the extraordinary advances we have been able to make in modeling how our neurons network, thanks to the continuing revolution in computer technology. We are coming to understand computational systems in our head like never before. Language now seems to be substantially hardwired into our brains. Even consciousness itself is no longer hopelessly shrouded in a mystical fog, but is rather something that emerges from observable physical systems. But no one until now has taken all this new work together and used it to elucidate what is for me the most beautiful human obsession. Your brain on music is a way to understand the deepest mysteries of human nature. That is why I wrote this book. This book was written for the general reader and not for my colleagues, so I have tried to simplify topics without *oversimplifying* them. All the research described herin has been vetted by the peer-review process and appeared in refereed journals. The full details of "your brain on music" are contained in the notes at the end of the book.

By better understanding what music is and where it comes from, we may be able to better understand our motives, fears, desires, memories, and even communication in the broadest sense. Is music listening more along the lines of eating when you're hungry, and thus satisfying an urge? Or is it more like seeing a beautiful sunset or getting a backrub, which triggers sensory pleasure systems in the brain? Why do people seem to get stuck in their musical tastes as they grow older and cease experimenting with new music? This is the story of how brains and music coevolved—what music can teach us about the brain, what the brain can teach us about music, and what both can teach us about ourselves.

1. What Is Music?

From Pitch to Timbre

What is music? To many, “music” can only mean the great masters—Beethoven, Debussy, and Mozart. To others, “music” is Busta Rhymes, Dr. Dre, and Moby. To one of my saxophone teachers at Berklee College of Music—and to legions of “traditional jazz” aficionados—anything made before 1940 or after 1960 isn’t *really* music at all. I had friends when I was a kid in the sixties who used to come over to my house to listen to the Monkees because their parents forbade them to listen to anything but classical music, and others whose parents would only let them listen to and sing religious hymns, in both cases fearing the “dangerous rhythms” of rock and roll. When Bob Dylan dared to play an electric guitar at the Newport Folk Festival in 1965, people walked out and many of those who stayed, booed. The Catholic Church banned music that contained polyphony (more than one musical part playing at a time), fearing that it would cause people to doubt the unity of God. The church also banned the musical interval of an augmented fourth, the distance between C and F-sharp and also known as a tritone (the interval in Leonard Bernstein’s *West Side Story* when Tony sings the name “Maria”). This interval was considered so dissonant that it must have been the work of Lucifer, and so the church named it *Diabolus in musica.* It was pitch that had the medieval church in an uproar. And it was timbre that got Dylan booed. It was the latent African rhythms in rock that frightened white suburban parents, perhaps fearful that the beat would induce a permanent, mind-altering trance in their innocent children. What are rhythm, pitch, and timbre—are they merely ways of describing different mechanical aspects of a song, or do they have a deeper, neurological basis? Are all of these elements necessary?

The music of avant-garde composers such as Francis Dhomont, Robert Normandeau, or Pierre Schaeffer stretches the bounds of what most of us think music is. Going beyond the use of melody and harmony, and even beyond the use of instruments, these composers use recordings of found objects in the world such as jackhammers, trains, and waterfalls. They edit the recordings, play with their pitch, and ultimately combine them into an organized collage of sound with the same type of emotional trajectory—the same tension and release—as traditional music. Composers in this tradition are like the painters who stepped outside of the boundaries of representational and realistic art—the cubists, the Dadaists, many of the modern painters from Picasso to Kandinsky to Mondrian.

What do the music of Bach, Depeche Mode, and John Cage fundamentally have in common? On the most basic level, what distinguishes Busta Rhymes's "What's It Gonna Be?!" or Beethoven's "Pathétique" Sonata from, say, the collection of sounds you'd hear standing in the middle of Times Square, or those you'd hear deep in a rainforest? As the composer Edgard Varèse famously defined it, "Music is organized sound."

This book drives at a neuropsychological perspective on how music affects our brains, our minds, our thoughts, and our spirit. But first, it is helpful to examine what music is made of. What are the fundamental building blocks of music? And how, when organized, do they give rise to music? The basic elements of any sound are loudness, pitch, contour, duration (or rhythm), tempo, timbre, spatial location, and reverberation. Our brains organize these fundamental perceptual attributes into higher-level concepts —just as a painter arranges lines into forms—and these include meter, harmony, and melody. When we listen to music, we are actually perceiving multiple attributes or "dimensions."

Before getting to the brain basis of all this, I'd like to take this chapter to define the musical terms and quickly review some basic ideas in music theory, and illustrate them with musical examples. (Musicians may want to skip or skim this chapter.) First here is a brief summary of the main terms.

~ *Pitch* is a purely psychological construct, related both to the actual frequency of a particular tone and to its relative position in the musical scale. It provides the answer to the question "What note is that?" ("It's a C-sharp.") I'll define frequency and musical scale below. (When a trumpet player blows in his instrument and makes a single sound, he makes what most of us call a *note*, and what scientists call a *tone*. The two terms, *tone* and *note* refer to the same entity in the abstract, but we reserve the word *tone* for what you hear, and the word *note* for what you see written on a musical score.) In the nursery rhymes "Mary Had a Little Lamb" and "Are You Sleeping?" pitch is the *only* thing that varies in the first seven notes—the rhythm stays the same. This demonstrates the power—and fundamentality—of pitch in defining a melody or song.

~ *Rhythm* refers to the durations of a series of notes, and to the way that they group together into units. For example, in the "Alphabet Song" (the same as "Twinkle, Twinkle Little Star") the first six notes of the song are all equal in duration as we sing the names of the letters A B C D E F and then we hold the letter G for twice the duration. Then we're back to the standard duration for H I J K, and then the following four letters are sung with half the duration, or twice as fast per letter: L M N O and then ending on a held P (leading generations of schoolchildren to spend several early months believing that there was a letter in the English alphabet called ellemmenno). In the Beach Boys' song "Barbara Ann," the first seven notes are all sung on the same pitch, with only the rhythm varying. In fact, the seven notes after that are all sung on the same pitch as well (in the melody), as Brian Wilson is joined by other voices singing other notes (harmony). The Beatles have several songs in which pitch is held constant and only rhythm varies across several notes: the first four notes of "Come Together"; the six notes of "Hard Day's Night" following the lyric "It's been a"; the first six notes of "Something."

~ *Tempo* refers to the overall speed or pace of the piece. If you were tapping your foot, dancing, or marching to the piece, it's how fast or slow these regular movements would be.

~ *Contour* describes the overall shape of a melody, taking into account only the pattern of "up" and "down" (whether a note goes up or down, not the amount by which it goes up or down).

- *Timbre* (rhymes with *amber*) distinguishes one instrument from another—say, trumpet from piano—when both are playing the same written note. It is a kind of tonal color that is produced in part by overtones from the instrument's vibrations (more on that later). It also describes the way that a single instrument can change sound as it moves across its range—say the warm sound of a trumpet low in its range versus the piercing sound of that same trumpet playing its highest note.
- *Loudness* is a purely psychological construct that relates (nonlinearly and in poorly understood ways) to how much energy an instrument creates—how much air it displaces—and what an acoustician would call the amplitude of a tone.
- *Reverberation* refers to the perception of how distant the source is from us in combination with how large a room or hall the music is in; often referred to as "echo" by laypeople, it is the quality that distinguishes the spaciousness of singing in a large concert hall from the sound of singing in your shower. It has an underappreciated role in communicating emotion and creating an overall pleasing sound.

Psychophysicists—scientists who study the ways that the brain interacts with the physical world—have shown that these attributes are *separable.* Each can be varied without altering the others, allowing the scientific study of one at a time. I can change the pitches in a song without changing the rhythm, and I can play a song on a different instrument (changing the timbre) without changing the duration or pitches of the notes. The difference between *music* and a random or disordered set of sounds has to do with the way these fundamental attributes combine, and the relations that form between them. When these basic elements combine and form relationships with one another in a meaningful way, they give rise to higher-order concepts such as meter, key, melody, and harmony.

- *Meter* is created by our brains by extracting information from rhythm and loudness cues, and refers to the way in which tones are grouped with one another across time. A waltz meter organizes tones into groups of three, a march into groups of two or four.
- *Key* has to do with a hierarchy of importance that exists between tones in a musical piece; this hierarchy does not exist in-the-world, it exists only in our minds, as a function of our experiences with a musical style and musical idioms, and mental schemas that all of us develop for understanding music.
- *Melody* is the main theme of a musical piece, the part you sing along with, the succession of tones that are most salient in your mind. The notion of melody is different across genres. In rock music, there is typically a melody for the verses and a melody for the chorus, and verses are distinguished by a change in lyrics and sometimes by a change in instrumentation. In classical music, the melody is a starting point for the composer to create variations on that theme, which may be used throughout the entire piece in different forms.
- *Harmony* has to do with relationships between the pitches of different tones, and with tonal contexts that these pitches set up that ultimately lead to expectations for what will come next in a musical piece—expectations that a skillful composer can either meet or violate for artistic and expressive purposes. Harmony can mean simply a parallel melody to the primary one (as when two singers harmonize) or it can refer to a chord progression—the clusters of notes that form a context and background on which the melody rests.

I'll be elaborating on all of these as we go along.

The idea of primitive elements combining to create art, and of the importance of relationships between elements, also exists in visual art and dance. The fundamental elements of visual perception include color (which itself can be decomposed into the three dimensions of hue, saturation, and lightness), brightness, location, texture, and shape. But a painting is more than these—it is not just a line here and another there, or a spot of red in one part of the picture and a patch of blue in another. What makes a set of lines and colors into art is the *relationship* between this line and that one; the way one color or form echoes another in a different part of the canvas. Those dabs of paint and lines become art when form and flow (the way in which your eye is drawn across the canvas) are created out of lower-level perceptual elements. When they combine harmoniously they give rise to perspective, foreground and background, and ultimately to emotion and other aesthetic attributes. Similarly, dance is not just a raging sea of unrelated bodily movements; the relationship of those movements to one another is what creates integrity and integrality, a coherence and cohesion that the higher levels of our brain process. And as in visual art, music plays on not just what notes are sounded, but which ones are not. Miles Davis famously described his improvisational technique as parallel to the way that Picasso described his use of a canvas: The most critical aspect of the work, both artists said, was not the objects themselves, but the space between objects. In Miles's case, he described the most important part of his solos as the empty space between notes, the "air" that he placed between one note and the next. Knowing precisely when to hit the next note, and allowing the listener time to anticipate it, is a hallmark of Davis's genius. This is particularly apparent in his album *Kind of Blue.*

To nonmusicians, terms such as *diatonic, cadence,* or even *key* and *pitch* can throw up an unnecessary barrier. Musicians and critics sometimes appear to live behind a veil of technical terms that can sound pretentious. How many times have you read a concert review in the newspaper and found you have no idea what the reviewer is saying? "Her sustained *appoggiatura* was flawed by an inability to complete the *roulade.*" Or, "I can't believe they modulated to C-sharp minor! How ridiculous!" What we really want to know is whether the music was performed in a way that moved the audience. Whether the singer seemed to inhabit the character she was singing about. You might want the reviewer to compare tonight's performance to that of a previous night or a different ensemble. We're usually interested in the music, not the technical devices that were used. We wouldn't stand for it if a restaurant reviewer started to speculate about the precise temperature at which the chef introduced the lemon juice in a hollandaise sauce, or if a film critic talked about the aperture of the lens that the cinematographer used; we shouldn't stand for it in music either.

Moreover, those who study music—even musicologists and scientists—disagree about what is meant by some of these terms. We employ the term *timbre,* for example, to refer to the overall sound or tonal color of an instrument—that indescribable character that distinguishes a trumpet from a clarinet when they're playing the same written note, or what distinguishes your voice from Brad Pitt's if you're saying the same words. But an inability to agree on a definition has caused the scientific community to take the unusual step of throwing up its hands and defining timbre by what it is not. (The official definition of the Acoustical Society of

America is that timbre is everything about a sound that is not loudness or pitch. So much for scientific precision!)

What is pitch and where does it come from? This simple question has generated hundreds of scientific articles and thousands of experiments. Almost all of us, even without musical training, can tell if a singer is offkey; we might not be able to say whether she is sharp or flat, or by how much, but after the age of five, most humans have as well a refined ability to detect tones that are out of tune as to discriminate a question from an accusation (in English, a rising pitch indicates a question, a straight or slightly falling pitch indicates an accusation). This comes from an interaction between our exposure to music and the physics of sound. What we call pitch is related to the frequency or rate of vibration of a string, column of air, or other physical source. If a string is vibrating so that it moves back and forth sixty times in one second, we say that it has a frequency of sixty cycles per second. The unit of measurement, cycles per second, is often called Hertz (abbreviated Hz) after Heinrich Hertz, the German theoretical physicist who was the first to transmit radio waves (a dyed-in-the-wool theoretician, when asked what practical use radio waves might have, he reportedly shrugged, "None"). If you were to try to mimic the sound of a fire engine siren, your voice would sweep through different pitches, or frequencies (as the tension in your vocal folds changes), some "low" and some "high."

Keys on the left of the piano keyboard strike longer, thicker strings that vibrate at a relatively slow rate. Keys to the right strike shorter, thinner strings that vibrate at a higher rate. The vibration of these strings displaces air molecules, and causes them to vibrate at the same rate—with the same frequency as the string. These vibrating air molecules are what reach our eardrum, and they cause our eardrum to wiggle in and out at the same frequency. The only information that our brains get about the pitch of sound comes from that wiggling in and out of our eardrum; our inner ear and our brain have to analyze the motion of the eardrum in order to figure out what vibrations out-there-in-the-world caused the eardrum to move that way. Although I said that air molecules vibrate, other molecules will too—we can hear music under water or in other fluids if the water (or other fluid) molecules are caused to vibrate. But in the vacuum of space, with no molecules to vibrate, there is no sound. (The next time you're watching *Star Trek* and hear the roar of the engines in space, you'll have some good Trekkie Trivia to share.)

By convention, when we press keys nearer to the left of the keyboard, we say that they are "low" pitch sounds, and ones near the right side of the keyboard are "high" pitch. That is, what we call "low" are those sounds that vibrate slowly, and are closer (in vibration frequency) to the sound of a large dog barking. What we call "high" are those sounds that vibrate rapidly, and are closer to what a small yip-yip dog might make. But even these terms *high* and *low* are culturally relative—the Greeks talked about sounds in the opposite way because the stringed instruments they built tended to be oriented vertically. Shorter strings or pipe organ tubes had their tops closer to the ground, so these were called the "low" notes (as in "low to the ground,") and the longer strings and tubes—reaching up toward Zeus and Apollo—were called the "high" notes. *Low* and *high*—just like *left* and *right*—are effectively arbitrary terms that ultimately have to be memorized. Some writers have argued that "high" and "low" are intuitive labels, noting that what we call high-pitched sounds come from birds (who are high up in trees or in the sky) and what we call low-pitched sounds often come from large, close-to-the-ground mammals such as bears or the low sounds of an earthquake. But this is not convincing, since low sounds also come from up high (think of thunder) and high sounds can come from down low (crickets and squirrels, leaves being crushed underfoot).

As a first definition of *pitch,* let's say it is that quality that primarily distinguishes the sound that is associated with pressing one piano key versus another.

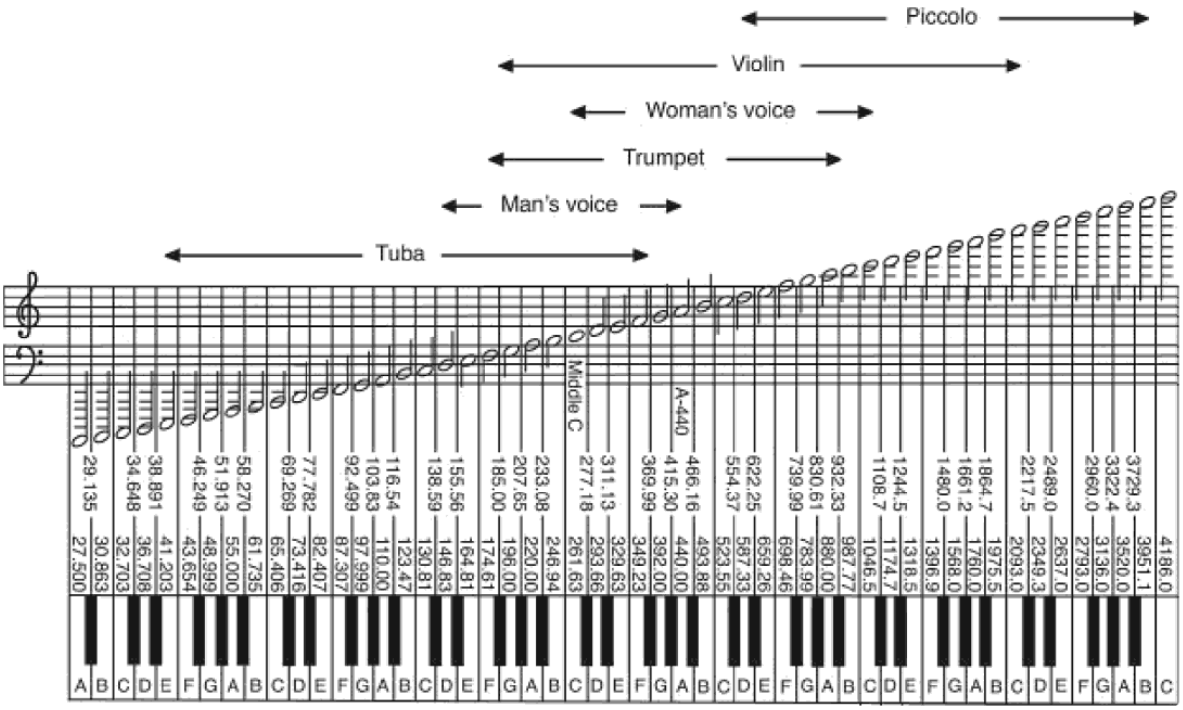
Pressing a piano key causes a hammer to strike one or more strings inside the piano. Striking a string displaces it, stretching it a bit, and its inherent resiliency causes it to return toward its original position. But it overshoots that original position, going too far in the opposite direction, and then attempts to return to its original position again, overshooting it again, and in this way it oscillates back and forth. Each oscillation covers less distance, and, in time, the string stops moving altogether. This is why the sound you hear when you press a piano key gets softer until it trails off into nothing. The distance that the string covers with each oscillation back and forth is translated by our brains into loudness; the rate at which it oscillates is translated into pitch. The farther the string travels, the louder the sound seems to us; when it is barely traveling at all, the sound seems soft. Although it might seem counterintuitive, the distance traveled and the rate of oscillation are independent. A string can vibrate very quickly and traverse either a great distance or a small one. The distance it traverses is related to how hard we hit it—this corresponds to our intuition that hitting something harder makes a louder sound. The rate at which the string vibrates is principally affected by its size and how tightly strung it is, not by how hard it was struck.

It might seem as though we should simply say that pitch is the same as frequency; that is, the frequency of vibration of air molecules. This is almost true. Mapping the physical world onto the mental world is seldom so straightforward, as we'll see later. However, for most musical sounds, pitch and frequency are closely related.

The word pitch refers to the mental representation an organism has of the fundamental frequency of a sound. That is, pitch is a purely psychological phenomenon related to the frequency of vibrating air molecules. By "psychological," I mean that it is entirely in our heads, not in the world-out-there; it is the end product of a chain of mental events that gives rise to an entirely subjective, internal mental representation or quality. Sound waves—molecules of air vibrating at various frequencies—do not themselves have pitch. Their motion and oscillations can be measured, but it takes a human (or animal) brain to map them to that internal quality we call pitch.

We perceive color in a similar way, and it was Isaac Newton who first realized this. (Newton, of course, is known as the discoverer of the theory of gravity, and the inventor, along with Leibniz, of calculus. Like Einstein, Newton was a very poor student, and his teachers often complained of his inattentiveness. Ultimately, Newton was kicked out of school.)

Newton was the first to point out that light is colorless, and that consequently color has to occur inside our brains. He wrote, "The waves themselves are not colored." Since his time, we have learned that light waves are characterized by different frequencies of oscillation, and when they impinge on the retina of an observer, they set off a chain of neurochemical events, the end product of which is an internal mental image that we call color. The essential point here is: What we perceive as color is not made up of color. Although an apple may appear red, its atoms are not themselves red. And similarly, as the philosopher Daniel Dennett points out, heat is not made up of tiny hot things.


Piccolo
Violin
Woman's voice
Trumpet
Man's voice
Tuba
Middle C
A-440
29.135
34.648
38.891
46.249
51.913
58.270
69.269
77.782
92.499
103.83
116.54
138.59
155.56
185.00
207.65
233.08
277.18
311.13
369.99
415.30
466.16
554.37
622.25
739.99
830.61
932.33
1108.7
1244.5
1480.0
1661.2
1864.7
2217.5
2489.0
2960.0
3322.4
3729.3
27.500
30.863
32.703
36.708
41.203
43.654
48.999
55.000
61.735
65.406
73.416
82.407
87.307
97.999
110.00
123.47
130.81
146.83
164.81
174.61
196.00
220.00
246.94
261.63
293.66
329.63
349.23
392.00
440.00
493.88
523.55
587.33
659.26
698.46
783.99
880.00
987.77
1046.5
1174.7
1318.5
1396.9
1568.0
1760.0
1975.5
2093.0
2349.3
2637.0
2793.0
3136.0
3520.0
3951.1
4186.0
A B C D E F G A B C D E F G A B C D E F G A B C D E F G A B C D E F G A B C D E F G A B C D E F G A B C

A bowl of pudding only has taste when I put it in my mouth—when it is in contact with my tongue. It doesn't have taste or flavor sitting in my fridge, only the potential. Similarly, the walls in my kitchen are not "white" when I leave the room. They still have paint on them, of course, but *color* only occurs when they interact with my eyes.

Sound waves impinge on the eardrums and pinnae (the fleshy parts of your ear), setting off a chain of mechanical and neurochemical events, the end product of which is an internal mental image we call pitch. If a tree falls in a forest and no one is there to hear it, does it make a sound? (The question was first posed by the Irish philosopher George Berkeley.) Simply, no—sound is a mental image created by the brain in response to vibrating molecules. Similarly, there can be no pitch without a human or animal present. A suitable measuring device can register the frequency made by the tree falling, but truly it is not pitch unless and until it is heard.

No animal can hear a pitch for every frequency that exists, just as the colors that we actually see are only a small portion of the entire electromagnetic spectrum. Sound can theoretically be heard for vibrations from just over 0 cycles per second up to 100,000 cycles per second or more, but each animal hears only a subset of the possible sounds. Humans who are not suffering from any kind of hearing loss can usually hear sounds from 20 Hz to 20,000 Hz. The pitches at the low end sound like an indistinct rumble or shaking—this is the sound we hear when a truck goes by outside the window (its engine is creating sound around 20 Hz) or when a tricked-out car with a fancy sound system has the subwoofers cranked up really loud. Some frequencies—those below 20 Hz—are inaudible to humans because the physiological properties of our ears aren't sensitive to them. The beats we hear on 50 Cents' "In da Club" or N.W.A.'s "Express Yourself" are near the low end of our range of hearing; the ending of "A Day in Life" on the CD of the Beatles' *Sgt. Pepper's Lonely Hearts Club Band* has a few seconds of sound at 15 KHz, inaudible to most adults over 40! (If the Beatles believed to never trust anyone over 40, this may have been their test, but Lennon reportedly just wanted something to make people's dogs perk up.)

The range of human hearing is generally 20 Hz to 20,000 Hz, but this doesn't mean that the range of human pitch perception is the same; although we can hear sounds in this entire range, they don't all sound musical; that is, we can't unambiguously assign a pitch to the entire range. By analogy, colors at the infrared and ultraviolet ends of the spectrum lack definition compared to the colors closer to the middle. The figure on page 23 shows the range of musical instruments, and the frequency associated with them. The sound of the average male speaking voice is around 110 Hz, and the average female speaking voice is around 220 Hz. The hum of fluorescent lights or from faulty wiring is 60 Hz (in North America; in Europe and countries with a different voltage/current standard, it can be 50 Hz). The sound that a singer hits when she causes a glass to break might be 1000 Hz. The glass breaks because it, like all physical objects, has a natural and inherent vibration frequency. You can hear this by flicking your finger against its sides or, if it's crystal, by running your wet finger around the rim of the glass in a circular motion. When the singer hits just the right frequency—the resonant frequency of the glass—it causes the molecules of the glass to vibrate at their natural rate, and they vibrate themselves apart.

A standard piano has eighty-eight keys. Very rarely, pianos can have a few extra ones at the bottom and electronic pianos, organs, and synthesizers can have as few as twelve or twenty-four keys, but these are special cases. The lowest note on a standard piano vibrates with a frequency of 27.5 Hz. Interestingly, this is about the same rate of motion that constitutes an important threshold in visual perception. A sequence of still photographs—slides—displayed at or about this rate of presentation will give the illusion of motion. “Motion pictures” are a sequence of still images presented at a rate (twenty-four frames per second) that exceeds the temporal resolving properties of the human visual system. In 35 mm film projection, each image is presented for ≈1/48th of a second, alternating with a black frame of roughly equal duration as the lens is blocked between successive still images. We perceive smooth, continuous motion when in fact there is no such thing actually being shown to us. (Old-timey movies seem to flicker because their frame rate, at 16–18 fps was too low, and our visual system picked up on the discontinuities.) When molecules vibrate at around this speed we hear something that sounds like a continuous tone. If you put playing cards in the spokes of your bicycle wheel when you were a kid, you demonstrated to yourself a related principle: At slow speeds, you simply hear the click-click-click of the card hitting the spokes. But above a certain speed, the clicks run together and create a buzz, a tone you can actually hum along with; a pitch.

When that lowest note on the piano plays, and vibrates at 27.5 Hz, to most people it lacks the distinct pitch of sounds toward the middle of the keyboard. At the lowest and the highest ends of the piano keyboard, the notes sound fuzzy to many people with respect to their pitch. Composers know this, and they either use these notes or avoid them depending on what they are trying to accomplish compositionally and emotionally. Sounds with frequencies above the highest note on the piano keyboard, around 6000 Hz and more, sound like a high-pitched whistling to most people. Above 20,000 Hz most humans don’t hear a thing, and by the age of sixty, most adults can’t hear much above 15,000 Hz or so due to a stiffening of the hair cells in the inner ear. So when we talk about the range of musical notes, or that restricted part of the piano keyboard that conveys the strongest sense of pitch, we are talking about roughly three quarters of the notes on the piano keyboard, between about 55 Hz and 2000 Hz.

Pitch is one of the primary means by which musical emotion is conveyed. Mood, excitement, calm, romance, and danger are signaled by a number of factors, but pitch is among the most decisive. A single high note can convey excitement, a single low note sadness. When notes are strung together, we get more powerful and more nuanced musical statements. Melodies are defined by the pattern or relation of successive pitches across time; most people have no trouble recognizing a melody that is played in a higher or lower key than they’ve heard it in before. In fact, many melodies do not have a “correct” starting pitch, they just float freely in space, starting anywhere. “Happy Birthday” is an example of this. One way to think about a melody, then, is as an *abstract prototype* that is derived from specific combinations of key, tempo, instrumentation, and so on. A cognitive psychologist would say that a melody is an auditory object that maintains its identity in spite of transformations, just as a chair maintains its identity when you move it to the other side of the room, turn it upside down, or paint it red. So, for example, if you hear a song played louder than you are accustomed to, you still identify it as the same song. The same holds for changes in the absolute pitch values of the song, which can be changed so long as the relative distances between them remain the same.

The notion of relative pitch values is seen readily in the way that we speak. When you ask someone a question, your voice naturally rises in intonation at the end of the sentence, signaling that you are asking. But you don’t try to make the rise in female deer, Re, a drop of golden sun ...”). As frequencies get higher, so do the letter names; B has a higher frequency than A (and hence a higher pitch) and C has a higher frequency than either A or B. After G, the note names start all over again at A. Notes with the same name have frequencies that are double (or half) the frequencies of each other. One of the several notes we call A has a frequency of 110 Hz. The note with half that frequency—55 Hz—is also an A and the note with twice 110 Hz—220 Hz—is an A as well. If we keep doubling the frequencies we get more As at 440 Hz, 880 Hz, 1760 Hz, and so on.

Here is a fundamental quality of music. Note names repeat because of a perceptual phenomenon that corresponds to the doubling and halving of frequencies. When we double or halve a frequency, we end up with a note that sounds remarkably similar to the one we started out with. This relationship, a frequency ratio of 2:1 or 1:2, is called the octave. It is so important that, in spite of the large differences that exist between musical cultures—between Indian, Balinese, European, Middle Eastern, Chinese, and so on—every culture we know of has the octave as the basis for its music, even if it has little else in common with other musical traditions. This phenomenon leads to the notion of circularity in pitch perception, and is similar to circularity in colors. Although red and violet fall at opposite ends of the continuum of visible frequencies of electromagnetic energy, we see them as perceptually similar. The same is true in music, and music is often described as having two dimensions, one that accounts for tones going up in frequency (and sounding higher and higher) and another that accounts for the perceptual sense that we’ve come back home again each time we double a tone’s frequency.

When men and women speak in unison, their voices are normally an octave apart, even if they try to speak the exact same pitches. Children generally speak an octave or two higher than adults. The first two notes of the Harold Arlen melody “Over the Rainbow” (from the movie *The Wizard of Oz*) make an octave. In “Hot Fun in the Summertime” by Sly and the Family Stone, Sly and his backup singers are singing in octaves during the first line of the verse “End of the spring and here she comes back.” As we increase frequencies by playing the successive notes on an instrument, there is a very strong perceptual sense that when we reach a doubling of frequency, we have come “home” again. The octave is so basic that even some animal species—monkeys and cats, for example—show octave equivalence, the ability to treat as similar, the way that humans do, tones separated by this amount.

An *interval* is the distance between two tones. The octave in Western music is subdivided into twelve (logarithmically) equally spaced tones. The intervallic distance between A and B (or between “do” and “re”) is called a whole step or a tone. (This latter term is confusing, since we call any musical sound a tone; I’ll use the term *whole step* to avoid ambiguity). The smallest division in our Western scale system cuts a whole step perceptually in half: the half step, or semitone, which is one twelfth of an octave. (I’ll use the word *semitone* because it is more common, and because there is no ambiguity about what it means.)

Intervals are the basis of melody, much more so than the actual pitches of notes; melody processing is relational, not absolute, meaning that we define a melody by its intervals, not the actual notes used to create them. Four semitones always create the interval known as a major third regardless of whether the first note is an A or a G# or any other note. See the table of the intervals as they’re known in our (Western) musical system.

The table could continue on: Thirteen semitones is a minor ninth, fourteen semitones is a major ninth, etc., but these names are typically used only in more advanced discussions. The intervals of the perfect fourth and perfect fifth are so called because they sound particularly pleasing to many people, and since the ancient Greeks, this particular feature of the scale is at the heart of all music. (There is no "imperfect fifth," this is just the name we give the interval.) Ignore the perfect fourth and fifth or use them in every phrase, they have been the backbone of music for at least five thousand years.

Although the areas of the brain that respond to individual pitches have been mapped, we have not yet been able to find the neurological basis for the encoding of pitch relations; we know which part of the cortex is involved in listening to the notes C and E, for example, and for F and A, but we do not know how or why both intervals are perceived as a major third, or the neural circuits that create this perceptual equivalency. These relations must be extracted by computational processes in the brain that remain poorly understood.

Distance in semitones	*Interval name*
0	unison
1	minor second
2	major second
3	minor third
4	major third
5	perfect fourth
6	augmented fourth, diminished fifth, or tritone
7	perfect fifth
8	minor sixth
9	major sixth
10	minor seventh
11	major seventh
12	octave

If there are twelve named notes within an octave, why are there only seven letters (or do-re-mi syllables)? After centuries of being forced to eat in the servants' quarters and to use the back entrance of the castle, this may just be an invention by musicians to make nonmusicians feel inadequate. The additional five notes have compound names, such as E♭ pronounced "E-flat") and F# (pronounced "F-sharp"). There is no reason for the system to be so complicated, but it is what we're stuck with.

The system is a bit clearer looking at the piano keyboard. A piano has white keys and black keys spaced out in an uneven arrangement—sometimes two white keys are adjacent, sometimes they have a black key between them. Whether the keys are white or black, the perceptual distance from one adjacent key to the next always makes a semitone, and a distance of two keys is always a whole step. This applies to many Western instruments; the distance between one fret on a guitar and the next is also a semitone, and pressing or lifting adjacent keys on woodwind instruments (such as the clarinet or oboe) typically changes the pitch by a semitone.

The white keys are named A, B, C, D, E, F, and G. The notes between—the black keys—are the ones with compound names. The note between A and B is called either A-sharp or B-flat, and in all but formal music theoretic discussions, the two terms are interchangeable. (In fact, this note could also be referred to as C double-flat, and similarly, A could be called G double-sharp, but this is an even more theoretical usage.) Sharp means high, and flat means low. B-flat is the note one semitone lower than B; A-sharp is the note one semitone higher than A. In the parallel do-re-mi system, unique syllables mark these other tones: di and ra indicate the tone between do and re, for example.

The notes with compound names are not in any way second-class musical citizens. They are just as important, and in some songs and some scales they are used exclusively. For example, the main accompaniment to “Superstition” by Stevie Wonder is played on only the black keys of the keyboard. The twelve tones taken together, plus their repeating cousins one or more octaves apart, are the basic building blocks for melody, for all the songs in our culture. Every song you know, from “Deck the Halls” to “Hotel California,” from “Ba Ba Black Sheep” to the theme from *Sex and the City,* is made up from a combination of these twelve tones and their octaves.

To add to the confusion, musicians also use the terms *sharp* and *flat* to indicate if someone is playing out of tune; if the musician plays the tone a bit too high (but not so high as to make the next note in the scale) we say that the tone being played is sharp, and if the musician plays the tone too low we say that the tone is flat. Of course, a musician can be only slightly off and nobody would notice. But when the musician is off by a relatively large amount—say one quarter to one half the distance between the note she was trying to play and the next one—most of us can usually detect this and it sounds off. This is especially apparent when there is more than one instrument playing, and the out-of-tune tone we are hearing clashes with in-tune tones being played simultaneously by other musicians.

The names of pitches are associated with particular frequency values. Our current system is called A440 because the note we call A that is in the middle of the piano keyboard has been fixed to have a frequency of 440 Hz. This is entirely arbitrary. We could fix A at any frequency, such as 439, 444, 424, or 314.159; different standards were used in the time of Mozart than today. Some people claim that the precise frequencies affect the overall sound of a musical piece and the sound of instruments. Led Zeppelin often tuned their instruments away from the modern A440 standard to give their music an uncommon sound, and perhaps to link it with the European children’s folk songs that inspired many of their compositions. Many purists insist on hearing baroque music on period instruments, both because the instruments have a different sound and because they are designed to play the music in its original tuning standard, something that purists deem important.

We can fix pitches anywhere we want because what defines music is a set of pitch relations. The specific frequencies for notes may be arbitrary, but the distance from one frequency to the next—and hence from one note to the next in our musical system—isn’t at all arbitrary. Each note in our musical system is equally spaced to our ears (but not necessarily to the ears of other species). Although there is not an equal change in cycles per second (Hz) as we climb from one note to the next, the distance between each note and the next sounds equal. How can this be? The frequency of each note in our system is approximately 6 percent more than the one before it. Our auditory system is sensitive both to relative changes and to proportional changes in sound. Thus, each increase in frequency of 6 percent gives us the impression that we have increased pitch by the same amount as we did last time.

The idea of proportional change is intuitive if you think about weights. If you're at a gym and you want to increase your weight lifting of the barbells from 5 pounds to 50 pounds, adding 5 pounds each week is not going to change the amount of weight you're lifting in an equal way. After a week of lifting 5 pounds, when you move to 10 you are doubling the weight; the next week when you move to 15 you are adding 1.5 times as much weight as you had before. An equal spacing—to give your muscles a similar increase of weight each week—would be to add a constant percentage of the previous week's weight each time you increase. For example, you might decide to add 50 percent each week, and so you would then go from 5 pounds to 7.5, then to 11.25, then to 16.83, and so on. The auditory system works the same way, and that is why our scale is based on a proportion: Every tone is 6 percent higher than the previous one, and when we increase each step by 6 percent twelve times, we end up having doubled our original frequency (the actual proportion is the twelfth root of two = 1.059463 ...).

The twelve notes in our musical system are called the chromatic scale. Any scale is simply a set of musical pitches that have been chosen to be distinguishable from each other and to be used as the basis for constructing melodies.

In Western music we rarely use all the notes of chromatic scale in composition; instead, we use a subset of seven (or less often, five) of those twelve tones. Each of these subsets is itself a scale, and the type of scale we use has a large impact on the overall sound of a melody, and its emotional qualities. The most common subset of seven tones used in Western music is called the major scale, or Ionian mode (reflecting its ancient Greek origins). Like all scales, it can start on any of the twelve notes, and what defines the major scale is the specific pattern or distance relationship between each note and its successive note. In any major scale, the pattern of intervals—pitch distances between successive keys—is: whole step, whole step, half step, whole step, whole step, whole step, half step.

Starting on C, the major scale notes are C - D - E - F - G - A - B - C, all white notes on the piano keyboard. All other major scales require one or more black notes to maintain the required whole step/half step pattern. The starting pitch is also called the tonic of the scale.

The particular placement of the two half steps in the sequence of the major is crucial; it is not only what defines the major scale and distinguishes it from other scales, but it is an important ingredient in musical expectations. Experiments have shown that young children, as well as adults, are better able to learn and memorize melodies that are drawn from scales that contain unequal distances such as this. The presence of the two half steps, and their particular positions, orient the experienced, acculturated listener to where we are in the scale. We are all experts in knowing, when we hear a B in the key of C—that is, when the tones are being drawn primary from the C major scale—that it is the seventh note (or "degree") of that scale, and that it is only a half step below the tonic, even though most of us can't name the notes, and may not even know what a tonic or a scale degree is. We have assimilated the structure of this and other scales through a lifetime of listening and passive (rather than theoretically driven) exposure to the music. This knowledge is not innate, but is gained through experience. By a similar token, we don't need to know anything about cosmology to have learned that the sun comes up every morning and goes down at night—we have learned this sequence of events through largely passive exposure.

Different patterns of whole steps and half steps give rise to alternative scales, the most common of which (in our culture) is the minor scale. There is one minor scale that, like the C major scale, uses only the white notes of the piano keyboard: the A minor scale. The pitches for that scale are A - B - C - D - E - F - G - A. (Because it uses the same set of pitches, but in a different order, A minor is said to be the "relative minor of the C major scale.") The pattern of whole steps and half steps is different from that of the major scale: whole–half–whole–whole–half–whole–whole. Notice that the placement of the half steps is very different than in the major scale; in the major scale, there is a half step just before the tonic that "leads" to the tonic, and another half step just before the fourth scale degree. In the minor scale, the half steps are before the third scale degree and before the sixth. There is still a momentum when we're in this scale to return to the tonic, but the chords that create this momentum have a clearly different sound and emotional trajectory.

Now you might well ask: If these two scales use exactly the same set of pitches, how do I know which one I'm in? If a musician is playing the white keys, how do I know if he is playing the A minor scale or the C major scale? The answer is that—entirely without our conscious awareness—our brains are keeping track of how many times particular notes are sounded, where they appear in terms of strong versus weak beats, and how long they last. A computational process in the brain makes an inference about the key we're in based on these properties. This is another example of something that most of us can do even without musical training, and without what psychologists call declarative knowledge—the ability to talk about it; but in spite of our lack of formal musical education, we know what the composer intended to establish as the tonal center, or key, of the piece, and we recognize when he brings us back home to the tonic, or when he fails to do so. The simplest way to establish a key, then, is to play the tonic of the key many times, play it loud, and play it long. And even if a composer thinks he is writing in C major, if he has the musicians play the note A over and over again, play it loud and play it long; if the composer starts the piece on an A and ends the piece on an A, and moreover, if he avoids the use of C, the audience, musicians, and music theorists are most probably going to decide that the piece is in A minor, even if this was not his intent. In musical keys as in speeding tickets, it is the observed action, not the intention, that counts.

For reasons that are largely cultural, we tend to associate major scales with happy or triumphant emotions, and minor scales with sad or defeated emotions. Some studies have suggested that the associations might be innate, but the fact that these are not culturally universal indicates that, at the very least, any innate tendency can be overcome by exposure to specific cultural associations. Western music theory recognizes three minor scales and each has a slightly different flavor. Blues music generally uses a five note (pentatonic) scale that is a subset of the minor scale, and Chinese music uses a different pentatonic scale. When Tchaikovsky wants us to think of Arab or Chinese culture in the *Nutcracker* ballet, he chooses scales that are typical to their music, and within just a few notes we are transported to the Orient. When Billie Holiday wants to make a standard tune bluesy, she invokes the blues scale and sings notes from a scale that we are not accustomed to hearing in standard classical music.

Composers know these associations and use them intentionally. Our brains know them, too, through a lifetime of exposure to musical idioms, patterns, scales, lyrics, and the associations between them. Each time we hear a musical pattern that is new instrument, you're actually hearing many, many pitches at once, not a single pitch. Most of us are not aware of this consciously, although some people can train themselves to hear this. The one with the slowest vibration rate—the one lowest in pitch—is referred to as the fundamental frequency, and the others are collectively called overtones.

To recap, it is a property of objects in the world that they generally vibrate at several different frequencies at once. Surprisingly, these other frequencies are often mathematically related to each other in a very simple way: as integer multiples of one another. So if you pluck a string and its slowest vibration frequency is one hundred times per second, the other vibration frequencies will be 2 × 100 (200 Hz), 3 × 100 Hz (300 Hz), etc. If you blow into a flute or recorder and cause vibrations at 310 Hz, additional vibrations will be occurring at twice, three times, four times, etc., this rate: 620 Hz, 930 Hz, 1240 Hz, etc. When an instrument creates energy at frequencies that are integer multiples such as this, we say that the sound is harmonic, and we refer to the pattern of energy at different frequencies as the overtone series. There is evidence that the brain responds to such harmonic sounds with synchronous neural firings—the neurons in auditory cortex responding to each of the components of the sound synchronize their firing rates with one another, creating a neural basis for the coherence of these sounds.

The brain is so attuned to the overtone series that if we encounter a sound that has all of the components except the fundamental, the brain fills it in for us in a phenomenon called *restoration of the missing fundamental.* A sound composed of energy at 100 Hz, 200 Hz, 300 Hz, 400 Hz, and 500 Hz is perceived as having a pitch of 100 Hz, its fundamental frequency. But if we artificially create a sound with energy at 200 Hz, 300 Hz, 400 Hz, and 500 Hz (leaving off the fundamental), we still perceive it as having a pitch of 100 Hz. We don't perceive it as having a pitch of 200 Hz, because our brain "knows" that a normal, harmonic sound with a pitch of 200 Hz would have an overtone series of 200 Hz, 400 Hz, 600 Hz, 800 Hz, etc. We can also fool the brain by playing sequences that deviate from the overtone series such as this: 100 Hz, 210 Hz, 302 Hz, 405 Hz, etc. In cases like these, the perceived pitch shifts away from 100 Hz in a compromise between what is presented and what a normal harmonic series would imply.

When I was in graduate school, my advisor, Mike Posner, told me about the work of a graduate student in biology, Petr Janata. Although he hadn't been raised in San Francisco like me, Petr had long bushy hair that he wore in a ponytail, played jazz and rock piano, and dressed in tiedye: a true kindred spirit. Peter placed electrodes in the inferior colliculus of the barn owl, part of its auditory system. Then, he played the owls a version of Strauss's "The Blue Danube Waltz" made up of tones from which the fundamental frequency had been removed. Petr hypothesized that if the missing fundamental is restored at early levels of auditory processing, neurons in the owl's inferior colliculus should fire at the rate of the missing fundamental. This was exactly what he found. And because the electrodes put out a small electrical signal with each firing—and because the firing *rate* is the same as a *frequency* of firing (as we saw above)—Petr sent the output of these electrodes to a small amplifier, and played back the sound of the owl's neurons through a loudspeaker. What he heard was astonishing; the melody of "The Blue Danube Waltz" sang clearly from the loudspeakers: ba da da da da, deet deet, deet deet. We were *hearing* the firing rates of the neurons and they were identical to the frequency

PENGUIN BOOKS

UK | USA | Canada | Ireland | Australia
India | New Zealand | South Africa

Penguin Books is part of the Penguin Random House group of companies whose addresses can be found at global.penguinrandomhouse.com.

Penguin
Random House
UK

First published in the United States of America by Dutton, a member of Penguin Group (USA) Inc. 2006
First published in hardback in Great Britain by Atlantic Books, an imprint of Grove Atlantic Ltd. 2007
Published in Penguin Books 2019

Cover design by Mecob. Cover image © Thomas Vogel / Getty Images

ISBN: 978-0-241-98736-0