PREFACE:
SHAPING EXPERIENCE

A FEW years ago, I was invited to speak at a popular science festival held in London. I'm a professor of cognitive philosophy (an odd title that reflects a rather eclectic set of interests spanning philosophy, neuroscience, psychology, and artificial intelligence) and I was about to give a talk on one of my favorite topics—the human brain as a "prediction machine." The festival, run by a popular science magazine, was called New Scientist Live. Every year, *New Scientist* (the magazine) invites experts in many different fields to give public presentations. This year, it was held in the huge ExCel center in London's docklands. Entering the ExCel center was like arriving at multiple conjoined ocean liners each hosting a different large-scale event. As a university professor, I'm no stranger to public speaking. But standing backstage at one of the larger auditoriums and thinking about the packed audience behind the curtain, I couldn't help but get the jitters. Maybe I should have made some last-minute changes to my slides. Maybe I ought to have worn a less startling shirt. Was there someone I forgot to thank? Suddenly, my anxious train of thought was interrupted by my phone buzzing in my pocket.

But my phone was not in my pocket. As I quickly remembered, not only had I removed it and placed it under the podium, I had also set it safely on airplane mode for the entire event. But buzzing I had felt—and clear, strong buzzing too. What I had experienced was a thoroughly modern phenomenon, a remarkably common trick of the mind now known as "phantom vibration syndrome." Given that I am a chronic long-term phone user, my brain has slowly come to expect the frequent intrusion of pocket-buzz, and I'm not the only one. A 2012 study found that 89 percent of college undergraduates reported feeling phantom phone vibrations, and it's been found to be particularly prevalent among medical interns, where fake buzzing is strongly associated with stress.* In 2013, the term was rated "word of the year" by Australia's *Macquarie Dictionary.*

It was fitting that these phantom vibrations should intrude just as I was about to launch my presentation. For although such phenomena are well known within psychology and neuroscience, they now fall into place as part of a much grander theory, one that I have been helping construct for the past decade. According to that overarching theory (the topic of my talk) phantom vibrations are just one vivid demonstration of the way all human experience is built. According to the new theory (called "predictive processing"), reality as we experience it is built from our own predictions. It was my habitual expectation of pocket-buzz that, combined with the stress of the occasion, created a clear buzzing sensation out of whole cloth.

Predictive processing speaks to one of the most challenging questions in science and philosophy—the nature of the relationship between our minds and reality. The theory, which has been steadily gaining momentum, changes our understanding of this relationship in ways that have far-reaching implica-

* All references, evidence, and supporting materials are gathered in the endnotes at the back of the book, where they are arranged by the relevant page number and a short identifying extract from the text.

tions. Contrary to the standard belief that our senses are a kind of passive window onto the world, what is emerging is a picture of an ever-active brain that is always striving to predict what the world might currently have to offer. Those predictions then structure and shape the whole of human experience, from the way we interpret a person's facial expression, to our feelings of pain, to our plans for an outing to the cinema.

Nothing we do or experience—if the theory is on track—is untouched by our own expectations. Instead, there is a constant give-and-take in which what we experience reflects not just what the world is currently telling us, but what we—consciously or nonconsciously—were expecting it to be telling us. One consequence of this is that we are never simply seeing what's "really there," stripped bare of our own anticipations or insulated from our own past experiences. Instead, all human experience is part phantom—the product of deep-set predictions. We can no more experience the world "prediction and expectation free" than we could surf without a wave.

When I stood backstage at the New Scientist Live festival, the stress of waiting to give my presentation sent my prediction machinery into overdrive. Given my lifetime of experience, I would not expect the floor to suddenly turn to jelly underneath me, or an anvil fall cartoon-like on my head. But my phone does vibrate in my pocket annoyingly often, causing my brain to form a kind of baseline prediction of frequent vibrations. Stress and caffeine (I had plenty of both) tend to amplify such effects, and signals from an anxious gut feed directly into the prediction machinery in our heads. When all those factors came together, that baseline prediction of pocket-buzzing briefly became my reality. But just as quickly as it occurred, I was able to reorient myself toward the facts, and recognize it as an illusion.

The illusion occurred because predictive brains are guessing machines, proactively anticipating signals from the body and the surrounding world. That guessing is only as good as the assumptions it makes, and even a well-informed best guess will frequently miss the mark. After all, there was no phone in my pocket. When the brain's best guessing misses the mark, the mismatch with the actual sensory signal carries crucial new information. That information (prediction error) can be used to try again—to make a better guess at how things really are. But experience still reflects the brain's current best guessing. It is just that each new round of guessing is a little bit better informed.

This challenges a once traditional picture of perception. Whereas sensory information was often considered to be the starting point of experience, the emerging science of the predictive brain suggests a rather different role. Now, the current sensory signal is used to refine and correct the process of informed guessing (the attempts at prediction) already taking place. It is now the predictions that do much of the heavy lifting. According to this new picture, experience—of the world, ourselves, and even our own bodies—is never a simple reflection of external or internal facts. Instead, all human experience arises at the meeting point of informed predictions and sensory stimulations.

This is a profound change in our understanding of the mind that fundamentally alters how we should think about perception and the construction of human reality. For much of human history, scientists and philosophers saw perception as a process that worked mostly "from the outside in," as light, sound, touch, and chemical odors activate receptors in eyes, ears, nose, and skin, progressively being refined into a richer picture of the wider world. Even well into the twenty-first century, leading models in both neuroscience and artificial intelligence retained core elements of that view.

The new science of predictive processing flips that traditional story on its head. Perception is now heavily shaped from the opposite direction, as predictions formed deep in the brain reach down to alter responses all the way down to areas closer to the skin, eyes, nose, and ears—the sensory organs that take in signals from the outside world. Incoming sensory signals help correct errors in prediction, but the predictions are in the driver's seat now. This means that what we perceive today is deeply rooted in what we experienced yesterday, and all the days before that. Every aspect of our daily experience comes to us filtered by hidden webs of prediction—the brain's best expectations rooted in our own past histories.

To see just how important this could be, imagine a world in which the weather forecast played a significant role in causing—not simply predicting—the weather itself. In that strange world, a confident forecast of rain helps bring about changes to the flows of matter and energy that determine the changing weather. There, a confident forecast of rain has causal powers that make *rain itself* a little more likely. There, as here, the weather forecast depends on a model (never perfect) of the way existing weather conditions are most likely to change and evolve over time. But in that world the weather you get (here and now) reflects a kind of combination of the effects of the prior forecast itself and preexisting conditions out in the world.

We do not live in that bizarre world. The weather we get is not affected by our best model-based predictions of that weather. But our mental world shares something of that remarkable profile. When the brain strongly predicts a certain sight, a sound, or a feeling, that prediction plays a role in shaping what we seem to see, hear, or feel.

Emotion, mood, and even planning are all based in predictions too. Depression, anxiety, and fatigue all reflect alterations to the hidden predictions that shape our experience. Alter those predictions (for example, by "reframing" a situation using different words) and our experience itself alters.

Consider the prickly rush of adrenaline I felt just before going onstage to deliver that speech. I had practiced attending to that prickly feeling while verbally reframing it not as a portent of failure but as a sign of my own chemical readiness to deliver a good performance. This helps alter my self-predictions, leading to a more relaxed and fluent performance. We'll explore several such interventions, stressing both their surprising scope and their undoubted limits.

What is your relationship to the reality you perceive? In what ways do you shape it, and, by extension, in what ways do you shape yourself, often without even knowing it? In this book, I draw on paradigm-shifting research to confront these crucial questions and ask what these insights mean for neuroscience, psychology, psychiatry, medicine, and how we live our lives. We'll look hard at experiences of the body and self, from chronic pain to psychosis, and see how work on the predictive brain helps explain a wide spectrum of human behaviors and neurodiversity. We'll reassess our own experiences of the world, from social anxiety and emotional feedback loops to the many forms of bias that can creep into our judgments. We'll also explore some ways that predictive brains might support "extended minds," blurring the boundaries between ourselves and our best-fitted tools and environments.

The book ends by putting some key insights into action, looking at ways to "hack" the predictive mind by changing our practices, reframing our experiences using different kinds of language, and the controlled use of psychedelic drugs. As these themes converge, we glimpse the shape of a new and more deeply unified science of the mind—one that does justice to the range and diversity of human experience, and that has real implications for how to think about ourselves and improve our lives.

THE

EXPERIENCE MACHINE

1

UNBOXING THE PREDICTION MACHINE

IT'S MORNING and I'm still asleep in my bedroom, a daunting pile of work perched uneasily beside the bed. Waking dozily from sleep I hear some gentle birdsong. Or at least, that's how it seems to me at first. But I soon discover that I am mistaken. I listen harder and realize that all is deathly quiet. Not even my cats' early mewling for food breaks the silence. I was hallucinating birdsong.

Fortunately, there is a simple explanation. My partner recently decided to ease the process of waking up in the morning by using a smartphone app that plays a birdsong instead of a traditional alarm. The app alarm starts off as a gentle chirping that very gradually, and very slowly, builds to something approximating a full morning chorus. Today, the alarm was not actually going off—it was far too early. Nor does the sound of real birdsong ever make it through the double-glazing. But I have become so used to waking up to the gently increasing tweeting of the alarm that my brain has started to play a trick on me. I now find that I quite often awake well in advance of the start of the actual alarm, already seeming to hear the faint onset of those prerecorded chirps.

These are genuine auditory hallucinations, caused by my new, strong expectation of waking to the subtle sound of the birds. There is probably nothing sinister about my proneness to this hallucination. It has long been known that hallucinations, both auditory and visual, can be quite easily induced by the right kind of training. But these, as well as a myriad of other intriguing phenomena, are lately falling into place as signs of something much larger—something that lies at the very heart of all human experience.

The idea (the main topic of this book) is that human brains are prediction machines. They are evolved organs that build and rebuild experiences from shifting mixtures of expectation and actual sensory evidence. According to that picture, my own unconscious predictions about what I was likely to be hearing as I awoke pulled my perceptual experience briefly in that direction, creating a short-lived hallucination that was soon corrected as more information flowed in through my senses. That new information (signifying the lack of birdsong) generated "prediction error signals" and these—on this occasion at least—were all it took to bring my experience back into line with reality. The hallucination gave way to a clear experience of a silent room. But in other cases, as we'll see, mistaken predictions can become entrenched and contact with reality (itself a complex and vexed notion) harder to achieve. Even when there are no mistakes involved, and we are seeing things "as they are," our brain's predictions are still playing a central role. Predictions and prediction errors are increasingly recognized as the core currency of the human brain, and it is in their shifting balances that all human experience takes shape.

This book is about those balances and an emerging science that turns much of what we thought we knew about perceiving our worlds upside down. According to that science, the brain is constantly trying to guess how things in the world (and our own body) are most likely to be, given what has been learned from previous encounters. Everything that I see, hear, touch, and feel—so this new science suggests—reflects hidden wells of prediction. If the expectations are sufficiently strong, or (as in early chirps of the bird alarm) the sensory evidence sufficiently subtle, I may get things wrong, in effect overwriting parts of the real sensory information with my brain's best guess at how things ought to be.

This does not mean that successful sensing is simply a form of hallucination, though the mechanisms are related to those of hallucination. We should not downplay the importance of all that rich sensory information arriving at the eyes, ears, and other senses. But it casts the process of seeing—and of perceiving more generally—in a new and different way. It casts it as a process led by our brain's own best predictions: predictions that are then checked and corrected using the sensory inputs as a guide. With the prediction machinery up and running, perception becomes a process structured not simply by incoming sensory information but by difference—the difference between the actual sensory signals and the ones the brain was expecting to encounter.

Since brains are never simply "turned on" from scratch—not even first thing in the morning when I awake—predictions and expectations are always in play, proactively structuring human experience every moment of every day. On this alternative account, the perceiving brain is never passively responding to the world. Instead, it is actively trying to hallucinate the world but checking that hallucination against the evidence coming in via the senses. In other words, the brain is constantly painting a picture, and the role of the sensory information is mostly to nudge the brushstrokes when they fail to match up with the incoming evidence.

This new understanding of the process of perceiving has real importance for our lives. It alters how we should think about the evidence of our own senses. It impacts how we should think about the way we experience our own bodily states—of pain, hunger, and other experiences such as feeling anxious or depressed. For the way our bodily states feel to us likewise reflects a complex mixture of what our brains predict and what the current bodily signals suggest. This means that we can, at times, change how we feel by changing what we (consciously or unconsciously) predict.

This does not mean we can simply "predict ourselves better," nor does it mean we can alter our own experiences of pain or hunger in any way we choose. But it does suggest some principled and perhaps unexpected wiggle room—room that, with care and training, we might turn to our advantage. Handled carefully, a better appreciation of the power of prediction could improve the way we think about our own medical symptoms and suggest new ways of understanding mental health, mental illness, and neurodiversity.

The Smart Camera Model of Seeing

The idea that the brain is basically a giant prediction machine is relatively recent. Prior to that, it was widely believed that sensory information is processed in a mostly "feedforward" manner—that is, taken from our senses and directed "forward" into the brain. To take the best-studied example, visual information (that older picture suggests) is first registered at the eyes and then processed in a step-by-step fashion deeper and deeper inside the brain, which is slowly extracting more and more abstract forms of information. Beginning with patterns of incoming light, the brain might first extract information about simple features such as lines, blobs, and edges, then assemble these into larger and more complex wholes. I'm calling this the "smart camera" account of seeing. But this was clearly no camera, but rather a very smart intelligent system. Nonetheless, as in a simple camera, the direction of influence flowed mostly inward, moving forward from the eyes into the brain. Only at some point quite late in this process would lifetime memory and world knowledge become engaged, enabling you (the perceiver) to understand how things are in your world.

Versions of the smart camera (feedforward) view have been influential in philosophy, neuroscience, and AI. Such a view is intuitive because we typically think of perception as all about the flow of information from the world to the mind. That picture can be found, for example, in Descartes's 1664 *Treatise on Man*. There, Descartes depicts perception as the complex opening and closing of networks of inner tubes imprinting an image of the world first onto the sense organs (such as the eyes) and then via a network of tiny tunnels deeper and deeper into the brain. As impressions from the outside world (and from within the body) flowed forward into the brain, they were said to be preserved in our minds much the way pushing your fingers into wax preserves information about their shape.

It was never clear how Descartes's mechanism would work. But what remained even as much more sophisticated scientific understandings emerged was the core idea of the perceiving brain as a relatively passive organ taking sensory inputs from the world and then "processing" them in a predominantly feed-forward (outer to inner) fashion. That idea was pretty much standard in late-twentieth-century cognitive neuroscience. This was probably because it appeared as a governing principle of David Marr's hugely influential computer model of vision.

Marr was a towering figure, whose work in neuroscience, computer vision, and AI ranks among the most important contributions ever made to cognitive science. In Marr's depiction, visual processing starts by detecting basic ingredients in some incoming signal—an ordered array of pixels, for instance. From there, layered processing slowly builds toward a more complex understanding. For example, the next stage might look for places where pixel intensities display rapid changes from their neighbors—usually a clue to the presence of a boundary or an edge out there in the world. As processing moves forward, step by step and deeper into the brain, further patterns are detected, such as the recurring sequences that characterize stripes. Vision is here a matter of subjecting the raw signal to a series of operations, such as edge or stripe detection, that slowly reveal more and more complex patterns in the environment—the source of the incoming signal. Eventually, the complex detected patterns are brought into contact with knowledge and memory to deliver (though revealingly, this part of the puzzle was never satisfactorily solved) a kind of 3D picture of the worldly scene.

Marr's computer model (like any computer model) had the distinct virtue of specifying the key computations that might be involved in those early stages of processing, though the shape of those crucial final steps remained something of a mystery. The Marr model was for many years the standard picture not just in artificial vision but in neuroscience too. Even into the twenty-first century the visual system was primarily regarded as a mostly feedforward analyzer of incoming sensory information along the lines that Marr had described.

Notably absent from Marr's model, however, was another direction of influence—one running backward, from deep within the brain down toward the eyes and other sensory organs. The number of neuronal connections carrying signals backward in this way is estimated to exceed the number of connections carrying signals forward by a very substantial margin, in some places by as much as four to one. What is all that downward connectivity feeding information from deep in the brain to regions closer to the sensory peripheries doing? This wiring runs in the opposite direction to the wiring needed to perform the processing tasks described in Marr's early computational model, yet it reaches right down to those very regions.

Real neural wiring like that is costly to install and maintain. The brain, weighing in at about 2 percent of human body weight, is estimated to account for around 20 percent of total bodily energy consumption. It is by far our most "expensive" adaptive accessory. Yet a huge amount of that expense is now known to be devoted to establishing and maintaining an immense web of downward (and sideways) connectivity, spanning not just early visual processing but the whole of the brain. This is a puzzle. It was puzzling enough to lead the artificial intelligence pioneer Patrick Winston to comment, even as recently as 2012, that with so much information apparently flowing in the other (downward) direction, we confront "a strange architecture about which we are nearly clueless." Things look different, however, once we recognize the attractions of a bold new claim: that brains are nothing other than large-scale prediction machines.

Flipping the Flow

It now seems that the core operating principle of the perceiving brain is pretty much the opposite of the smart camera view. Instead of constantly expending large amounts of energy on processing incoming sensory signals, the bulk of what the brain does is learn and maintain a kind of model of body and world—a model that can then be used, moment by moment, to try to predict the sensory signal. These predictions help structure everything we see, hear, touch, and feel. They were at work when I heard nonexistent birdsong in the morning. They were at work when I felt phantom vibrations from a smartphone that was not even in my pocket. But they are also at work, as we'll see, when I hear actual birdsong, feel real phone vibrations, and see the various objects scattered about on my university desktop.

A predictive brain is a kind of constantly running simulation of the world around us—or at least, the world as it matters to us. Incoming sensory information is used to keep the model honest—by comparing the prediction to the sensory evidence and generating an error signal when the two don't match up. Despite the wiring costs, constant prediction brings many efficiencies, as we'll shortly see. It also—and perhaps more importantly—makes us flexible, able to adapt our responses in ways that reflect the demands of our current tasks and context. Instead of steadily extracting a rich picture of the world from a barrage of sensory clues, the rich evolving picture of the world is the starting point, and the sensory information is used to test, probe, and tweak that picture. Before new sensory signals arrive, the predictive brain is already busy painting a rich picture of how things are most likely to be.

This explains, in broad outline, the need for all that downward connectivity. It is carrying predictions from deep in the brain, pushing them toward the sensory peripheries. It also explains the huge energy outlay used simply to sustain the brain's intrinsic activity. That activity is necessary to maintain the model that issues moment-by-moment predictions. As a brain encounters new sensory information its job is to determine if there is anything in that incoming signal that looks like important "news"—unpredicted sensory information that matters to whatever it is that we are trying to see or do. There is increasing consensus that something like this is the primary way our brains process sensory information. Unpacking that hypothesis, the last ten to fifteen years has seen an explosion of work in computational and cognitive neuroscience that now makes detailed and testable sense of this, thereby solving the mystery of Winston's "strange architecture." That work goes by various names including "predictive processing," "hierarchical predictive coding," and "active inference." I'll mostly stick with "predictive processing" as a handy label for this family of theories.

According to this view, the smart camera picture of perception was a big mistake. Despite its intuitive appeal, the right way to think about perception is not (for the most part) as a process that runs primarily from the eyes and other sense organs inward. Nor is the brain ever just sitting there waiting patiently for sensory information to arrive. Instead, it is actively anticipating the sensory information, using everything it knows about patterns and objects in the world—the twittering sounds of birds (and of my partner's early morning alarm), the all-too-frequent intrusion of phone vibrations, and the organization of the various objects on my office desk. It is also making constant use of the active body, moving head, eyes, and limbs in ways that harvest new and better information. Instead of being a passive receiver and processor of sensory information, a brain like that is a tireless predictor (and, as we'll later see, a skilled and active interrogator) of its own sensory streams.

Bad Radios and Controlled Hallucinations

The contemporary picture of the predictive brain has historical roots in the nineteenth-century ideas of a German physicist and polymath named Hermann von Helmholtz. Helmholtz was the inventor of the ophthalmoscope used by opticians to examine the eye and formulated the law of conservation of energy. He was also interested in theories of perception and argued that we perceive the world only thanks to a kind of unconscious reasoning or inference in which the brain is asking itself, "Given everything I know, how must the world be for me to be receiving the pattern of signals currently present?" This is the question that perceptual systems are built to resolve.

You might not realize how common this is in our everyday life. If you listen to a familiar song on a radio with bad reception, the words and rhythms sound surprisingly clear. But try to listen to a brand-new song with that same reception quality and the sounds seem much more indistinct, the vocals hard to distinguish. In each case your brain, just as Helmholtz argued, is using what it knows to try to infer which words and sounds are the most likely cause of the somewhat patchy auditory signals currently being picked up by your ears. But the brain's guessing is much better for the familiar song—making it sound that much clearer. In fact, that guessing is altering the brain's responses all the way "down" to early auditory processing areas, so as to bring those responses more into line with the expected sounds. In a very real sense, your brain is now playing much of the song for itself, so the poor incoming signal is cleaned up using stored knowledge about the world.

This is the brain doing what it does best, churning out "good hallucinations" by filling in and fleshing out the missing signal according to what it expects to hear. Our brain knows about the way the song sounds and the various subtleties of that specific singer's rendition, and it can use all that prior knowledge to actively predict the most likely shape of the auditory signals as the song plays. If the world doesn't send strong counterevidence, those predictions sculpt experience, making the song sound clearer to you.

It's important to emphasize that this is not a trick of memory, so much as a fascinating window on the way perception itself works. The brain's predictions for the familiar song help it carve out the signal from the noise, rendering the sounds more clearly than the bad signal would otherwise allow. Perception of this kind is highly active. It involves sending complex predictions down the chain from higher processing areas toward the sensory peripheries, generating error messages whenever a serious mismatch is detected. This backward flow is sometimes referred to by cognitive scientists as the "top-down" flow of information. While all this goes on, the human perceiver is also active, trying to gather key pieces of sensory information by means of bodily action such as turning the head or moving the eyes. These actions too are chosen and launched by the predictive machinery, creating a unified web of mental and bodily activity. We'll have lots more to say about the role of action as our story proceeds.

Putting predictions in the driver's seat in this way makes ordinary perception into what has sometimes been colorfully described as a "controlled hallucination"—the brain is guessing at how the world is by using sensory evidence mostly as a way to correct and finesse the guessing. When inner guessing completely rules the roost, we are just hallucinating, full stop. But when it is appropriately sensitive to sensory stimulations—via prediction error signals—the guessing is controlled, and the world becomes known to the mind. When we heard that familiar song on the bad radio receiver, we were benefiting from just this kind of "good hallucination." The phantom phone vibrations we met in the Preface, though in that case misleading, were generated in just the same kind of way. All human experience, if predictive processing is on track, is built in this way. We see the world by predicting the world. But where prediction errors ensue, the brain must predict again.

The Frugal Brain

Making perception turn on prediction has another important benefit too. It enables the brain to process incoming sensory information in a way that is quite remarkably efficient. The discipline that most famously examines the issue of communication efficiency is information science, which has played a major role in developing very frugal ways to transmit signals. In the mid-twentieth century, global telecommunications systems were strained by ever-increasing demand. The problem for the telecom giants was how to convey increasingly large amounts of information using just the noisy and limited channels provided by old-fashioned telephone cables. That's where