sweet anticipation

music and the psychology of expectation

david huron

Sweet Anticipation

Sweet Anticipation

Music and the Psychology of Expectation

David Huron

A Bradford Book
The MIT Press
Cambridge, Massachusetts
London, England

First MIT Press paperback edition, 2007

MIT Press books may be purchased at special quantity discounts for business or sales promotional use. For information, please email special_sales@mitpress.mit.edu or write to Special Sales Department, The MIT Press, 55 Hayward Street, Cambridge, MA 02142.

This book was set in Stone Sans and Stone Serif by SNP Best-set Typesetter Ltd., Hong Kong, and was printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Huron, David.
Sweet anticipation : music and the psychology of expectation / David Huron.
p. cm.
"A Bradford book."
Includes bibliographical references (p.) and index.
ISBN 978-0-262-08345-4 (hc : alk. paper)—978-0-262-58278-0 (pb : alk. paper)
1. Music—Psychological aspects. 2. Expectation (Psychology). I. Title.

ML3838.H87 2006 781′.11—dc22 2005054013

10 9 8 7 6

Contents

Preface

This book describes a psychological theory of expectation. I call it the *ITPRA theory*—a name that will be explained in the first chapter. When I began this research, my interests were limited to music. I was aiming to better understand how listeners form music-related expectations, and how these expectations might account for various emotional responses. As the work progressed, the ITPRA theory evolved into a general theory of expectation and so expanded beyond my parochial concerns about music. Although my principal motivations remain musical, this book should prove informative to a more general audience of readers interested in cognitive science and evolutionary psychology.

For musicians and music scholars, this book offers psychologically based insights into such venerable topics as meter, syncopation, cadence, tonality, atonality, and form. Detailed accounts of musical tension, deception, and surprise are given, and I suggest how music is able to evoke emotions such as spine-tingling chills and musically induced laughter. I provide moment-by-moment analyses of the psychological effects of common musical devices such as the appoggiatura, suspension, and anticipation, and discuss the role of expectation in crafting effective performance interpretations. In addition, I suggest how the organization of the brain might account for the taken-for-granted aesthetic distinctions of *work*, *genre*, and *rendition*. The book offers psychological interpretations of various historical events in Western music, notably the advent of musical modernism as exemplified in the works of Wagner, Stravinsky, and Schoenberg. Finally, I speculate about how the psychology of expectation might be exploited to create entirely novel musics. In general, the book attempts to show how both biology and culture contribute to the subjective phenomenal experiences that make listening to music such a source of pleasure.

For psychologists and cognitive scientists, this book offers a general theory of expectation. I suggest how the phenomenon of surprise can lead to fear, laughter, frisson, or awe. I endeavor to reconcile competing theories of emotion, most notably cognitive appraisal theories with physiologically inspired theories such as the James–Lange theory. Along the way, I propose a new interpretation of the Mere Exposure effect and suggest that misattribution is an artifact of the biological world's response to the problem of induction. An important part of the book deals with the origin of auditory expectations. In general, the experimental evidence reported here supports the emerging consensus for statistical learning. Finally, I address the problem of how expectations are mentally represented. I note that the apparently chaotic patterns found in the research on musical development are consistent with neural Darwinist theories. I further note that the differences between innate and learned representations (such as those observed in auditory localization) can be explained by the Baldwin effect. In general, I attempt to tell an uninterrupted story, from the patterns of the objective world, through imperfectly learned heuristics used for predicting that world, to the phenomenal *qualia* we experience as we apprehend the world.

In retrospect, music provided me with a serendipitous starting place for theorizing more generally about the psychology of expectation, I think for three reasons: First, most everyday experiences are too complicated to provide fruitful cases for analysis. For example, one of the most important expectation-related experiences is *surprise.* Unfortunately, many of the surprises that people experience involve complex social contexts that are often difficult to interpret. Even the simple "peek-a-boo" surprise between parent and infant involves a social dynamic that makes it hard to study. Although music is not simple, there are often fewer confounding factors to consider. A second advantage is that many musicians actively seek to provoke emotional responses in their listeners. Although different musicians pursue different goals, manipulating expectations has been a common technique used to create an emotional effect. In general, without manipulation, causality is difficult to infer; so this aspect of music often provides helpful clues about the cascade of causal events. Finally, music typically provides detailed records (in the form of notated scores) that chronicle the precipitating events as they unfold in time. Musical scores provide a convenient database for testing specific hypotheses about expectation. In short, music offers a number of advantages as a case study of the psychology of expectation. My hope is that psychologists will find the theory engaging, even if they have no interest in music.

While my main audience is intended to be musicians, I have tried to make the underlying theory of expectation accessible to nonmusician readers. Where possible, I have bracketed the technical musical descriptions as independent chapters. In chapters 1 and 2 the theory itself is described in general terms with little reference to music. Chapter 3 describes the experimental methods used to study the phenomenon of expectation. Chapters 4, 5, and 6 identify five general patterns of expectation exhibited by listeners familiar with Western music. Chapters 7, 8, 11, and 12 expand on the basic theory. While the theory itself is described in general terms, the illustrations in these chapters are drawn almost entirely from the field of music. Parallel examples in visual perception, linguistics, social behavior, and ethology will readily come to mind for those readers who are knowledgeable in such areas. Nonmusicians may wish to skip the applied discussion of music, especially chapters 9, 10, and 13 to 16. The concluding chapter (17) provides an analytic summary of the basic theory.

For readers who don't read music, the notational examples may feel irksome or irrelevant. Let me assure readers that the musical examples are genuinely illustrative. The examples have been "field tested," and for most listeners they evoke fairly reliable phenomenal experiences. To help readers grasp them more fully, recorded versions of all of the notated examples are available on the World Wide Web.[1] Professional musicians themselves may want to refer to the recorded examples since they often include performance nuances that enhance the effect.

The purpose of this book is to describe a set of psychological mechanisms and illustrate how these mechanisms work in the case of music. Some of these mechanisms are able to evoke particular emotional responses. However, this book does not provide a comprehensive theory of music and emotion; there are many other factors that contribute to musically evoked emotion that do not arise from expectation. In discussing emotional responses to music, much of the book will concentrate on how expectations are able to generate *pleasurable* experiences for listeners. The discussion will focus on general principles and will not deal directly with individual differences, such as why a person might dislike a particular tune. The emphasis on pleasure may seem controversial to some readers. Although pleasure is recognized as an important psychological motivator, it is a factor that has sometimes been overlooked or denigrated by arts scholars as merely hedonistic. It should be noted that pleasure does not trump all other values: the best music is not necessarily music that fills its listeners with pleasure. But without a significant dose of pleasure, no one would bother about music.

Pleasure does not preclude effort. Minds need to *reach*, not simply *grasp*. Brains need to be *challenged*, not simply *pampered*. If the arts are to achieve all that can be achieved, it would be wrongheaded to focus on the limitations of human minds. But neither is it the case that anything is possible. Humans are biological beings living in a social and physical world. In pursuing some artistic goal, there are often constraints that must be taken into account if the goal is to be attained.[2] To ignore these constraining phenomena is to exist in a naive delusional world. One must not mistake ignorance for imagination. Nor should wishful thinking masquerade as artistic insight.

It is essential that musicians understand that I am attempting to describe psychological processes, not aesthetic goals. My musical aim in this book is to provide musicians with a better understanding of some of the tools they use, not to tell musicians what goals they should pursue. If we want to expand artistic horizons and foster creativity there is no better approach than improving our understanding of how minds work.

Many artists have assumed that such knowledge is unnecessary: it is *intuition* rather than *knowledge* that provides the foundation for artistic creation. I agree that intuition is essential for artistic production: in the absence of knowledge, our only recourse is to follow our intuitions. But intuition is not the foundation for artistic freedom or creative innovation. Quite the contrary. The more we rely on our intuitions, the more our behaviors may be dictated by unacknowledged social norms or biological predispositions. Intuition is, and has been, indispensable in the arts. But intuition needs to be supplemented by knowledge (or luck) if artists are to break through "counterintuitive" barriers into new realms of artistic expression.

Accordingly, the best I can aim for in writing a book like this is to provide artists with some conceptual tools that might lead to something new. In the concluding chapter I highlight some of the opportunities afforded to both musicians and music scholars by taking the cognitive sciences to heart.

Acknowledgments

Research is never done in a vacuum and the research reported in this book is no exception. Much of this work was inspired by research carried out in my Ohio State University laboratory by postdoctoral fellow Paul von Hippel and doctoral student Bret Aarden. Paul von Hippel took my intuition about the possible influence of regression to the mean in melodic organization and wove a marvelous story about how listeners hear melodies. In particular, von Hippel's experiments made clear the discrepancy between what people hear and what they expect. Bret Aarden took my interest in reaction-time measures in judging melodic intervals and turned the paradigm into a truly useful tool for investigating musical expectation. His work has transformed the way we understand previous work on tonality. I am indebted to both Paul and Bret for being such tolerant listeners and for mentoring me as much as I mentored them.

Although I have always preferred so-called structural theories of tonality to functional theories, I have benefited enormously by having David Butler (the principal advocate of functional tonality) as a departmental colleague. Professor Butler's knowledgeable criticisms of structural theories led me to a better understanding of the importance of parallel mental representations. My discussion of rhythmic expectation builds on the research of another colleague, Dr. Mari Riess Jones. Along with her collaborator, Dr. Ed Large, she assembled a theory of rhythmic attending that provides the core for understanding the "when" of expectation.

Throughout the book I report a variety of statistical measures based on several databases of encoded musical scores. I am indebted to the late Helmut Schaffrath for making available his Essen Folksong Collection. I am also indebted to the Center for Computer Assisted Research in the Humanities at Stanford University for providing access to the MuseData electronic scores. Particular thanks go to Walter Hewlett, Eleanor Selfridge-Field, and Craig Sapp. In addition, my thanks go to Tom Recchia for encoding the database of pop chord progressions.

Without wishing to minimize my debt to my colleagues and collaborators, I must acknowledge that the most important people in any research program are one's critics.

Over the years, I have come to greatly value the preventive medicine provided by the peer review process. The sad truth about writing a book is that it is immensely difficult to cajole knowledgeable people into providing critical feedback. The distinguished hearing scientist, Georg von Békésy, once lamented that his successful career had resulted in the loss of his best critics. For von Békésy, former critics were ultimately transformed into friends, and with that, he felt that the quality of his research had suffered. One way to compensate for the paucity of critics is to encourage friends to put aside their affections and pull out their scalpels. I am grateful to those of my associates who recognize the value of mixing encouragement with pointed criticism. My thanks to Bob Snyder, Ian Quinn, Elizabeth Margulis, David Temperley, Fred Lerdahl, Zohar Eitan, Marc Perlman, Dirk-Jan Povel, Peter Desain, Ryan Jordan, Donald Gibson, William Conable, Peter Culicover, Kristin Precoda, Simon Durrant, James Wright, Jonathan Berger, Joy Ollen, Randolph Johnson, Joshua Veltman, Judy Feldmann, Marion Harrison, Freya Bailes, and Bruno Repp.

Sweet Anticipation

1 Introduction

The world provides an endless stream of unfolding events that can surprise, delight, frighten, or bore. Such emotions provide the intimate experiences that define our personal lives. Sometimes emotions are overwhelming—as when we experience great pleasure or great agony. More often, emotions add subtle nuances that color our perceptions of the world. Emotions add depth to existence; they give meaning and value to life.

How do emotions arise? What purposes do they serve? What accounts for the distinctive feelings we experience? These questions have stimulated philosophers for centuries. More recently, these questions have inspired the curiosity of psychologists and cognitive scientists. But they are also questions that attract the attention of the "practitioners" of emotion. Playwrights, novelists, poets, film directors, musicians, choreographers, comedians, and theatrical magicians all have a professional interest in what distinguishes delight from boredom. Therapists, game designers, carnival operators, and traffic engineers have good reasons to try to understand what causes people to be surprised or fearful. Even advertisers and politicians have practical motivations for understanding how the flux of events shape human emotional experiences.

It is no coincidence that the performing arts have figured prominently in attempts to understand the dynamics of emotion. Over hundreds of years, poets, actors, comedians, and musicians have developed a sort of folk psychology about how certain emotions can be generated. Of the many arts, music has perhaps faced the most onerous challenges. Where the poet or playwright can evoke sadness by narrating a recognizably sad story, musicians must create sadness through abstract nonrepresentational sounds. Where a comedian might evoke laughter through parody, wordplay, or absurd tales, musicians must find more abstract forms of parody and absurdity. Where magicians evoke awe by appearing to transgress the laws of physics, no comparable recipe exists for creating musical awe. Despite the difficulties, musicians have amply demonstrated an exquisite skill in evoking the profoundly sad, the twistedly absurd, and the deeply awe-inspiring.

In each of the arts, codes of practice, heuristic rules of thumb, and speculative theories have been passed from teacher to student across the generations. These folk psychologies are based on a combination of intuition and tried-and-true techniques. In music, composers absorb a number of clichés—useful devices that are most easily observed in film scores. Trained musicians will readily recognize some commonplace examples: *tragedy* can be evoked by using predominantly minor chords played with rich sonorities in the bass register. *Suspense* can be evoked using a diminished seventh chord with rapid tremolo. *Surprise* can be evoked by introducing a loud chromatic chord on a weak beat.

For many thoughtful musicians, such clichés raise the question, "Why do these techniques work?" To this question, an ethnomusicologist might add a second: "Why do they often *fail* to work for listeners not familiar with Western music?" And an experienced film composer might insist on adding a third: "Why do they sometimes fail to work, even for those who are familiar with Western music?" In addressing these questions, intuition and folk psychology provide important starting points. But if we want to probe these questions in depth, we must ultimately embrace a more systematic approach. In theorizing about music and emotion, it is inevitable that we must move beyond folk psychology to psychology proper.

Many of the arts achieve specific emotional effects through a sort of stylized depiction or representation of common emotional displays. The mime exaggerates human body language and facial expressions. The cartoonist distills these same expressions into a few suggestive pen strokes. Even when a dancer aims for a strictly formal performance, her body movements will still tend to imply natural gestures or socially defined expressions. Music too involves mimicry of some natural emotional expressions. But aesthetic philosophers and music commentators have long noted that music is not a "representational" art in the way that painting or sculpture can be. How is music so successful in evoking emotions when its capabilities for representing the natural world seem so constrained?

In the 1950s, the renowned musicologist Leonard Meyer drew attention to the importance of *expectation* in the listener's experience of music. Meyer's seminal book, *Emotion and Meaning in Music,* argued that the principal emotional content of music arises through the composer's choreographing of expectation. Meyer noted that composers sometimes thwart our expectations, sometimes delay an expected outcome, and sometimes simply give us what we expect. Meyer suggested that, although music does contain representational elements, the principal source for music's emotive power lies in the realm of expectation.[1]

As a work of music theory, Meyer's approach was pioneering in its frequent appeals to psychological explanations. Despite Meyer's interest in psychology, however, *Emotion and Meaning in Music* was written at a time when there was little pertinent psychological research to draw on. In the intervening decades, a considerable volume of experimental and theoretical knowledge has accumulated.[2] This research provides an opportunity to revisit Meyer's topic and to recast the discussion in light of contemporary findings. The principal purpose of this book is to fill in the details and to describe a comprehensive theory of expectation—a theory I have dubbed the "ITPRA" theory.

Of course, expectations are not the province of music alone; expectation is a constant part of mental life. A cook expects a broth to taste a certain way. A pedestrian expects traffic to move when the light turns green. A poker player expects an opponent to bluff. A pregnant woman expects to give birth. Even as you read this book, you have many unconscious expectations of how a written text should unfold. If my text were abruptly to change topics, or if the prose suddenly switched to a foreign language, you would probably be dismayed. Nor do the changes need to be dramatic in order to have an effect. Some element of surprise would occur if a sentence simply ended. Prematurely.

Any theory of musical expectation necessarily presupposes a general theory of expectation. The ITPRA theory is intended to provide such a general theory. The theory is ambitious in scope and aims to account for all of the main psychological phenomena related to expectation. In particular, the ITPRA theory endeavors to account for the many emotion-related elements of expectation. The theory attempts to explain how expectations evoke various feeling states, and why these evoked feelings might be biologically useful.

The story of expectation is intertwined with both biology and culture. Expectation is a biological adaptation with specialized physiological structures and a long evolutionary pedigree. At the same time, culture provides the preeminent environment in which many expectations are acquired and applied. This is especially true in the case of music, where the context for predicting future sounds is dominated by cultural norms. In attempting to understand expectation, it is essential to take both biology and culture seriously. Accordingly, my text will freely meander through such topics as physiological and evolutionary psychology, learning, enculturation, style, and music history.

From an evolutionary perspective, the capacity to form accurate expectations about future events confers significant biological advantages. Those who can predict the future are better prepared to take advantage of opportunities and sidestep dangers. Over the past 500 million years or so, natural selection has favored the development of perceptual and cognitive systems that help organisms to anticipate future events. Like other animals, humans come equipped with a variety of mental capacities that help us form expectations about what is likely to happen. Accurate expectations are adaptive mental functions that allow organisms to prepare for appropriate action and perception.

But what about the emotional "feelings" that are often conjured up as a result of expectations? What gives *anticipation* or *surprise* their distinctive phenomenological characters? The story of emotion is intertwined with the psychology of behavioral motivation. Emotions are motivational amplifiers.[3] Emotions encourage organisms to pursue behaviors that are normally adaptive, and to avoid behaviors that are normally maladaptive. In this regard, the emotions evoked by expectation do not differ in function from other emotions. As we will see, the emotions accompanying expectations are intended to reinforce accurate prediction, promote appropriate event-readiness, and increase the likelihood of future positive outcomes. We will discover that music-making taps into these primordial functions to produce a wealth of compelling emotional experiences. In this way, musicians are able to create a number of pleasurable emotional experiences, including surprise, awe, "chills," comfort, and even laughter.

The biological purpose of expectation is to prepare an organism for the future. A useful place to begin is to consider, in general, what it means to be prepared.

Preparation

When you switch on a light, electrical energy streams down a convoluted path of wires from a distant power station. The speed with which this happens is impressive. The electricity flows at nearly the speed of light, which means that the power you consume was generated less than one one-thousandth of a second earlier. There is no time at the power station to "gear up" for your demand. The turbine generators must already be producing the electricity that the power company thinks you (and other customers) might need. Any energy generated that is not used by current customers is simply wasted: fuel is burned for no good reason. Clearly, power companies have a strong incentive to anticipate precisely how much power should be produced at any given moment in time.

All biological organisms consume power—power to maintain metabolisms, to move muscles, and to spark nervous systems. This power is expensive. It must be generated from the food the animal consumes, and gathering food is difficult, time-consuming, and very often dangerous. As with the electrical grid, the amount of power required by an organism changes from moment to moment, so it is important for the animal to avoid waste by matching the amount of energy generated with the amount the animal needs.

Commercial power producers employ teams of statisticians whose sole job is to try to predict power demands. They estimate what time people will get up on Saturday morning, how many people are likely to watch the big game on TV, and whether the outside temperature will entice customers to turn on their air conditioners. The predictive models used by utility companies are elaborate and impressive feats of human ingenuity. But like so many other human creations, the complexity and efficiency of these predictive models pale when compared with the achievements of nature.

Organisms are constantly trying to conserve energy. Bodies (including brains) drift toward low states of arousal when no action or thought is needed. In a static unchallenging environment, minds grow bored and bodies grow limp. We respond to these environments by invoking nature's all-purpose energy-conservation strategy—*sleep*. Of course, sometimes the events of the world do require some appropriate action, and so the body and mind must be roused in preparation. Like a machine that has been turned off, a certain amount of time is needed for us to "power up."

When you unexpectedly hear the sound of a barking dog, your heart will quicken and the volume of blood flowing to your muscles will increase. At the same time, an important hormone, norepinephrine, will be released in your brain making you more alert and attentive. In truly dangerous situations, this response, quick as it is, may prove to be too slow. Like a power "brown out," the demands of the body might momentarily exceed the supply of resources. Many animals have become another animal's dinner in the split second required to respond to danger. If only one could have known in advance to increase the power output and pay closer attention. If one could have anticipated the danger, a more effective response might have been rallied.

Over the eons, brains have evolved a number of mechanisms for predicting the future. The biological purpose of these mechanisms is to prepare the body and mind for future events while simultaneously minimizing the consumption of metabolic resources. From a physiological perspective, there are two interrelated systems that influence metabolic consumption: *arousal* and *attention*. The arousal system controls heart rate, respiration, perspiration, and many other functions associated with movement. The attention system is more subtle. Attention spurs the brain to be more engaged with the world. Instead of looking at nothing in particular, our gaze becomes focused. Instead of tuning out a conversation, we pay close attention to what is being said. Instead of daydreaming, we become grounded in the here and now. All of this takes energy.

Arousal and attention levels fluctuate according to both the actual and the anticipated demands of the environment. When we think of arousal and attention reacting to the environment, there is a tendency to think foremost of them as *increasing*. However, the arousal and attention systems can also *reduce* or *inhibit* responsiveness. The experiences of boredom and sleepiness are no less manifestations of metabolic fine-tuning than are the experiences of excitement and exhilaration.

We may also tend to think of arousal and attention as systems that deal necessarily with the uncertainties of life. But even if we knew with exact precision and certainty all of the future events in our lives, we would still need anticipatory mental and corporeal changes to fine-tune our minds and bodies to the upcoming events. Suppose, for example, that I know that at 9:18 A.M. I will encounter an obstacle on the path requiring me to steer my bicycle around it. This godlike foreknowledge does not absolve me from having to attend to the object and make the appropriate motor movements at the appointed time. Nor can I execute any of the needed mental or corporeal maneuvers before they are required. So perfect knowledge of the future would not change the fact that attention and arousal levels must fluctuate according to the moment.

Of course, such perfect knowledge of the future doesn't exist; we do live in a world in which the future is uncertain, and this uncertainty does make it more difficult to produce the optimum arousal and attention. How do we prepare for a future that has untold possibilities? Sometimes this uncertainty doesn't matter. There are some situations where the precise outcome is highly uncertain, but where all of the potential outcomes would require the same type of mental and physical preparation. In a casino, a roulette croupier has no idea which number will appear on the wheel, but the croupier's ensuing actions are highly practiced: collect the chips from the losing bets and reward any successful bets. While the croupier's actions are obviously guided by the result on the roulette wheel, the croupier's response depends very little on the specific outcome—unlike the responses of the gamblers!

These sorts of situations are not commonplace, however. More commonly, different outcomes will require different optimum responses. The body typically faces a quandary: which of several possible outcomes does one prepare for? In preparing the body and mind for these outcomes, our instincts are depressingly pessimistic. Like a grumbling naysayer, nature tends to assume the worst. Consider, for example, the slamming of a door. Even though we may see that the door is about to slam shut, it is difficult to suppress the impending startle or defense reflex. We know the door poses no danger to us, but the sound of the slamming door provokes a powerful bodily response anyway. Despite our annoyance, nature knows best: it is better to respond to a thousand false alarms than to miss a single genuinely dangerous situation.

As we will see later, nature's tendency to overreact provides a golden opportunity for musicians. Composers can fashion passages that manage to provoke remarkably strong emotions using the most innocuous stimuli imaginable. As every music-lover knows, simple sequences of sounds hold an enormous potential to shape feelings. As we will see, it is nature's knee-jerk pessimism that provides the engine for much of music's emotional power—including feelings of joy and elation.

The object of expectation is an event in time. Uncertainty accompanies not only *what* will happen but also *when* it will happen. Sometimes the *when* is certain but not the *what*. Sometimes the *what* is known, but not the *when*. Later, we will see how music manipulates both kinds of uncertainty, and how the different what/when combinations produce different emotional responses.

Along with *what* and *when,* brains also predict *where* and *why*—but these are more specialized operations. For sound stimuli, the *where* expectations are associated with physiologically ancient structures for sound localization. Musicians have sometimes manipulated the locations of sounds (as in the antiphonal works of Giovanni Gabrieli or the electroacoustic works of Karlheinz Stockhausen), but they have less often manipulated listener *expectations* of location. The *why* expectations are associated with physiologically recent structures associated with conscious thought. In contrast to the *what* and *when* of prediction, the *where* and *why* components of auditory expectation have played little role in musical organization and experience. But they represent opportunities for future enterprising composers.

Emotional Consequences of Expectations

As I have noted, the ability to anticipate future events is important for survival. Minds are "wired" for expectation. Neuroscientists have identified several brain structures that appear to be essential for prediction and anticipation. These include the substantia nigra, the ventral tegmental area, the anterior cingulate cortex, and the lateral prefrontal cortical areas.[4] Most people will regard such biological facts as uninteresting details. For most of us, the more compelling details pertain to the subjective experience. From a phenomenological perspective, the most interesting property of expectation is the feeling that can be evoked. What happens in the future matters to us, so it should not be surprising that how the future unfolds has a direct effect on how we feel.

Why precisely do expectations evoke various feeling states? I propose that the emotions evoked by expectation involve five functionally distinct physiological systems: imagination, tension, prediction, reaction, and appraisal. Each of these systems can evoke responses independently. The responses involve both physiological and psychological changes. Some of these changes are autonomic and might entail changes of attention, arousal, and motor movement. Others involve noticeable psychological changes such as rumination and conscious evaluation.

Outcomes matter, so the evoked emotions segregate into positive and negative kinds. That is, the feeling states are *valenced*. Positive feelings reward states deemed to be adaptive, and negative feelings punish us for states deemed to be maladaptive. The word "deemed" here is important. Positive feelings are evoked not by results that are objectively adaptive, but by results that the brain, shaped by natural selection, presumes to be adaptive. From time to time the evoked emotions are wrongheaded. For example, a family pet may experience acute distress when being taken to the veterinarian—despite the fact that the medical attention objectively increases the animal's well-being. Like the family pet, we can feel that our world is falling apart even while good things are happening to us. Each of the five response systems makes different assumptions about what is good or bad. So different emotions can be evoked by each of the five systems.

The five response systems can be grouped into two periods or epochs: *pre-outcome* responses (feelings that occur prior to an expected/unexpected event) and *post-outcome* responses (feelings that occur after an expected/unexpected event). Our discussion begins with two types of pre-outcome responses: those of the imagination and the tension systems.

1 Imagination Response

Some outcomes are both uncertain and beyond our control. The weather provides a good example. It may or may not rain, but you are helpless to influence either outcome. Other outcomes, however, may lie within our control. If it rains, you might get wet; but if you carry an umbrella you can reduce the probability of that outcome. In short, people have no control over "rain," but we sometimes have control over "getting wet."

At some point in animal evolution, the ability to predict aspects of the future led to the emergence of other mental mechanisms that attempted to ensure that particular future outcomes were more likely to occur than others. Once an animal is able to predict that some events are likely, there is a lot to be gained if one behaves in a fashion that increases the likelihood of a favorable outcome.

Imagining an outcome allows us to feel some vicarious pleasure (or displeasure)—as though that outcome has already happened. You might choose to work overtime because you can imagine the embarrassment of having to tell your boss that a project remains incomplete. You might decide to undertake a difficult journey by imagining the pleasure of being reunited with a loved one. This *imagination response* is one of the principal mechanisms in behavioral motivation. Through the simple act of daydreaming, it is possible to make future outcomes emotionally palpable. In turn, these feelings motivate changes in behavior that can increase the likelihood of a future favorable result.[5]

Neurological evidence for such an imagination response is reported by Antonio Damasio, who has described a clinical condition in which patients fail to anticipate the feelings associated with possible future outcomes.[6] In one celebrated case, Damasio described a patient ("Elliot") who was capable of feeling negative or positive emotions after an outcome had occurred, but was unable to "preview" the feelings that would arise if a negative outcome was imminent. Although Elliot was intellectually aware that a negative outcome was likely, he failed to take steps to avoid the negative outcome because, prior to the outcome, the future negative feelings were not palpable and did not seem to matter. Damasio's clinical observations have established that it is not the case that we simply think about future outcomes; when imagining these outcomes, we typically are also capable of feeling a muted version of the pertinent emotion. We don't simply *think* about future possibilities; we *feel* future possibilities.

The imagination response provides the biological foundation for deferred gratification. Feelings that arise through imagination help individuals to forgo immediate pleasures in order to achieve a greater pleasure later. Without this imaginative emotional capacity, our lives would be dominated entirely by petty excitements. From time to time, pop psychologists and self-appointed spiritual advisors have advocated that people focus on living in the here and now and let go of their concerns for the future. Damasio's patients have achieved exactly such a state. For these individuals, the future is a gray abstraction that is irrelevant to the business of living. As a consequence, they lose their friends, go bankrupt, and live lives in which present-tense joys become increasingly hard to achieve because they are unable to plan ahead. It is important to pause and smell the roses—to relish the pleasures of the moment. But it is also crucial to take the imaginative step of planting and nurturing those roses.

If we think of positive and negative feelings as hills and valleys in a complex landscape, the imagination response helps us avoid getting stranded at the top of the nearest hill. Imaginative *thought* allows us to see the higher peaks that might be experienced if only we are willing to first descend into one or more valleys. But it is imaginative *emotions* that motivate us to undertake the difficult journey to reach those higher peaks.

2 Tension Response

A second form of pre-outcome emotional response originates in the mental and corporeal preparation for an anticipated event. At a party, a friend approaches you with a balloon in one hand, and a sharp pin poised for action in the other hand. The grin on your friend's face suggests that the balloon is not likely to remain inflated for long. You squint your eyes, put your fingers in your ears, and turn your face away.

Preparing for an expected event typically involves both motor preparation (arousal) and perceptual preparation (attention). The goal is to match arousal and attention to the expected outcome and to synchronize the appropriate arousal and attention levels so that they are reached just in time for the onset of the event. Usually, events require some increase in arousal. Heart rate and blood pressure will typically increase, breathing will become deeper and more rapid, perspiration will increase, and muscles will respond faster. In addition, pupils may dilate, eyes may focus, the head may orient toward (or away from) the anticipated stimulus, and distracting thoughts will be purged. These (and other) changes help us to react more quickly and to perceive more accurately.

If we want to conserve the maximum amount of energy, then we ought to wait until the last possible moment before increasing attention or arousal. If it only takes a second or two to reach an optimum arousal, then we shouldn't begin increasing arousal until a second or two prior to the outcome. This simple ideal is confounded, however, by uncertainty—uncertainty about *what* will happen, and uncertainty about *when* it will happen. When we are uncertain of the timing of the outcome, we must raise arousal or attention levels in advance of the earliest anticipated moment when the event might happen. If the actual event is delayed, then we might have to sustain this heightened arousal or attention for some time while we continue to wait for the event.

I once saw a couple moving from their second-storey apartment. Having tired of running up and down the stairs, they had resorted to dropping bundles of clothing from their apartment balcony. She would toss bags over the railing while her partner would try to catch them before they hit the ground causing the plastic bags to split. Unfortunately, the physical arrangement prevented the two of them from making eye contact. As a consequence, her partner stood on the ground with his arms perpetually outstretched, unsure of when the next bag would drop out of the sky. I recall this incident because the man looked so silly—like something out of a Laurel and Hardy film. At one point, I could see that the woman had gone back into the apartment to fetch some more bags, but the man was still staring intently upward, arms outstretched, rocking back and forth in anticipation. He was wasting a great deal of energy because the timing of expected events was so uncertain.

Apart from uncertainty regarding *timing*, we may have difficulty tailoring the *level* of arousal or attention. When the exact nature of the outcome is uncertain, it can be difficult to match precisely the arousal and attention levels to the ultimate outcome. The safest strategy is to prepare for whatever outcome requires the highest arousal and/or attention level. In a baseball game, a fielder can clearly see the pitcher's windup and whether or not the batter swings. There is little doubt about the timing of outcomes. The uncertainty resides principally with the *what*: Will the batter hit the ball? And if so, will the ball be hit into the fielder's area of play? The actual probability of a batter hitting any given pitch into the vicinity of the fielder is comparatively low. Nevertheless, the fielder's best preparation is to assume the worst—namely, a hit into the fielder's area. Unfortunately, this vigilance comes at a cost. With each pitch, arousal and attention peak and then subside (presuming no action is needed). Maintaining high levels of arousal and attention will cause the player to expend a lot of energy. In an important championship game, a fielder is apt to be exhausted by the end of the game, even if the player never had to field a ball.

In the case of the man catching bags of clothing, the uncertainty relates mostly to the *when*, not the *what*: a bag of clothing would surely drop out of the sky, but the timing was uncertain. In the case of the baseball fielder, the uncertainty pertains mostly to the *what*, not the *when*: the ball can only be hit after the pitcher throws the ball. But the outcome of each pitch is uncertain.

The most uncertain situations are those where both the *when* and the *what* are unknown. A soldier on guard duty, for example, might have reason to fear a possible attack. Although the shift may pass uneventfully, the heightened attention and arousal engendered by the expectation of a possible attack is likely to produce acute mental and physical exhaustion.

As it turns out, the physiological changes characteristic of high arousal are also those associated with stress. Not all high arousal is stressful: positively valenced emotions such as joy and exuberance will evoke high arousal with little stress. But anticipating negative events is sure to be stressful. In dangerous situations, organisms respond with one of three classic behaviors: *fighting*, *fleeing*, or *freezing*. The greatest stress tends to occur when high arousal coincides with low movement. Consequently, it is the *freeze* response that engenders the most stress. Fighting and fleeing are active responses, while the freeze response is often symptomatic of helplessness. This is thought to be the reason why the worst health effects of stress are to be found in those people who are unable to do anything to alleviate their stressful conditions.

When anticipating some future event, our physiological state is often akin to that of the freeze response. We may experience elevated heart rate and perspiration without any motor movement. The word "dread" captures the stressful feeling that accompanies anticipating a bad future outcome. By contrast, anticipating something positive evokes a feeling something like being "heartened." But even anticipating something positive has some accompanying stress. In the pre-outcome period, nothing is certain, and so our heartened state is likely to be mixed with a nagging fear that an anticipated positive result may not actually come to pass.

Since stress commonly accompanies the rise of anticipatory arousal, I have chosen the word "tension" to characterize these sorts of pre-outcome responses. Both the baseball fielder and the man catching bags of clothing were experiencing distinct physiological states in anticipation of future outcomes.

Unlike the imagination response, the tension response is linked to the period immediately prior to the anticipated moment of outcome. As the arousal and attention levels move toward an optimum level in anticipation of the outcome, the physiological changes themselves evoke characteristic feeling states. The feelings that accompany the tension response are artifacts. The evoked feeling states have no particular function by themselves, but are simply consequences of the physiological changes that accompany preparation for an anticipated outcome.

The "artifact" status of certain emotions was famously proposed by William James and Carl Lange roughly a century ago.[7] In an often quoted passage, James argued that fear was evoked by the act of trembling, sorrow was evoked by the act of crying, and so on.[8] This "James–Lange" theory of emotion has a checkered history. Some important research supports the theory.[9] One example is found in a simple experiment carried out by Fritz Strack and his colleagues where participants were asked to hold a pencil in their mouth.[10] In one condition, participants held the pencil using their teeth without allowing their lips to touch the pencil. In a contrasting condition, participants held the pencil with their lips only. Strack showed that the manner by which participants hold the pencil has a direct effect on how they feel. Grasping a pencil between your teeth causes you to feel happier than grasping it with your lips. The difference can be traced to the flexing of the zygomatic muscles: holding a pencil between your teeth produces something very similar to smiling. It is not just that you smile when happy—you can feel happy *because* you smile.[11]

The research by Strack and others notwithstanding, there is also research that is wholly inconsistent with the James–Lange theory.[12] It is probably the case that the sort of physiologically induced emotions described by James and Lange are limited to a handful of particular circumstances. I propose that the tension response is one of the circumstances in which the James–Lange theory holds. Simply flexing muscles in anticipation of catching a ball will change a person's feeling state. The evoked feelings will depend on which muscles are flexed. Flexing abdominal muscles will tend to evoke a different affect than squinting eyes, smiling, or clutching a steering wheel.

There are several factors that influence the character and magnitude of the tension response. These include the degree of uncertainty, the importance of the possible outcomes, the difference in magnitudes between the best and worst plausible outcomes, and the estimated amount of time before the outcome is realized. Sometimes outcomes are utterly certain and have little consequence. These situations evoke little tension. In other cases, we may have little idea about what will happen. If one or more of the possible outcomes involves a high stake (something very good or very bad), then we will tend to be more alert and aroused as the moment approaches when the outcome will be made known.

In general, organisms should try to avoid situations of high uncertainty. High uncertainty requires arousal and vigilance, both of which incur an energy cost. Consequently, it would be adaptive for an organism to experience high tension responses as unpleasant. That is, even if only positive outcomes are possible, high uncertainty will lead to a certain amount of unpleasant stress.[13]

3 Prediction Response

Once some event occurs, there ensues a convoluted sequence of physiological and psychological changes. It is useful to distinguish three post-outcome responses.

As you might suppose, organisms respond better to expected events than unexpected events. Accurate predictions help an organism to prepare to exploit opportunities and circumvent dangers. When a stimulus is expected, appropriate motor responses are initiated more rapidly and more accurately. In addition, a stimulus is more accurately perceived when it is predictable.

Since accurate predictions are of real benefit to an organism, it would be reasonable for psychological *rewards and punishments to arise in response solely to the accuracy of the expectation*. Following a snow storm, for example, I might predict that I will slip and fall on the sidewalk. In the event that I actually fall, the outcome will feel unpleasant, but the experience will be mixed with a certain satisfaction at having correctly anticipated this dismal outcome. This expectation-related emotion might be dubbed the *prediction response*. When the stimulus is expected, the emotional response is postively valenced; when the stimulus is unexpected, the emotional response is negatively valenced.

Psychological evidence in support of a prediction response is found in the classic work of George Mandler.[14] An abundance of subsequent experimental research has affirmed the importance of this response. In fact, this response is considered so important in the extant literature on expectation that it is commonly referred to as the *primary affect*.[15] Confirmation of expected outcomes generally induces a positive emotional response even when the expected outcome is bad. It is as though brains know not to shoot the messenger: accurate expectations are to be valued (and rewarded) even when the news is not good. We will devote an extended discussion to this important response in chapter 8.

4 Reaction Response

The most obvious emotions in the post-outcome epoch are those that pertain to the pleasantness or unpleasantness of the outcome itself. Once an outcome is known, our emotions reflect some sort of assessment of the new state. For example, we might experience fear when encountering a snake, sadness when receiving a poor grade, or joy when meeting an old friend. These emotional responses occur only after the outcome is known.

Extensive research has established that there are two types of responses to the advent of events. One type of response is very fast. The other type of response is more leisurely. The fast response represents a "quick-and-dirty" assessment of the situation followed by an immediate somatic (bodily) response. The second response represents a more "thoughtful" assessment of the situation—a response that takes into account complex social and environmental factors. I propose to call the fast response a *reaction response,* and the more complex slower response an *appraisal response.*

Reaction responses exhibit three characteristic features: (1) The response has a fast onset. Typically, the onset of the response begins less than 150 milliseconds following the onset of the outcome. Although the onset of the response is fast, the somatic changes arising from the response might continue for several seconds afterward. (2) The response is not mediated by consciousness. No conscious thought or rumination is involved. Some reaction responses can even occur when we are asleep. (3) The response is defensive or protective in function. The reaction assumes a worst-case scenario, and responds accordingly.

An example of a reaction response is a *reflex*. Suppose that you accidentally touch a hot oven. A well-documented reflex will cause your hand to be abruptly withdrawn from the hot surface. Surprisingly, this reflex is so fast that it happens in less time than it takes for a neural signal to travel from the hand up to the brain and then back down from the brain to the muscles of the arm. Physiologists have determined that the reflex originates in the spine rather than in the brain. So-called *reflex arcs* in the spine connect the sensory neurons in the hand to the motor neurons of the arm. You have withdrawn your hand before your brain even registers the sensation of the hot surface. The reflex has a fast onset, is not mediated by consciousness, and has a defensive function.

Reflexes are examples of reaction responses, but not all reaction responses are technically reflexes. As we will see later, reaction responses can also be *learned*—which is not the case with reflexes. More specifically, we will see that learned *schemas* are used in reaction responses. These learned reaction responses are easiest to observe in situations of surprise. By way of example, consider wrong with speak. Violations of grammar—such as in the preceding sentence—evoke a mild but rapid surprise. Of course English grammar is entirely learned, so the reaction can't be considered a reflex—despite its speed and automaticity. The surprise here arises from a discrepancy between an actual outcome and a highly practiced schema.

Learned schemas span a huge range of behaviors. Schemas can relate to practiced motor skills (such as brushing your teeth) or perceptual norms (such as watching traffic flows). Schemas can involve social norms (such as polite greeting rituals) or cultural norms (such as framing an object so that it is recognized as "art"). As long as the schema is well entrenched in a mind, it becomes possible to provoke reaction responses by violating the schematic expectation. In chapter 2 we will consider such reactive surprises in greater detail, and focus specifically on the feeling states that can be evoked.

5 Appraisal Response

Suppose you answer the phone and are pleasantly surprised to hear the voice of a close friend. Within a second, your pleasure turns to acute embarrassment as you realize that you have forgotten your friend's recent birthday. Or, imagine an experienced biologist is walking in a forest and is startled when a large spider drops onto the sleeve of her jacket. Her negative feelings immediately turn to joy as she realizes that she may have discovered a new species of spider.

Our initial reactions to events are susceptible to revision or augmentation. What we find initially exciting or startling may be completely transformed by further thought. The *reaction response* is quick and unconscious. Once conscious thought is engaged, the assessment of a situation is the province of the *appraisal response*. The above examples are illustrations of when the appraisal response and the reaction response evoke contrasting emotions. But the two responses may also reinforce one another. A near accident in an automobile might quickly evoke a feeling of fear. The subsequent recognition that you were not wearing a seat belt and that any accident would have likely proved fatal might provoke an even stronger sense of fear. Moreover, further conscious thought might lead you to realize that you are behind in your life insurance payments, and that had you died, your children would not have been adequately provided for—hence evoking even greater fear.

As you continue to ruminate about a situation, several successive appraisal responses might ensue. The important point is that appraisal responses can involve conscious thought that often draws on complex social and contextual factors. By contrast, the reaction response involves no conscious thought.

The reaction response and the appraisal response are independent and need not be consistent with each other. As we have seen, a single outcome can produce a negatively valenced reaction response and a positively valenced appraisal response (or vice versa). We will see many examples of such paradoxical feeling states in later chapters. In addition, different people may experience similar reaction responses, but contrasting appraisals. Consider, for example, two office workers who are both startled by the unexpected ringing of their telephones. After the initial start, the worker in the sales department may become excited because the call represents the possibility of making a sale (with an accompanying commission). But the worker in the customer service department might react more negatively, since the call likely represents a customer with a complaint.

In general, positive and negative emotions act as behavioral reinforcements. The pain caused by biting your tongue teaches you to chew carefully and avoid tissue damage. Bad tastes and bad smells reinforce the aversion to ingesting unhealthy foods. The pleasure caused by engaging in sex encourages procreation. The enjoyment of playing with our children encourages parental investment and nurturing. Positive emotions encourage us to seek out states that increase our adaptive fitness. Negative emotions encourage us to avoid maladaptive states.

The ITPRA Theory of Expectation

To summarize: I have distinguished five expectation-related emotion response systems. Each response system serves a different biological function. The purpose of the *imagination response* is to motivate an organism to behave in ways that increase the likelihood of future beneficial outcomes. The purpose of the *tension response* is to prepare an organism for an impending event by tailoring arousal and attention to match the level of uncertainty and importance of an impending outcome. The purpose of the *prediction response* is to provide positive and negative inducements that encourage the formation of accurate expectations. The purpose of the *reaction response* is to address a possible worst-case situation by generating an immediate protective response. The purpose of the *appraisal response* is to provide positive and negative reinforcements related to the biological value of different final states. All of these goals are biologically adaptive. Table 1.1 summarizes these five response systems and presents them in their approximate order in time.

Table 1.1

Response system	Epoch	Biological function
(I) *imagination response*	pre-outcome	future-oriented behavioral motivation; enables deferred gratification
(T) *tension response*	pre-outcome	optimum arousal and attention in preparation for anticipated events
(P) *prediction response*	post-outcome	negative/positive reinforcement to encourage the formation of accurate expectations
(R) *reaction response*	post-outcome	neurologically fast responses that assume a worst-case assessment of the outcome
(A) *appraisal response*	post-outcome	neurologically complex assessment of the final outcome that results in negative/positive reinforcements

Informally, we might characterize the "feeling" components to these responses by posing five questions:

1. What do you think might happen, and how do you feel about that prospect?
2. Are you ready for what's about to happen? How do the preparations make you feel?
3. Did you "place a good bet"—did you predict the outcome accurately? Are you pleased or disappointed by the accuracy of your wager?
4. Assuming the worst, how have you reacted? How does this reaction make you feel?
5. Upon reflection, how do you feel about how things have turned out?

Once again, these five response systems are evoked at different times in the expectation cycle. The imaginative function may begin years prior to an expected event. A person might imagine the sense of achievement associated with graduating from college or paying off a mortgage. As an anticipated event approaches, the emotions evoked by the imagination become dwarfed by the feelings evoked by the mental and corporeal preparations for the actual event—especially if the outcome is uncertain. These preparatory responses relate predominantly to a sense of stress or tension. Once the outcome is known, three response systems are set in motion. One component simply responds to the accuracy of the prediction. In tandem with this prediction response are the emotional states evoked by the reaction and appraisal responses. A short-lived reaction response is typically replaced by the more nuanced appraisal response. Like the imagination phase, the appraisal emotions have the potential to last for years. One may still feel good about some long-ago success, or feel regret about some long-ago failure. The time course of these different emotional responses is illustrated in figure 1.1. It is this time-course that leads to the acronym ITPRA: Imagination–Tension–Prediction–Reaction–Appraisal. Since the prediction and reaction responses

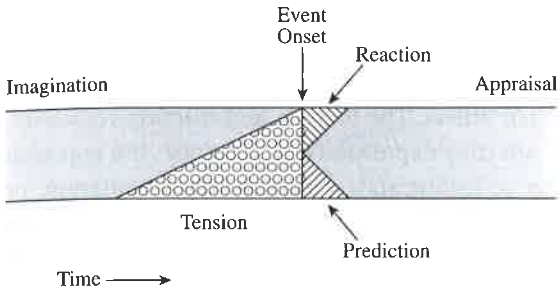

Figure 1.1
Schematic diagram of the time-course of the "ITPRA" theory of expectation. Feeling states are first activated by imagining different outcomes (**I**). As an anticipated event approaches, physiological arousal typically increases, often leading to a feeling of increasing tension (**T**). Once the event has happened, some feelings are immediately evoked related to whether one's predictions were borne out (**P**). In addition, a fast reaction response is activated based on a very cursory and conservative assessment of the situation (**R**). Finally, feeling states are evoked that represent a less hasty appraisal of the outcome (**A**).

occur in tandem, one might equally call it ITRPA (reversing the R and P), but I prefer the more pronounceable ITPRA.

As I have noted, I propose that these five response systems arise from five functionally distinct neurophysiological systems. Each response system solves an important problem in tailoring behavior so that it is optimally adapted to a given environment. Since each response system addresses a different biological problem it is possible that each system represents a distinct evolved adaptation. One might even propose a plausible order of evolution for these systems. The oldest response system is probably the (unconscious) reaction response. Clearly, an organism must take appropriate actions in response to what actually happens in the world; outcomes (and our responses to them) are what matter most. An organism that always assumes the worst outcome has a better chance of surviving those occasional situations that are truly dangerous. Since these hasty reaction responses are commonly exaggerated, some basic elements of the appraisal response probably evolved next. This would have begun as an (unconscious) inhibitory function, suppressing those reaction responses that are excessively conservative. The tension response was likely next to evolve. Simple classical conditioning might allow an organism to anticipate what happens next, and there are clear advantages to tailoring the arousal and attention to the expected event. Since the prediction response provides a way to evaluate the predictions implicit in the tension response, the prediction response must have appeared after the advent of the tension response. Finally, the imagination response is probably the most recent evolutionary addition. Once one achieves some modicum of success in predicting the future, there is obvious value in trying to change the future through our own actions.

Each of these five proposed systems is able to evoke various feeling states—although some systems are more constrained than others. The tension and reaction responses, for example, have a limited range of affective expressions. By contrast, the appraisal response is able to evoke a huge range of feeling states, from jealousy, contempt, or loneliness, to compassion, pride, or humor.[16] For any given situation, these five proposed systems combine to create a distinctive limbic cocktail. Actually, "cocktail" isn't quite the right word, because it is a dynamic phenomenon rather than a simple static mixture. Expectation-related emotions can begin long before an event occurs and can linger long afterward. Within this time span, a dynamically evolving sequence of feelings can arise.

As we will see later, these systems combine to produce a wealth of different feeling experiences in different circumstances. Of all the "practitioners" of emotion, musicians, I believe, have proved the most adept at manipulating the conditions for these different dynamic responses. Although I have used nonmusical examples in this chapter, the principal focus of this book will be on using the ITPRA theory to explain many aspects of musical organization. In chapters 13 to 15 we will see how musicians make use of these psychological systems, and in chapter 16 we will see how the psychology of expectation has shaped a major event in Western music history. But before we apply the ITPRA theory to specific musical circumstances, there are a number of supporting topics that need to be addressed. How does a listener know *what* to expect? How are expectations acquired? Are all expectations learned, or are some innate? How are expectations mentally represented? How are expectations tailored to a particular context? How do the different response systems interact? These and other questions will occupy the next several chapters.

2 Surprise

When I was in sixth grade, we held a surprise party for our teacher, Ms. Bradley. One of my classmate's parents had invited Ms. Bradley to dinner. Hiding in the basement was my entire sixth-grade class. Under the pretense that dinner wasn't yet ready, Ms. Bradley was led down into the darkened room. When the lights came on we all shouted "SURPRISE!" at the top of our lungs. The effect couldn't have been more satisfying to a child's mind: Ms. Bradley nearly jumped out of her skin. There was an expression of sheer terror on her face: her chin had dropped, leaving her mouth wide open, and the whites of her eyes were visible all the way around her pupils. But the expression of terror quickly dissolved into laughter as she regained her composure.

When a surprising event occurs, two brain processes are initiated: a rapid reaction response and a slower appraisal response. Research by Joseph LeDoux and his colleagues at New York University has helped to trace the corresponding pathways through the brain.[1] Figure 2.1 provides a schematic diagram of the reaction and appraisal tracks. Whether the stimulus is visual or auditory in origin, information is first relayed to the thalamus, where some basic processing occurs. At this point the fast and slow tracks are known to diverge. The fast track proceeds directly to the amygdala, which plays an important role in the assignment of affective significance to sensory stimuli. If fear is evoked, activity can be seen in the midbrain periaqueductal gray—which coordinates a number of defense responses. In addition, the paragigantocellularis lateralis will be activated. This is a region of the brain that initiates changes to the viscera that contribute to the sensation of fear. When an event makes you feel something in the pit of your stomach, it is the paragigantocellularis lateralis that is responsible.

The slow track amounts to a long detour between the thalamus and the amygdala. The detour here is via the cerebral cortex—the massive outer layer of the brain, which, among other things, is responsible for conscious thought. Here the stimulus causes activation throughout large areas of the cortex, especially in the frontal lobe region—areas associated with cognitive rumination and evaluation. All sorts of complex social,

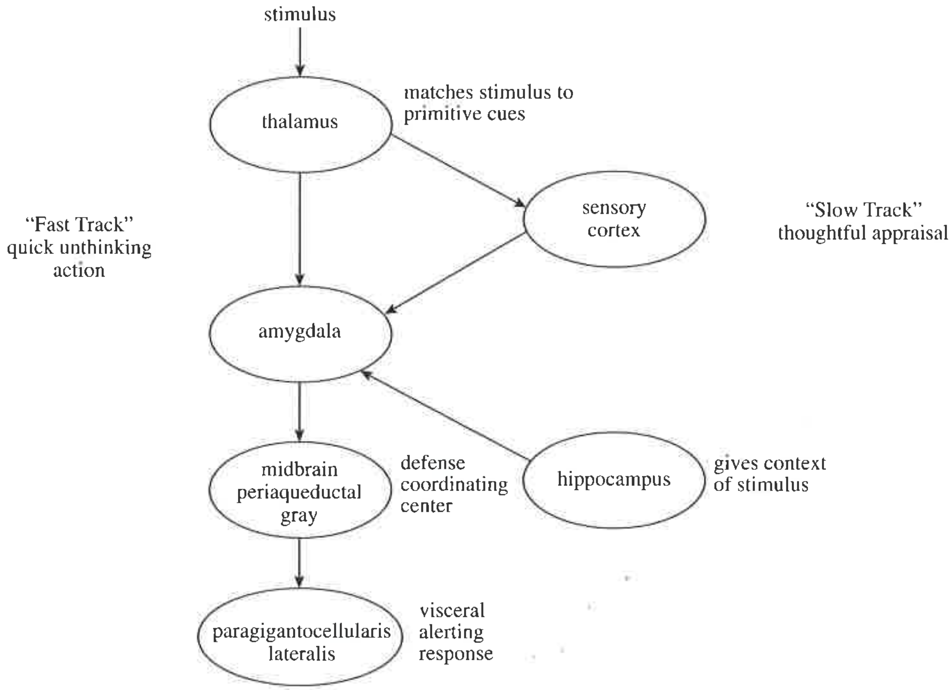

Figure 2.1

Schematic diagram of the brain mechanisms involved in the fear response. Concurrent fast and slow pathways are activated. The fast track involves a direct connection between the thalamus and the amygdala. The slow track leads through the sensory cortex—the large mass located on the exterior surface of the brain.

environmental, and behavioral considerations are brought to bear in appraising the situation. All of this neural activity makes this track the more complicated pathway.

The fast and slow pathways converge in the lateral nucleus of the amygdala. In fact, the two pathways converge onto single neurons in the amygdala.[2] If the slow track leads to the assessment that fear is unwarranted, then the cortex will generate an inhibitory signal that effectively turns off the amygdala and arrests the fear response. Compared with reaction responses, such cognitive processes result in a more accurate, nuanced, and realistic appraisal of dangers, risks, or opportunities afforded by some situation. Unfortunately, this appraisal process is comparatively slow.

For both the fast and slow pathways, the context of the stimulus is known to influence the fear response. Several studies have shown that the context for fear associations depends on the hippocampus. That is, the hippocampus helps us interpret the fear-evoking stimulus.[3]

As noted, the rapid reaction response to surprise assumes the worst. The body responds quickly—under the presumption that the outcome is bad and that the body must belatedly rally resources to deal with an unanticipated situation. When it comes to the unexpected, our physiological reflexes reveal that biology is deeply suspicious and pessimistic: bodies don't like surprises.

One consequence of this rapid reaction response to surprise is activation of the sympathetic nervous system: heart rate increases, and there is increased perspiration, increased glucose uptake, and an increase in the rate and volume of respiration. The purpose of the respiratory response is to bathe the body's tissues in oxygen (and purge carbon dioxide) in preparation for a possibly major energy expenditure. In addition to this increase in arousal, another consequence of surprise is an increase in attention. Sensory systems become more acute. Irrelevant thoughts are purged and attentional resources are focused on the immediate events.

The heightened arousal and attention associated with surprise can be observed directly in the characteristic facial expression for surprise. Recall that Ms. Bradley's chin dropped so that her mouth was wide open, and her eyelids were fully retracted producing a "wide-eyed" look. The open mouth facilitates respiration, which is a component of raising arousal. The open eyes facilitate visual perception, which is a component of increased attentiveness.

Contrastive Valence

Surprising people for fun appears to be a cross-cultural universal.[4] Ronald Simons has documented many instances of "recreational" surprise in various cultures. Most cultures have an equivalent of the "peek-a-boo" interaction between parents and infants, and in nearly all cultures children take delight in sneaking up on each other.

It is not simply the case that surprises are fun for those doing the surprising. In some circumstances (such as Ms. Bradley's surprise party) surprises can also be fun for the person being surprised. But the observation that at least some surprises can be pleasurable raises a biological puzzle. Recall that the purpose of expectation is to enhance readiness. The phenomenon of "surprise" represents a failure of expectation. From a biological perspective, surprise is *always* a bad thing. Even when the surprising outcome turns out to be good, failing to anticipate the outcome means that the brain has failed to provide useful information about possible futures. Predictive failures are therefore cause for biological alarm. If an animal is to be prepared for the future, the best surprise is no surprise.

So if surprise is biologically bad, how is it possible for some surprises to activate the physiological machinery for pleasure? How can people possibly enjoy being surprised? It turns out that an answer to this question can be found on any basketball court. Consider the following experiment carried out by psychologist Peter McGraw.[5] McGraw asked amateur basketball players to take shots from different locations around the court. Before each shot, the player was asked to estimate the likelihood of scoring a basket. Following each shot, the player rated how good he or she felt. As you would expect, players are happiest when they successfully make a shot and are unhappy when they miss a shot. Emotions are positively valenced for positively appraised outcomes and negatively valenced for negatively appraised outcomes. However, the degree of satisfaction or dissatisfaction is also related to the player's expectation. McGraw found that basketball players experienced the greatest unhappiness when they missed shots judged easy, and were happiest when they scored baskets judged to be difficult. The magnitude of the emotional response is amplified when there is a large contrast between predicted and actual outcome. In general, *unexpected* fortune or misfortune causes the biggest emotional responses.[6] That is, low expectation amplifies the emotional response to the outcome.

When we are surprised, a limbic contrast will sometimes arise between the reaction response and the ensuing appraisal response. The fast biological response to surprise is perpetually negative. But the slower appraisal response might be neutral or even positive. At Ms. Bradley's surprise party, the momentary terror she experienced was quickly displaced by a highly positive appraisal response. Her amygdala was switched on and then abruptly switched off. As with McGraw's basketball players, it was the limbic contrast that rendered the experience so powerful. In less than a second, Ms. Bradley's brain went from experiencing profound terror to happy celebration.

In the case of McGraw's basketball players, the limbic contrast was between the prediction response and the appraisal response. In the case of Ms. Bradley's surprise party, there was similarly a contrast between the prediction and appraisal responses, but the main limbic contrast was between the (very negative) reaction response and the (very positive) appraisal response.

Limbic contrasts between the fast and slow pathways are not limited to the extreme circumstances of a surprise party. In our daily lives we experience hundreds of small moments of surprise: a telephone rings, a pen runs out of ink, the car ahead changes lanes, a petal falls from a cut flower. Many of these episodes fail to reach conscious awareness. Most are ultimately appraised as innocuous.

The interaction of the fast and slow pathways has repercussions for how listeners experience sound. If a nominally unpleasant sound is not expected by a listener, then the sound will be perceived as even more unpleasant or annoying. Conversely, if a nominally pleasant sound is not expected by a listener, it will tend to be perceived as more pleasant. A lengthy dissonant passage is likely to lead listeners to expect further dissonant sonorities. If the music shifts toward a more consonant texture, then the resulting contrast will tend to evoke a pleasing effect that can be greater than experiencing only the consonant passage.

More interestingly, even neutrally appraised sounds have the capactiy to generate limbic contrast. Whether or not a sound is regarded as inherently pleasant or unpleasant, if it is unexpected, it is capable of evoking a negatively valenced prediction response—which may contrast with an ensuing neutral appraisal.

An apparent problem for the contrastive valence theory is accounting for why an unexpected good outcome would be experienced as more positively valenced than an expected good outcome. In the case of an expected good outcome, a person should feel good about the outcome and should also feel good about the predictive accuracy. In the case of an unexpected good outcome, a person should feel good about the outcome but also feel bad about the poor predictive accuracy. Why then does an unexpected good outcome sometimes evoke a more positive emotion than an expected good outcome? It is possible that the contrast between the positive appraisal and negative reaction is the more powerful effect. But if so, it would be maladaptive to consistently experience a positive emotion whenever one's predictions prove wrong. Perhaps the only important lessons for an organism to learn occur when a wrong prediction accompanies a negative appraisal of the outcome.

A better understanding of the phenomenon comes from looking at some of the physiological research. When we experience stress, it is typically accompanied by the release of endogenous opiates such as endorphins. These natural opiates produce an analgesic effect that reduces the experience of pain. You've probably experienced a panic situation, such as where you've had to run to catch a bus. Only after you are settled in the bus do you realize that you inadvertently cut yourself in your frantic rush. Oddly, the pain is much more subdued than if the same injury had been inflicted while at rest. That is, the energetic panic had somehow attenuated the experience of pain. Working at Stanford University, Albert Bandura, Delia Cioffi, Barr Taylor, and Mary Brouillard carried out an experiment where individuals were artificially induced to fail at some task. Half of these individuals were administered a saline solution while the other half received naloxone—an opiate receptor antagonist that blocks the effect of any endorphins present. The participants were then given a test that measured their tolerance for pain. Those individuals who had received naloxone were significantly less pain tolerant, indicating that the stress caused by failing at the task had resulted in a release of endogenous opiates.[7]

The body appears to release opiates whenever we experience a negatively valenced emotion such as pain or fear. Suppose that a wild bear appeared and injured you. Immediately, analgesic opiates are released by the body to counteract the pain and allow you to continue to function. Now, suppose that it turned out that you weren't actually injured, but the opiates were released anyway. The net result is simply the opiate release—and the ensuing pleasant feelings. The physiological origin of contrastive valence might follow a path akin to this sequence of events.

The potential for increased pleasure provided by contrastive valence might explain a number of seemingly peculiar conscious strategies people use when mentally preparing for future outcomes. Recall that the imagination response occurs in the pre-outcome epoch—prior to the outcome event. Normally, the imagination response serves a motivational function. By imagining different outcomes (and experiencing a foretaste of the feelings associated with these outcomes) we are encouraged to take actions that will reduce future negative outcomes and increase the likelihood of future positive outcomes. Even for the most certain positive outcomes, if time permits, we will have an opportunity to consider the possibility that the positive outcome will not happen. The imagination response for such scenarios is sobering. We realize that even at the cost of reducing the predictive accuracy, it is better to lower the subjective probability of a positive outcome in order to mute the bad emotions that might ensue from an unexpected negative outcome.

Support for this view comes from a common mental strategy. When a good outcome seems highly probable, a person will sometimes mentally deny or minimize the likelihood of the positive result. ("Everyone seems certain that Caroline will agree to marry me, but I think there is a real chance that she will say no.") This conscious mental strategy has the effect of reducing the magnitude of the negative feelings should a negative outcome ensue, while simultaneously increasing the pleasure evoked by a positive outcome. Similarly, people will occasionally exaggerate negative predictions. ("My coat is ruined; I bet that stain will *never* come out.")

This view is reinforced by the reverse scenarios—found in the experience of *Schadenfreude*—pleasure in others' misfortune. When a rival seems poised for failure, we may minimize the likelihood of their negative outcome ("Aaron is certain to escape punishment for what he did"). This conscious mental strategy will minimize the disappointment if our rival escapes the negative outcome, and increase the pleasurable *Schadenfreude* should the negative outcome be realized. Conversely, when a rival seems poised for success, we may amplify the likelihood of their success ("Barbara is certain to inherit the house"). This will maximize the *Schadenfreude* if the rival in fact does not succeed, and increase the positively valenced prediction response if the outcome for our rival is positive ("I absolutely *knew* that Barbara was going to get the house").

Notice that these strategies are necessarily conscious. They supplant the underlying unconscious intuitions about the likelihood of various events and intentionally bias the subjective probabilities in order to maximize positively valenced emotional states. Since these strategies require conscious rumination, a certain amount of time is needed in order for a person to formulate the strategy. When events unfold quickly, there is insufficient time to arrange these strategies. By way of example, consider once again the case of the basketball players predicting the success of sinking baskets. The theory of contrastive valence can be used to make the following prediction. After a basketball player has predicted the likelihood of making a shot, suppose we stop the player from shooting. We then ask the player to pause and reflect again on the shot: "With a moment's thought here, would you like to consider revising your estimate of the likelihood of making this shot?" My prediction is that the players will tend to reduce the probability of making the shot. That is, in general, the longer the time interval for contemplation, the lower the assessed probability of a positive outcome. Moreover, if a large (monetary) reward is offered for completing a shot, I predict that the effect will be magnified: the larger the monetary reward, the greater the likelihood that the assessed probability will drop below the objective empirical likelihood of making the shot. That is, people will tend to adopt an overtly pessimistic attitude in order to minimize possible disappointment and maximize future pleasure.

All of this presumes a "contrastive valence" theory of the magnitude of positive and negative emotions. How we ultimately feel about an event is not simply tied to an appraisal of some absolute benefit or penalty associated with that event. Our feelings also seem to depend on limbic contrast.[8] This is not a new insight. In his *Essay on Human Understanding*, John Locke spoke of the pleasure arising from the removal or attenuation of pain. Edmund Burke explicitly argued that the aesthetic experience of the "sublime" depends on an initial sensation of fear that is ultimately appraised as inconsequential. Nor do these ideas originate with these eighteenth-century European philosophers. Throughout history, sages have recognized that pleasure is enhanced by contrast: happiness is not so much a state of being as it is a state of becoming. People who fast periodically probably benefit from contrastive valence. As the Spartans understood, hunger is the best spice.

Three Flavors of Surprise

Listening to music can give rise to an enormous range of emotions.[9] Music can engender a joyous exuberance or transport us into a deep sadness. It can evoke a calm serenity or generate spine-tingling chills. It can lead to a sense of ominous darkness or convey a mysterious sense of awe and wonder. Music can even cause listeners to laugh out loud.

Apart from these strong "whiz-bang" emotions, music is also able to generate a wealth of "microemotions"—more subdued or muted feelings that are not easy to describe but occur more frequently while listening to music. These more nuanced feelings will be considered in later chapters. For the remainder of this chapter I would like to focus on three strong emotions that are closely linked to surprise: *laughter*, *awe*, and *frisson*. *Laughter* is a state of jocular amusement that is characterized by a distinctive "ha-ha-ha" vocalization. *Awe* is a state of astonishment or wonder that may cause a person to gasp. *Frisson* is the feeling you have when "chills" run up and down your spine and the hair stands up on the back of your neck. (All of these responses will be described more carefully later.)

These strong emotions are not common while listening to music, but when they occur they are memorable. Such emotions are typically sparked by particular musical moments or passages. Of course, not all listeners will respond to these passages in the same way, and even a single individual listener may not have the same emotional experience each time the passage is encountered. When they are evoked, however, these emotions are distinctly pleasurable and so listeners often seek out these experiences. Although *laughter*, *awe*, and *frisson* appear to be very different from one another, I will suggest that they share a deep biological kinship. We will see that each of these emotions is related to a violation of expectation. All three are specialized varieties of surprise.

Extreme surprises lead to the characteristic "surprise" facial expression. Interestingly, the "surprise" expression is the same as the one for horror.[10] The common element is simple fear. When in a fearful state, an animal has recourse to one of three behaviors. One response is to stand one's ground and attempt to defeat, disarm, or neutralize the danger; the goal here is to prevail over the fear-inducing situation. Another response is to run away as quickly as possible; the goal here is to escape the fear-inducing situation. A third response is to hide by remaining perfectly immobile and quiet; the goal here is to escape notice while the fear-inducing situation passes. These three primordial behaviors of fight, flight, and freeze can be observed throughout the animal kingdom. Even insects respond to danger by generating one of these three behaviors.

In everyday life, these responses are rarely manifested as full-blown behaviors. Instead, they are occasionally glimpsed as momentary or fleeting states. In the case of the surprise party, the person being surprised will often exhibit the characteristic "surprise" face with the gaping mouth and wide-open eyes. The "surprised" facial expression can pass in the blink of an eye. It is sometimes so quick, that it is apparent only later when viewed on video. The surprised expression is soon replaced by smiles and laughter as the celebrated individual realizes the true meaning of the unanticipated event. But in that first brief moment of fear lies much that is important.

Laughter—A Pleasurable Panting

In chapter 14 we will see examples of musical passages that cause listeners to laugh out loud. People don't often laugh while listening to music, but musically induced laughter does occur from time to time. Occasionally, composers explicitly set out to create works that provoke laughter—such as Mozart's *Ein musikalischer Spass* ("A Musical Joke"), Haydn's string quartet opus 33, no. 2 ("The Joke"), or Peter Schickele's wacky *Pervertimento for Bagpipes, Bicycle and Balloons*. In chapter 14 we will see evidence that whenever music causes listeners to laugh, the response can be traced to a massive violation of expectation. But for now, let's simply accept my claim that musically induced laughter is one of the responses a listener can experience when surprised.

What is it about some violations of expectation that will cause listeners to laugh? A tempting answer is to say that the violations are "humorous." But this begs the question. Consider a more literal version of our question: Why do people sometimes make a "ha-ha-ha" sound in response to violated expectations? Why don't listeners make a "hissing" sound, grit their teeth, or tug on their earlobes? What is it about violated expectations that can evoke the distinctive "ha-ha-ha" vocalization we call laughter?

Scientific studies on laughter can help answer this question—but the story is a little circuitous.[11] Perhaps the single most important lesson from research on laughing is that laughter is predominantly a *social* response. Robert Provine estimates that people are thirty times more likely to laugh in the presence of another person than when they are alone.[12] Field studies have established that most laughter is not in response to humor. Social inferiors laugh more in the presence of their social superiors. Social inferiors are also more likely to laugh at the instigation of social superiors than vice versa. For those of us who love to laugh, the scientific research on laughter seems depressing: the principal function of laughter seems to be to dissipate social fears. Laughter is a common response to humor, but humor is not necessary for laughter. For example, people frequently laugh at the misfortunes of their enemies, or laugh nervously in the face of danger.

Although laughter is predominantly a social phenomenon, the laughter response itself is innate rather than learned. All over the world, people laugh with a characteristic "ha-ha-ha . . ."[13] Even congenitally deaf people who have never heard anyone laugh make the same sound. Of course there are many variations on a theme. Some people make little sound—simply allowing their chest to jiggle in paroxysms of laughter ("chuckling"). Other people squeal, bellow, or roar. Yet others exhibit all manner of giggling, yukking, snickering, and snorting. Despite these variations, the basic laughter pattern remains.

Provine describes laughter as a "punctuated exhaling." The onset-to-onset time between "ha's" occurs at regular intervals of about 210 milliseconds. If the vocalization is high pitched and quiet, the sound is regarded as "giggling." If the mouth remains closed during laughing, the sound is better described as "snickering." If the mouth opens and closes in synchrony with the "ha's" then the sound becomes one of "yukking." If the mouth is rounded and remains open, then the sound is modified to a distinctly Santa-like "ho-ho-ho." Each laugh-phrase is followed by a deep inhalation. If the inhalation relies on the nose alone, then sometimes a snorelike vibration of the soft palate will cause a distinctive "snorting" sound (often to the embarrassment of the person laughing). Sometimes the laughter is preceded by a high-pitched squeal.

In general, however, the basic laughter pattern consists of a punctuated exhaling that occurs at a rate of roughly five times per second.[14]

Humans are not the only animals that laugh. Recognizably laughlike responses are also observed in chimpanzees, bonobos, orangutans, and gorillas. For example, tickling a young chimpanzee will typically cause the chimp to produce a rapid "oo-oo-oo-oo" sound. Similar sounds are produced during rough-and-tumble play such as when an animal is being chased by a playmate. Interestingly, the rate of such punctuated breathing is comparable to the rate found in human laughter. However, the laughter of chimpanzees, bonobos, orangutans, and gorillas differs in one important respect from that of humans. Humans laugh by exhaling only. The "oo-oo-oo-oo" sound of nonhuman primates involves a rapid alternation of inhaling and exhaling that doubles the tempo of the vocalizations to about ten per second.[15]

Robert Provine has argued that the human exhale-only version of laughter is a recent evolutionary development since the inhale–exhale version is shared by the other primates. Says Provine, "The human variant with its looser respiratory-vocal coupling evolved sometime after we branched from chimpanzees about six million years ago. Evidence of the primacy of the chimp form comes from the identification of chimplike laughter in orangutans and gorillas, apes that split off from the chimpanzee/human line several million years before chimpanzees and humans diverged."[16]

The inhale–exhale laughter of chimpanzees, bonobos, orangutans, and gorillas is essentially a form of *panting*. In fact, when Provine played ape recordings to university students and asked them to identify the sounds, the most common response was to describe the sounds as "panting." Going beyond Provine's research, we might observe that panting is a common component of physiological arousal: panting prepares an animal for physical exertion by drawing in large quantitites of oxygen and expelling carbon dioxide. When an animal fails to anticipate an event, this failure represents a potential danger. Like increased heart rate and increased perspiration, panting is a wholly appropriate response to surprise.

There are different types of laughter. *Nervous laughter* tends to occur when we are aware of an impending threat or danger. There is *slapstick laughter* where we laugh at physically awkward or calamitous body movements of others (or sometimes ourselves). *Sadistic laughter* occurs when we laugh at the misfortunes of our enemies. If our enemy is present, we may direct the response *at* the person as *mocking laughter*. *Surprise laughter* tends to occur when an event (such as a bursting balloon) is completely unexpected. *Social laughter* is the sort of polite or gregarious laughter that signals our participation or desired membership in a particular social group. In social laughter, we may be aware that nothing is particularly funny, but we may still laugh as a way of reducing any social tension. Finally, *humor laughter* is an overt form of pleasurable entertainment, usually relying on some form of joke-telling. Humor does not require laughter (we can find things amusing without laughing). Nor does laughter necessarily indicate humor.

These different forms of laughter would seem to defy any single explanation. What is it that all of these situations share? Why do nervousness, surprise, sadism, slapstick, humor, and social politeness all tend to lead to a characteristic "ha-ha-ha" respiratory reflex? I think that Aristotle already had the answer 2,500 years ago. In his *Poetics*, Aristotle described laughter as "a species of the base or ugly . . . some blunder or ugliness that does not cause pain or disaster."[17]

Aristotle's view was reiterated by the eighteenth-century German philosopher Immanuel Kant when he characterized laughter as arising from "the sudden transformation of a strained expectation into nothing." The key here is the contrast between the fast *reaction response* and the slower *appraisal response.*

The different forms of laughter all share two features: (1) Each situation is tinged with risk or fear, although this fear may arise simply because of a failure to anticipate the future (i.e., surprise). Remember that the reaction response always assumes the worst in surprising situations. (2) However, cognitive reflection either eliminates or inhibits the assessment of fear or risk. In the case of slapstick, the fast reaction response arises from *empathy*—the brain's disposition to "mirror" the psychological experiences of others. For example, research by Tania Singer and her colleagues at University College, London, has shown that similar brain regions are activated when we see someone else experience pain as when we ourselves experience pain.[18] The fast reaction response responds to the physical danger (whether to ourselves or others), but the ensuing cognitive appraisal recasts the apparent danger as inconsequential. The same pattern characterizes *Schadenfreude* when we laugh sadistically at the misfortunes of those we exclude from our social in-group: these are the misfortunes of *others*—not ourselves. Once again, there is a rapid recognition of the existence of harm, followed by an appraisal that we ourselves are not endangered—we might even benefit from the situation. In the case of nervous and social laughter, these rely on a cognitive appraisal that a social superior will not mistreat us, for example.

Consider the simple game of "peek-a-boo." The game is surprisingly resilient. Pediatrician Dr. Robert Marion describes how his sympathy for a distressed infant led to a long game of peek-a-boo: "I covered my face with a hospital towel and pulled that towel away for nearly three hours, to the delighted squeals of this one-year-old."[19] How is it that the sudden appearance of a face can cause such delight? Once again, it is the combination of "innocent surprise." The surprise causes a fast biological alarm, followed by the appraisal that no harm will happen. But how is it possible to sustain such a response for nearly three hours?[20]

There are excellent reasons why the fast-track brain should not lower its guard. The sole purpose of these reaction responses is to defend against those rare circumstances that are truly dangerous. The whole fast-track system is designed to tolerate large numbers of false alarms. Unlike the townsfolk in Aesop's fable, the fast-track pathway tends to respond to the cry "wolf" no matter how many times it occurs. The fast-track system maintains its vigilance. It never grows tired of assuming the worst. If the fast-track system did adapt or habituate to all the false alarms, then the reaction responses would not be available for those rare occasions when they are really needed.

Musical surprises fall almost exclusively into the category of innocuous risks.[21] In generating musical humor, composers are taking advantage of the biology of pessimism. As we have noted, surprise is always a sign of adaptive failure, and so the initial limbic response is necessarily negative. However, the slower appraisal mechanism intercedes and the ensuing contrastive valence results in a broadly pleasant experience.

Laughter's Origin

If laughter is an innate reflex, then laughter must originate as an evolutionary adaptation with its own unique biological history of development. The idea that laughter originated as a type of fear response does not seem to square well with our human experience of humor. Laughing is fun. How is it possible that a species of fear became transmuted into something pleasant? If I might be permitted some evolutionary speculation, I would like to suggest that the development of human laughter might have evolved along the following path:

1. Unvocalized panting occurs in response to surprise. This panting is part of a generalized increase in physiological arousal. Like the gaping mouth of the "surprised" face, panting prepares the animal for action.
2. For highly social animals (like humans and great apes) the biggest dangers come from other members of our species (conspecifics).
3. Threats from other conspecifics also evoke unvocalized panting. The threatening animal recognizes the panting as a successful provoking of a momentary state of fear in the threatened animal. The evoked fear means that the threatened animal has been successfully cowed. Panting becomes a signal of social deference. As an aside here, we might note that dogs exhibit "social panting"—where submissive animals begin panting when a dominant animal (sometimes a human owner) appears.[22]
4. Being able to evoke panting in another animal reassures the dominant animal of its dominant status. Similarly, panting in the presence of another animal serves to communicate one's submissiveness. For both animals, this communication is valuable because it establishes the social hierarchy.

5. In order to enhance the communication of deference, panting becomes *vocalized*—that is, the vocal cords are activated. Vocalized panting becomes a specific signal of deference or submissiveness.
6. Vocalized panting generalizes to most surprising circumstances.
7. In some primates, "panting-laughter" is reserved specifically for surprise linked to nondangerous outcomes. In highly socialized animals, most dangers are social in origin, so panting-laughter is commonly associated with social interaction.
8. In hominids, panting-laughter becomes explicitly social. Mutual panting-laughter within a group becomes an important signal of reciprocal alliance, social cohesion, and peaceful social relations.
9. The contrast between negative reaction feelings and neutral/positive appraisal feelings evokes an especially pleasant state. Human culture expands on these agreeable feelings through the advent of "humor" as an intentional activity meant simply to evoke laughter.[23]
10. Laughter becomes commonplace in hominid social interaction. In order to reduce the energy cost of laughter, the inhaling–exhaling form is replaced by the more efficient vocalized exhaling (i.e., modern human laughter).

Evolutionary storytelling is fraught with dangers, so the above story should be viewed pretty skeptically. My story is offered only as an illustration of how the innate behavior we call laughter might have evolved from surprise-induced fear. Biology is a complicated business, and most physiological functions develop in strikingly convoluted ways. The actual origin of human laughter is probably much more complicated. But as a universal human reflex, there is no doubt that laughter has an evolutionary history.

Awe—Takes Your Breath Away

Laughter is not the only response that can be evoked by surprise. Imagine that you are making your way through a dense jungle where tangled branches render your progress fitful. The thick vegetation obscures your vision so you can only see a foot or two ahead. Suddenly, you break through a mass of vegetation and find yourself standing immediately on the edge of a high cliff. A few more steps and you would plunge over the precipice. While laughter might be a possible response to this situation, a more likely response is to gasp. A *gasp* is an abrupt inhaling followed by a momentary holding of one's breath.

This imaginary scenario differs from the usual fight-or-flight response. There is no need for an impending expenditure of energy. Instead, the danger is more suited to the often-overlooked *freeze* response. Standing at the cliff's edge, *panting* would serve no useful purpose. But freezing *is* an appropriate response. A single rapid inhale provides a reservoir of oxygen that allows us to hold our breath.

Other situations that might evoke a freeze response are easy to imagine. You might encounter a snake, or a sparking electrical cable, or someone pointing a gun. In Brazil, I once encountered a spider's web the size of my outstretched arms and legs—it stopped me dead in my tracks, and I distinctly recall holding my breath. The enormous web was beautiful—brilliantly illuminated in the sunshine. But it was simultaneously a fearsome sight.

Holding one's breath has several benefits. It reduces movement and sound, which makes it more difficult for a possible predator to see or hear us. Reducing movement and sound also makes it easier for us to listen intently, and to see more clearly. Marksmen are better able to hit the bull's-eye when they stop breathing and shoot between heartbeats. A cricket will merrily chirp away until approached. Then it will become silent until it assumes you have moved away.

The freeze response is most probable when the danger remains fixed. The danger associated with the cliff remains as long as we are near the edge. The danger associated with encountering a snake remains as long as the snake is nearby. Laughter, by contrast, is more likely to occur when an apparent or actual danger rapidly dissolves. The bursting of a balloon represents a momentary rather than sustained danger. I might slip on a staircase, but immediately recover my footing without falling—an event that then results in my laughing. A parent says "boo" to a child, who subsequently erupts in laughter.

In short, laughter is a response to an apparent or momentary danger, whereas the gasp is a response to a sustained danger. Following the gasp, subsequent appraisal of the situation will determine whether the danger is real or "manageable." If the danger is real, then the gasp will be a prelude to sustained fear. If the danger is manageable, then the gasp will be a prelude to *awe*. Wobbling on the edge of a cliff we may experience fear. Standing securely on the edge of a cliff we may experience awe. Being in close proximity to a venomous snake we may experience fear. If the snake is being held by an experienced snake handler then we may experience awe.

Like laughter, the gasp can be vocalized or unvocalized. The unvocalized rapid inhale is the quiet, clandestine form. Silence is an appropriate response to sustained danger. The vocalized "ah!" when we gasp is more unusual, because it is rare for humans to vocalize while inhaling. Vocalizing also makes it more difficult to remain unobserved. But the vocalized form of the gasp is more audible to others and so may have some communicative function. In short, the unvocalized gasp is a self-serving defense response, whereas the vocalized gasp is a more altruistic response that can alert others to potential danger.[24]

As in the case of laughter, the gasp can be evoked by stimuli that at first appear dangerous, but on reflection are recognized as not actually dangerous. That is, the reaction response is acutely negative, but the appraisal response is either neutral or positive. It is this contrast that transforms the experience into something pleasurable.

Accordingly, we might characterize awe as a form of pleasurable surprise, one that mixes a sense of apparent sustained danger with an appraisal that the situation is okay or good. This characterization of awe is reinforced by the word's etymology. In English, the word "awe" traces a circuitous historical path that nicely echoes the combination of positive and negative emotions. My Webster's dictionary reports the archaic meaning of the word "awe" as "the power to inspire dread." The word originates in the Greek *achos,* which means pain. Webster's defines the more modern meaning of "awe" as follows:

> emotion in which dread, veneration, and wonder are variously mingled: as 1: fearful reverence inspired by deity or by something sacred or mysterious 2: submissive and admiring fear inspired by authority or power (he stood in ~ of the king) 3: wondering reverence tinged with fear inspired by the sublime

Standing on the edge of the Grand Canyon is both dangerous and a wonderful opportunity to get a good look around. The sight is surely awesome—it takes your breath away. If God is to be both loved and feared, then meeting God would be a good reason for a person to feel awe.

Frisson—Thrills from Chills

Fear can provoke many different types of physiological reaction. When threatened by another animal, we might prepare to *fight* rather than flee or freeze. In the fight response, the first order of business is to produce an aggressive display. By signaling one's readiness to fight, it is possible that the threatening individual might back down, and so an actual fight can be avoided. Aggressive displays can include the displaying of teeth, making eye contact with the other animal, and generating low-pitched vocalizations. In addition, there are a series of behaviors that are all intended to make the individual appear bigger—and so more intimidating. This includes rearing up (to appear taller) and the bristling of hair (to appear more massive). Cats, by way of example, will arch their backs and make their hair stand on end.

All of these responses are also evident in humans. For example, human aggression can be signaled in the wrinkling of the upper lip, which, if the mouth is open, makes the upper cuspids (or canine teeth) more visible. Although humans are comparatively hairless as mammals go, fear can still cause the hair on the back of your neck to stand on end—a reflex technically known as *piloerection.*

Piloerection also occurs when a person is cold. Shivering causes the skin to tighten and hair follicles to pucker—causing the characteristic "goose flesh" texture. When it is cold, piloerection helps capture a layer of insulating air—a warming technique that works a lot better for animals with more hair than us humans. Even when piloerection occurs in response to fear (rather than cold), we still have a feeling which is often described as "chills." I think that this feeling is another example of the comparatively rare emotions that conform to the James–Lange theory. Recall from chapter 1 that James and Lange argued that the physiological response itself can precede and lead to a characteristic feeling. In chapter 1 we saw an example of such feelings in the Strack pencil task—where making people flex their cheeks (in a smile-like manner) causes them to feel happier. Here "chills" are evoked by piloerection. Fear doesn't make us "cold." Instead, fear causes piloerection, and this response evokes the phenomenal feeling of coldness—a feeling that would be a normal sensation when piloerection is evoked by cold temperatures.

This raises the question of why the linkage is not the other way around. Why is it that when we experience fear, we feel chills, but when we are cold, we don't feel fear? I suspect that the reason for this asymmetry harkens back to the very distant evolutionary past. Piloerection almost certainly arose first as a method of thermoregulation. For animals with lots of hair, making one's hair stand on end is a good response to a cold environment. Later, natural selection "discovered" that piloerection is a welcome addition for creating a convincing aggressive display (a physiological borrowing that biologists call a "preadaptation" or "exaptation"). Once the neural wiring was added so that aggression generated piloerection, the phenomenal sensation of chills simply came along as an artifact.

The phenomenon of "chills running up and down your spine" is technically referred to as *frisson*—a useful loan-word from French. The phenomenon has been variously described, including "thrills," "shivers," "chills," and "goose flesh." Psychophysiologists Gunther Bernatzky and Jaak Panksepp have proposed the memorable term "skin orgasm"—but their term has failed to gain currency.[25] A handful of studies regarding musically evoked frisson have been carried out, including work by Jaak Panksepp, Robert Zatorre, Anne Blood, Richard Gray, Avram Goldstein, and John Sloboda.[26] In chapter 14 we will examine a number of examples of musical passages that evoke frisson. We will see that the frisson response is correlated with two conditions: (1) loud passages, and (2) passages that contain some violation of expectation—such as an abrupt modulation.

Loudness is known to increase physiological arousal. There are good reasons for this connection: loudness is indicative of events in the environment that entail a large expenditure of physical energy. Whether physical energy is embodied in animate agents (such as a herd of elephants) or in inanimate objects (like boulders rolling down a slope), high levels of physical energy are more likely to pose a danger than low levels of energy. There are good reasons for organisms to be highly aroused by loud sounds.

In my own listening experience, I have found that I can reliably manipulate the magnitude of a frisson episode by adjusting the playback volume on my stereo. Louder reproduction enhances the frisson evoked by a musical passage, while quieter reproduction reduces the frisson. Incidentally, I've also observed that frisson is influenced by temperature: I am less likely to experience frisson when I am warm or hot. Cinemas and concert halls with lots of air conditioning might well enhance the emotional experiences of patrons.

Along with loudness, frisson is more likely to occur when there is an unexpected modulation, or an abrupt chromatic mediant chord, or an unexpected onset. In Beethoven's "Ode to Joy," many listeners have pointed to the pleasure evoked by a memorable moment when one of the phrases abruptly begins half a beat early.

As we have already noted, surprise is biologically bad. Both loudness and surprise evoke negatively valenced *reaction* responses. The loudness is symptomatic of high energy with a concomitant potential for physical injury. Surprise is symptomatic of an unpredictable environment. Together, these factors represent the potential for significant biological danger.

As in the case of laughter and awe, the frisson response originates in a reaction response shaped by fear. The fast-track brain responds to the combination of loudness and surprise with its usual pessimistic presumption. At the same time, the slower appraising mind concludes that the musical sounds are entirely safe. Once again, the negatively valenced piloerection response is in stark contrast with the neutral or positively valenced appraisal response. Once again, the magnitude of this contrast amplifies an overall sense of pleasure.

Fight, Flight, and Freeze: The Aesthetics of Pessimism

As noted earlier, physiologists have identified three classic responses to danger: the *fight*, *flight*, and *freeze* responses. The fight response begins with aggressive posturing and threat displays. The flight response is characterized by a quick increase in arousal, including rapid preparatory respiration. The freeze response is characterized by sudden motor immobility, including breath-holding.

My idea should by now be obvious. There is a striking similarity between the fight, flight, and freeze responses, and the experiences of frisson, laughter, and awe. The piloerection characteristic of frisson suggests a kinship with the fight response. The modified panting of laughter suggests a kinship with the flight response.[27] And the rapid inspiration and breath-holding of awe suggest a kinship with the freeze response.

When musicians create sounds that evoke laughter, awe, or frisson, they are, I believe, exploiting the biology of pessimism. The fast-track brain always interprets surprise as bad. The uncertainty attending surprise is sufficient cause to be fearful (at least until the more thorough appraisal process can properly evaluate the situation). Depending on the specific circumstances, that fear is expressed as one of the three primordial behaviors of fight, flight, or freeze.[28]

Because the fast-track brain never lowers its guard, musicians can rely on sounds to evoke pretty much the same response each time the music is heard. If the fast-track brain weren't so pig-headed in its pessimistic interpretation of surprise, then familiar musical works would rapidly lose their power to evoke the emotions of frisson, laughter, or awe. Of course, listening does change with exposure. But the fast-track brain responds primarily based on schematic expectations, and these schemas change only with extensive exposure. (I'll have more to say about this in later chapters.)

It might seem odd that the experiences of frisson, laughter, and awe rely on the evocation of fear. But this fear appears and disappears with great rapidity and does not involve conscious awareness. The appraisal response follows quickly on the heels of these reaction responses, and the neutral or positive appraisal quickly extinguishes the initial negative reaction. As listeners, we are left with the contrast in valence between the reaction/prediction and appraisal responses—a favorable contrast that leaves us with the sort of warm glow that contributes significantly to the attractiveness of music. In effect, when music evokes one of these strong emotions, the brain is simply realizing that the situation is very much better than first impressions might suggest. In this regard, music is similar to other forms of pleasurable risk-taking, such as hang gliding, skydiving, riding roller coasters, or eating chili peppers.[29]

The truly remarkable thing is that these powerful emotional responses can be evoked through the innocent medium of mere sounds. Of course, not just any sounds will do. Listeners must be enculturated into specific auditory environments where some events or patterns are more predictable than others. As we will see in the ensuing chapters, it is the learned schemas that provide the templates that enable the fast-track brain to make predictions, and in some cases, to be surprised.

If a musician wishes to evoke the experience of laughter, awe, or frisson, then the musician must be intimately familiar with the normative expectations of ordinary listeners. This is not a novel observation. Music scholars have long noted the importance of *convention* as a basis for generating various emotional responses. For example, in her book, *Haydn's Ingenious Jesting with Art*, musicologist Gretchen Wheelock describes many of the devices and elements in Haydn's music that relate to humor. Wheelock quite rightly emphasizes the necessity of convention to humor. Humor requires surprise; surprise requires an expected outcome; and an expected outcome requires an internalized norm. Composers must activate either normative schemas (such as styles) or commonplace clichés in their listeners if their violations of expectation are to have the desired effect. Leonard Meyer recognized this half a century ago.[30]

In each of laughter, awe, and frisson, an initially negative reactive response has been followed by a neutral or positive appraisal. What about the possible situation where an initially negative reaction response is followed by a *negative* appraisal? That is, what would happen (musically), if fight, flight, or freeze were allowed to be expressed unimpeded? Negative appraisals can arise for all sorts of reasons. A listener might find the style distasteful, regard the work as overvalued, dislike the musicians, or have unhappy past associations. A listener might be disappointed by the content of the lyrics, be embarrassed by the amateurism, be offended by an apparent plagiarism, or be disgusted by crass commercialism. But the sounds themselves will inevitably lead to the conclusion that the stimulus is "just music." Unlike the growling of a bear, the sounds do not represent an imminent danger. There is no need to run, flee, or hide from these sounds.

Of course it is possible to imagine fantastic scenarios where "just music" might be truly terrifying. For example, suppose you had just watched a horror film at the cinema with your friends, and then returned home alone. As you step into your darkened house, music from the horror film comes blaring at you from your stereo. (Perhaps you have inventive friends whose notion of "fun" suggests some need for psychological counseling.) In such a circumstance, most people would genuinely flee. But the ensuing flight would not really be evoked by the music. It would be evoked by the appraisal that either an intruder must be in your home, or that there really are evil spirits out to get you. These are both good reasons to feel genuine fear.

The "just music" assessment might explain why, of laughter, awe, and frisson, it is frisson that is the more common experience for listeners. An organism is most in command of a fearful situation when it chooses to fight, rather than flee or hide. That is, if the fear-inducing situation proves to be manageable, then one would expect an aggression display, rather than gasping or panting. The least fearful reaction response would generate frisson rather than laughter or awe.[31]

When surprised, how does a brain decide whether to initiate a fight, flight, or freeze response? Even though no conscious thought is entailed, some assessment must be involved—however crude the judgment. In general, if the danger is assessed to be relatively mild or manageable, then a fight response should be more likely. If the danger is assessed to be sustained, then the freeze response should be more likely. If the danger is assessed as an intermediate threat, then fleeing might be more likely. Almost certainly, context will play a role. When gathering with friends, for example, surprises are probably more likely to evoke laughter than awe or frisson. When alone in a dark alley, by comparison, the same surprise is not likely to result in laughter.

Interestingly, these same experiences of laughter, frisson, and awe can be evoked by purely intellectual pursuits—as when attempting to solve a problem. For example, consider the intellectual charge or pleasure that we feel when coming to understand, grasp, or solve some problem—an experience that is commonly dubbed the "insight" or "aha" phenomenon. The experience of solving a problem can also evoke laughter, frisson, or awe. But these different responses appear to be linked to distinctive characteristics of the problem-solving experience. For example, laughter is more likely to arise when the solution to a problem is suddenly recognized to be trivial. Frisson is more likely to arise when we make a connection that simplifies a problem. Awe is more likely to arise when we realize that the solution to a problem turns out to be massively complex. That is, frisson accompanies the experience of "gaining command" over a problem. Awe accompanies the experience of "losing command" over a problem. Laughter accompanies the experience of transforming a problem into something trivial.

A task as commonplace as solving a crossword puzzle can be the occasion to experience all three such emotions. We might have individual moments of insight. A chuckle may attend one solution, for example. As we finish the puzzle, we may have a frisson-like experience as we realize a previously unrecognized thematic unity that links the puzzle words together. Turning the page, we might have an awe-like response when we see that the next crossword puzzle is much larger (and harder).

In this chapter I have argued that laughter, frisson, and awe are three flavors of surprise. Since surprise represents a biological failure to anticipate the future, all surprises are initially assessed as threatening or dangerous. The body responds by initiating one of three primordial responses to threat: fight, flight, or freeze. The physiological basis for these responses can be seen in some characteristic behaviors: hair standing up on the back of your neck, shivers running up and down your spine, laughter, gasping, and breath-holding. In most real-world situations, evoking fight, flight, or freezing behaviors will prove to be excessive—an overreaction to innocuous situations. A slower cognitive process ultimately makes this assessment and begins to inhibit or modify the fast reaction response. Although the situation begins with a negatively valenced limbic response, it is replaced by a neutral or even positively valenced limbic response. The contrast between these successive assessments generates a subjective experience akin to relief. What begins as a brief moment of fear is transformed into a strikingly positive phenomenal experience.

In this chapter I have also argued that reaction responses are necessarily conservative or "pessimistic." To maintain their effectiveness, especially in environments that may generate lots of false alarms, reaction responses must be resistant to habituation or extinction. That is, reaction responses should be difficult to "unlearn." Notice, however, that this situation does not preclude learning. That is, there is no theoretical impediment to learning *new* ways of being surprised. In fact, there are significant biological benefits to be gained if an organism is able to learn new ways to become fearful. When it comes to fear, learning should be easy, but unlearning should be difficult.

At the beginning of this chapter I identified a number of types of emotions that can be evoked by music. But the discussion here has considered only three of these emotions: laughter, awe, and frisson. What about some of the other strong emotions that can be evoked by music, such as joy, exuberance, serenity, angst, sadness? Here I need to remind readers that the purpose of this book is to address the phenomenon of expectation (and the emotions that arise from expectation). While expectation probably plays a role in these other emotions, I suspect that the phenomenon of expectation is less relevant in evoking emotions such as sadness or exuberance.

Reprise

In this chapter, I have introduced the phenomenon of surprise. Expectations that prove to be correct represent successful mental functioning. The experience of surprise means that an organism has failed to accurately anticipate possible future events. From a biological perspective, surprise is always bad—at least initially. I have noted that there are different expressions of surprise and that these expressions echo the primordial behaviors of *fight*, *flight*, and *freeze*. Musical surprises are capable of initiating these responses, but the responses themselves are short-lived because an ensuing appraisal ultimately judges the stimuli as nonthreatening. The appraisal response inhibits the full expression of fight, flight, or freeze and also prevents the individual from becoming consciously aware of their brief brush with fear. Instead, the listener is left with a corresponding response of frisson, laughter, or awe. Evidence in support of this account can be found in the various physiological responses associated with fight, flight, and freeze that can also be observed among music listeners: piloerection, chills, changes of heart rate, laughter, gasping, and breath-holding.

I have suggested that the pleasure associated with these responses arises from limbic contrast—a phenomenon I've called *contrastive valence*. Pleasure is increased when a positive response follows a negative response. While surprise is biologically bad, surprise nevertheless plays a pivotal role in human emotional experience. Surprise acts as an emotional amplifier, and we sometimes intentionally use this amplifier to boost positive emotions. Suppose you had the opportunity to know in advance all of the future times and places when your most cherished goals or ambitions would be fulfilled. I doubt that many people would want such knowledge. Part of the joy of life is the surprise that accompanies achieving certain wishes. When all of the uncertainty is removed, the capacity for pleasure also seems to be diminished.

There is a tired old joke that begins by asking "Why do you keep beating your head against the wall?" Neuroscience gives some credence to the answer: "Because it feels so good when I stop." Of course there are more effective forms of head-banging. Music is one of them.

3 Measuring Musical Expectation

Without some way of gathering information about what individuals expect, all theories of expectation would remain purely speculative. If we want to hold our views accountable, we must be able to compare theories with evidence about how real minds anticipate the future. How then, do we go about determining what someone is expecting?

This question raises a host of related questions. What does it mean to have an expectation? How precise are expectations? Do we expect specific events, or do we expect "classes" or types of events? Can a person truly anticipate more than one possibility at a single time—that is, is it possible to have "plural" expectations? How do expectations manifest themselves as psychological or physiological states? How would one go about measuring what a person expects?

One definition of expectation might classify it as a form of mental or corporeal "belief" that some event or class of events is likely to happen in the future.[1] Such "beliefs" are evident in a person's "action-readiness"—that is, changes of posture, metabolism, or conscious thought that prepare the individual for certain possible outcomes but not for others. Such expectations can differ in strength of conviction or certainty.

Over the past four decades, researchers have devised a number of methods for gauging or estimating what people expect. The purpose of this chapter is to describe some of the experimental methods used to characterize listener expectations. Many of the same techniques are used to characterize nonauditory expectations, such as visual expectations. In addition, we will introduce some useful concepts from probability and information theory. These concepts will provide a convenient quantitative method for characterizing the range of expected possibilities, and for expressing the relative strength of conviction or certainty for various expectations.

Experimental Methods in Expectation

At least eight different experimental methods have been used to characterize a listener's expectations. Each method has strengths and weaknesses. Some methods are laborious whereas others are easy; some give fine-grained detail about the relative strengths of various possibilities, where others merely indicate that some outcome is possible for one listener. Some methods require the listener to reflect and introspect; others require no conscious thought at all. Some require the listener to be musically skilled; other methods require no special skills. Some methods are suitable only for adult listeners who can communicate verbally; other methods can be used with preverbal infants and nonhuman animals. Some methods require that the sound experience be periodically halted; other methods can be used without interrupting the listening experience. Each method is able to provide some useful information, but no method is a panacea. Becoming familiar with the different methods will help us better interpret the various experimental results, and also help us understand why different methods sometimes produce diverging—sometimes even conflicting—results.

It is important to understand that none of the following methods measures "expectation" in any direct sense. "Expectation" is a theoretical construct whose meaning and definition is open to debate. In experimental research, theoretical entities are rarely directly observable. Instead, researchers must operationally define some measureable quantity that is assumed to correlate with the theoretical construct. In reading the descriptions of the following experimental methods, it will become obvious that each measurement method is open to the legitimate charge "that's not what expectation really is." But in interpreting the experimental results we need to maintain some perspective. No method will capture the entire essence of expectation, but on occasion, some methods will allow clearer glimpses of how "expectation" operates.

1 Method of Tone Detection

Perhaps the earliest method for measuring auditory expectation was devised by Gordon Greenberg and Willard Larkin in the 1960s.[2] Working at the University of Illinois, Greenberg and Larkin had participants listen to tones in the presence of continuous loud noise. The listeners' task was simply to indicate whether or not they heard the tone.

Greenberg and Larkin discovered that listeners were better able to detect a tone if they expected a tone of a specific pitch to occur at a particular moment. They found that expectation allows listeners to direct their attention in both frequency and time: this directed attention has the effect of lowering the threshold of sensation for the sound. Greenberg and Larkin showed that there was a band of frequencies that was facilitated when listeners were expecting a particular frequency. If a listener expected a tone of 500 Hz, she was still able to detect a partially masked 550 Hz tone, but not a tone of 800 Hz. In effect, Greenberg and Larkin showed that listeners can direct their attention at particular frequency regions and time spans, and they used the method of detection to determine the shape and width of these "attentional bands."

Working at the Catholic University in Washington, Jim Howard, Alice O'Toole, Raja Parasuraman, and Kevin Bennett extended this method so that listeners were asked to detect a tone in some patterned context.[3] Listeners heard a twelve-note sequence presented along with a concurrent sustained noise. Two presentations of the sequence were given. One presentation was complete; the other presentation was missing one of the tones. Listeners were asked to indicate which of the two presentations was complete. The researchers established that the preceding sequence of tones facilitated the detection of some tones but not others. That is, they showed that the *melodic context* influences where listeners direct their attention.

The method of tone detection is rarely used today in experiments related to auditory expectation. But the work of Greenberg and Larkin remains important because it demonstrates two general principles concerning expectation. First, *accurate expectation facilitates perception.* When the events of the world conform to our expectations, we are better able to detect, perceive, and process these events. Over the past half century, this facilitating effect has been observed many times.[4] It is a principle that holds for both visual as well as auditory events. A second lesson from Greenberg and Larkin is that low-level sensory processes (like the hearing threshold for detecting a tone) are influenced by higher-level mental processes (like expectation). It is as though higher mental functions are able to reach down into the sensory apparatus and do some fine-tuning. Sensory systems don't just present information to the higher mental functions; in addition, higher mental functions can reconfigure a sensory system to focus on particular aspects of the sensorial world.

2 Method of Production

At the University of Washington in Seattle, James Carlsen, Pierre Divenyi, and Jack Taylor pioneered the simple technique of having listeners sing a continuation to some interrupted musical phrase.[5] Carlsen and his colleagues simply played a sequence of tones and asked listeners to sing what they thought would be an appropriate continuation. Carlsen used this method to compare the melodic continuations of American, German, and Hungarian listeners.[6] In analyzing the sung continuations, Carlsen found significant differences between the three groups, suggesting that one's cultural background influences listener expectations of what might happen next.

This method has a number of disadvantages. Notably it requires that participants have some singing ability (and be willing to sing while being recorded).[7] The method also relies on the participants' facility and comfort with improvising. Sung continuations can be confounded by vocal constraints. For example, if the antecedent context is low in pitch compared with the singer's vocal range, then there will be a natural tendency for the singer to produce a continuation that rises in pitch. Conversely, if the antecedent context is high in pitch compared with the singer's vocal range, then there will be a tendency for the singer to produce a continuation that falls in pitch.

Thus the melodic contour will reflect the participant's vocal range, rather than general melodic trends. (This problem can be controlled to some extent by determining the singer's vocal range prior to the experiment, and then tailoring the stimuli so they are positioned near the center of the participant's range.) When the antecedent context is short (such as a two-note interval), it may be impossible to infer the key that a participant might be using. In experiments by William Lake, this problem is eliminated by playing a tonic-establishing cadence before the start of the stimulus.[8] Another problem relates to deciphering what a singer sang. When singers are not well trained, the pitch and timing is often quite ambiguous. This introduces onerous technical challenges for the experimenter when transcribing what pitches and durations a singer produced or was intending to produce.

A variant of the method of production has been used by Mark Schmuckler. Instead of having participants sing, Schmuckler asked pianists to perform a continuation on a keyboard. Compared with sung continuations, the use of a keyboard circumvents the problems of pitch transcription, but it introduces a potential problem with skill. In order to reduce the need for keyboard skills, participants are given opportunities to try several different continuations, and rehearse their preferred continuation until they are satisfied with their response. This variation of the method of production has been popular with those researchers, such as Dirk-Jan Povel, who want to reduce the uncertainties introduced by pitch transcription.[9]

Yet another variant of this technique has been developed by Steve Larson at the University of Oregon. Larson contrived a task involving musical notation and expert music theorists.[10] He simply provided a notated antecedent context and asked music theorists to compose a suitable melodic continuation. In some ways, music theorists are ideal participants since they can draw on a lot of experience and know how to respond precisely. But music theorists often pride themselves on being musically clever, so a potential danger with this approach is that participants will be tempted to compose technical flights of creative fancy rather than commonplace or more intuitive melodic continuations. Larson explicitly instructed his participants to compose what they thought would be the most common or obvious continuation. As an inducement to this end, he offered a prize whose winner would be selected from the group of theorists who wrote the same (most frequent) continuation. That is, he gathered all of the responses, identified the most commonly occurring response, and then awarded a prize to one of the theorists (drawn at random) who had composed the most common continuation. One advantage of this approach over Carlsen's method is that it is possible to use complex harmonic or polyphonic stimuli (and responses) that would not be suitable for a sung response. However, there are a number of disadvantages to this method. The principal disadvantage is that notationally literate musicians often hold their own theories about melodic organization, and so the responses hold the potential to be confounded by theoretical preconceptions. Like

Schmuckler's keyboard task, the notation task lacks the spontaneity of improvised singing. That is, it encourages conscious, contrived, and reflective responses.

As we have seen, the principal drawback to the method of production is that participants in the experiment must have considerable musical competence—as vocalists, keyboardists, or by having facility with musical notation. This tends to limit the technique to participants who are relatively musical. Further, in requiring participants to "perform," the method of production also assumes that expectation facilitates not perception but *motor production*. Finally, Mark Schmuckler has pointed out that the method of production also requires a certain degree of conscious attention, whereas under normal listening conditions expectations may be largely unconscious and effortless.[11]

Compared with other methods, a unique benefit of the method of production is that it doesn't artificially limit a participant to producing a single tone following the given musical context. That is, whereas many other experimental methods assume that the preeminent expectation will pertain to the immediately succeeding tone, the method of production readily allows a participant to suggest several continuation notes as a coherent group. Later, we will see evidence indicating that a listener's strongest expectation may relate to an event that does not occur until after several intervening tones. The method of production provides better opportunities for an experimenter to study such possible long-term expectations, rather than focusing exclusively on note-to-note relationships.

3 Probe-Tone Method

Without question, the best-known experimental method for testing musical expectations is the probe-tone method pioneered by Roger Shepard and Carol Krumhansl.[12] Krumhansl and her colleagues at Cornell University have carried out numerous experiments using this technique. In simple terms, a musical context is presented—such as several chords or the initial notes of a melody. Following this context, a single tone or chord is played, and the listener is asked to judge this target (or "probe") sound according to some criterion. Often, the listener is asked to judge how well the tone or chord "fits" with the preceding musical context. The original contextual passage is then repeated and a different probe tone or chord is played. Following each presentation, the listener is asked to judge how well the new tone or chord fits with the preceding context.

In probe-tone experiments, a dozen or more repetitions of the same contextual passage may be presented—each presentation followed by a different probe. For example, several dozen possible continuations (probes) might be presented on successive trials. In this way, numerical ratings can be gathered for a large number of possible continuations. Thus, a significant advantage of the probe-tone method is that a detailed picture can be assembled where the listener provides information concerning several possible continuations, rather than only a single continuation. For example, with the probe-tone method a given participant might judge two or three continuations equally good. In the method of production, by contrast, participants must choose just one continuation, so it is problematic for the experimenter to infer that several different continuations would be equally acceptible for the participant. In addition, the probe-tone method can also establish which continuations sound "bad." That is, the method can be used to identify implausible as well as plausible continuations.

An obvious difficulty with the probe-tone method is that it is tedious. Each possible continuation must be tested separately. In practice, the total range of possibilities is reduced by the experimenter. For example, there are 88 tones available on a piano, but most of these tones are unlikely candidates to follow some melodic passage. Most tones will be implausibly high or low. Typically, the experimenter will reduce the candidate pitches to those within a two-octave range (one octave above or below) of the current pitch.

Yet another way to limit the number of possible continuations is to use so-called Shepard tones as probes. Shepard tones are specially constructed complex tones consisting of octave-spaced partials spanning the entire hearing range.[13] This encourages the listener to judge "goodness of fit" according to pitch-class rather than according to a single pitch. In Western music, there are only twelve pitch-classes, so using Shepard tones reduces all possible pitch continuations to just twelve. Using Shepard tones, however, means that the experimenter cannot directly infer the pitch direction (or contour) expected by the listener, given that each pitch class represents several possible pitches.

Apart from the tediousness of the probe-tone method, another difficulty is that it stops the music. When a listener judges "goodness of fit" one might imagine the response to arise from a combination of two sorts of judgments: (1) how well does this tone follow the previous note? and (2) how well does this tone terminate the tone sequence? Theoretically, it is possible that a tone follows well from the previous tone, but it might be judged as a poor fit because it evokes little sense of perceptual closure or completion. Conversely, a tone might follow poorly from the previous tone, yet evoke a strong sense of tonal closure, and so be rated highly by listeners. We will have more to say about these divergent interpretations in chapter 9.

Progressive probe-tone method In some cases, exhaustive experiments have been carried out to trace the changes in the listener's experience as the music progresses. For example, the first three notes of a melody may be played, followed by a probe tone. This procedure is repeated until a large number of continuation tones have been probed. Then the first *four* notes of the melody are played, again followed by one of several probe tones. This procedure continues for the first five notes, six notes, and so on. The progressive probe-tone method has been used to trace in detail such phenomena as how a modulating chord progression begins to evoke a different tonal center.[14]

Continuous probe method An obvious difficulty with the progressive probe-tone method is the tediousness of repeating the stimulus for each probe tone. If twelve probes are used following each note, a simple eight-note sequence will require 96 repetitions of the simulus in order to map the changing expectations over the course of the passage. In 2002, Carol Krumhansl and her colleagues introduced a variation of the probe-tone method in which a single probe tone (or chord) is sustained throughout the passage and the listener provides continuous responses as to the appropriateness of the probe at each moment in time.[15]

A problem with the continuous probe method is that it is hard to regard the responses as relating to expectations. Suppose, for example, that a tonic pitch is sounding continuously throughout a passage. As a cadence approaches, the dominant chord might sound. However, the harmonic clash between the dominant chord and the tonic pitch is not likely to result in a high rating for the tonic. Yet, one might presume that following a dominant chord, a high rating would be expected for the tonic. By comparison, if the penultimate chord is a subdominant chord, the probe-tone tonic is likely to receive a very high rating (because it is consonant with the sounding chord). Yet the dominant chord may well have evoked a greater expectation for an ensuing tonic than is the case for the subdominant chord. Said another way, one would expect the responses to continuous probe tones to be confounded by the resulting harmonic congruence: harmonic congruence is apt to play a much stronger role than expectation for subsequent events in determining a listener's response.

4 Betting Paradigm

Although the various probe tone methods do provide some information about the magnitude of various expectations, it would be useful to gather more precise measures of the subjective probabilities of different outcomes. In the *betting paradigm*, participants are given a "grub stake" of poker chips and asked to place bets on a set of possible continuations. Participants hear an antecedent passage and are invited to bet on what pitch they think will occur next.

I and my collaborators, Paul von Hippel and ethnomusicologist David Harnish, used this approach to compare the expectations for two cultural groups—Balinese musicians and American musicians. The experiment works as follows. Bets are placed on the keys of a mock-up of an instrument (in our experiment, a Balinese *peng ugal*). Bets need not all be placed on a single outcome. Instead, participants are free to distribute the poker chips across several possible continuations—varying the number of chips wagered according to the degree of certainty or uncertainty. Bets placed on the correct pitch are rewarded tenfold. Bets placed on incorrect pitches are lost. Participants are instructed to try to maximize their winnings.

As in the progressive probe-tone method, responses (wagers) can be collected following each note of a melody. The experiment begins with the participant hearing the first note of the melody while the pitch is indicated on a computer monitor. The participant is then invited to bet on what she or he thinks will be the second note. Once bets are placed, the actual second note is revealed, the winnings tabulated, and a sound recording of the melody is played, stopping before the third note. The participant is then invited to bet on what she or he thinks will be the third note. This process is repeated until a complete melody has been revealed.

In our experiment both American and Balinese musicians were tested on a traditional Balinese melody.[16] Throughout the experiment, participants could see the notation up to the current point in the melody, and could try out different continuations using a digital keyboard sampler that emulated the sound of the *peng ugal*. The betting context helped participants consider other possibilities apart from the first one that came to mind.

The principal benefit of the betting paradigm is that it allows the experimenter to calculate the subjective probabilities for different continuations. Assuming that the participant is behaving rationally, bets should be placed in proportion to the subjective likelihood of subsequent events. For example, if a participant thinks that a certain pitch is twice as likely as another pitch, then the participant ought to place twice as large a bet on the more probable pitch. Later in this chapter we will discuss how information theory provides a useful way to quantify such subjective probabilities.

A related advantage of the betting paradigm is that it allows the experimenter to measure confidence independent of relative probabilities. Suppose, for example, that the Balinese and American participants had roughly the same expectations ("A" is more likely than "B" which is more likely than "C"), but differed significantly in their confidence. A lack of confidence would be evident by participants' spreading their bets out more evenly. Once again, this can be calculated using information theory.

There are also several problems with the betting paradigm. For one thing, the procedure is even more tedious than the progressive probe-tone method. On average, we have found that it takes roughly three minutes for participants to complete their wagers for each note in the melody. A thirty-five-note melody thus can take as long as two hours to complete. Fortunately, the majority of our participants report that the task is fun, and that the gambling aspect of the task is highly motivating.

Like the notated version of the method of production, the betting paradigm encourages a conscious-reflective response rather than a spontaneous response. Unlike the method of production, the participant knows that there is a real melody involved, and is motivated to correctly anticipate the next note. Even more so than the method of production, the betting paradigm requires a degree of musical skill. In this regard, the probe-tone method is notably superior, not requiring any musical expertise.

A further problem with the betting paradigm is that the data may be confounded by a learning effect. Since the participants receive constant feedback about the accuracy of their wagers, they are likely to become progressively better at placing their bets as the experiment continues. Typically, experiments last more than an hour, so there is plenty of opportunity to improve and refine one's betting skills. Any apparent decrease in uncertainty as the melody progresses may therefore be an artifact of becoming a more savvy gambler.[17]

5 Head-Turning Paradigm

When we hear an unexpected sound, we will often turn our head in the direction of the sound. This basic reflex is referred to as the *orienting response*, and it is evident in all vertebrates, including young infants and adults. If a stimulus is repeated, after a while an individual will *habituate* to the stimulus and fail to orient to it. Further repetitions are unlikely to provoke a response. If a change is then made to the stimulus, and if the change is sufficiently novel, then a listener might reorient to the sound. This reorienting to a modified stimulus is called *dishabituation*. If an infant reorients to a modified sound, then one might interpret this as evidence that the infant didn't expect the sound.

Experiments employing a dishabituation paradigm typically repeat a stimulus until the participant becomes habituated to it. When habituation is complete then a new stimulus is introduced. If the new stimulus is perceived as the same (or similar) to the preceding stimuli then the participant will typically show no dishabituation. Conversely, if the new stimulus is perceived to differ from the preceding or expected stimuli then the participant is likely to show evidence of dishabituation by orienting to the stimulus.

The dishabituation paradigm is typically used when studying preverbal infants or nonhuman animals. The paradigm is used less commonly among adults since adults can verbally report perceived similarity or difference.

In research with infants, the head-turning paradigm has proved quite popular. There are a number of variants of this experimental method. In some cases, the experimenter merely tabulates whether or not a participant reorients to a modified stimulus. Another variant of the method measures the duration of orienting. This assumes that the greater the discrepancy between the expected and actual stimulus, the longer an infant will look in the direction of the stimulus. Yet another method will compare how quickly an individual habituates to two related events.[18]

Most of the experimental methods described in this chapter are not suitable for use with young children or infants. The principal advantage of the head-turning paradigm is that it can be used with infants and nonhuman animals. Also, unlike the bradycardic method (see below), it requires no special instrumentation apart from a video camera. One problem with the head-turning paradigm is that it requires that the participant first become habituated to the stimulus before some change is made. This makes the procedure extremely time-consuming. Newborn infants have difficulty controlling their head movements, so typically, the head-turning paradigm cannot be used until the infant is at least two or three months old.

An example of an experiment using the head-turning paradigm is one carried out by Michael Weiss, Philip Zelazo, and Irina Swain.[19] Weiss and his colleagues had infants listen to a nonsense word repeated until they habituated to it. The infants then heard either the original sound or one of four variants in which the pitch had been altered. Specifically, the frequency was modified by either 7, 14, 21, or 28 percent. They found that infants were most likely to reorient to the stimulus when the frequency of the nonsense word had been altered by 14 percent or more.

6 Bradycardic Response Method

Another version of the dishabituation paradigm examines changes of heart rate rather than head movements. When a stimulus deviates from an expected stimulus or attracts the attention of an individual, a measureable reduction of heart rate is often observed. Typically, such stimuli will result in a reduction of heart rate of about two to four beats per minute, followed by a recovery back to the normal rate. This response is referred to as *bradycardia*. Bradycardic changes of heart rate are associated with *interest* and *attending* to a stimulus.

Like the head-turning paradigm, the bradycardic response method is useful for studying the expectations of nonhuman animals, and especially for studying preverbal infants. Unfortunately, the method is tedious and the equipment can be cumbersome. Unlike the probe-tone method and the betting paradigm, each trial gives comparatively little information, and building a picture of infant expectation may require hundreds of trials from dozens of participants.

7 Reaction Time Method

Recall that Greenberg and Larkin showed that accurate expectation facilitates perception. When you hear an expected sound, you will typically be able to process it more quickly and respond to it faster (if a motor response is required). A quick reaction time is therefore correlated with high expectation. While reaction-time measures have long been used in experiments related to visual expectation, the method has gained favor only recently in research on auditory or musical expectation. Bret Aarden has shown that the method shows great promise for the study of melodic expectation.[20]

Aarden concocted a task where listeners were required to process and respond to sounds as quickly as possible. While they listened to an ongoing melody, Aarden asked his listeners simply to indicate whether the pitch contour of the melody had ascended, descended, or remained the same. In this task, there are three alternatives, and the listener must press one of three keys as quickly as possible after each note in the sequence. The responses are collected, including the elapsed time between the onset of the heard tone and the key press. The method is based on the assumption that if the pitch contour of a note moves as expected, then this will have a facilitating effect and so produce faster reaction times. Conversely, if a tone moves to an unexpected tone, this will increase the processing time and so result in a slower reaction time.[21]

The reaction time method has two notable advantages over other methods we have seen. First, it can be used in a continuous listening task where data are collected after every tone (except the first). The task is fairly difficult, so stimulus melodies are typically played using a tempo that is about 60 percent of the normal speed.[22] However, apart from the reduced tempo, there are no musical interruptions or pauses. As we will see later, compared with other methods, this method reduces the closure confound—where long pauses encourage listeners to respond to how well the tone provides a good ending point. A related advantage is that data collection is much faster than other methods. Rather than spending an entire experimental session on a single melody, a single session can collect data for dozens of melodies or tone sequences representing many different contexts. Another advantage is that the task happens so quickly that it is difficult for participants to engage in conscious reflection.

The reaction time method also has a number of drawbacks. First, unlike the betting paradigm or the probe-tone method, the reaction time method does not collect data for all of the different possibilities at each moment. Instead, we have a record of the processing time for specific contours within the heard sequence. When a listener makes a slow response, we have no idea of why this occurs—except that the heard tone evoked a more time-consuming mental process. There is no explicit information to tell us which alternative stimulus might have been processed more quickly. Finally, we must note that the reaction time method is fully premised on the idea that accurate expectations facilitate response times. At the moment, not enough is known about this relationship to know what sorts of confounds might be lurking.[23]

Despite these problems and caveats, the reaction time method shows excellent promise. As we will see in chapter 9, the results of the reaction time method and the probe-tone method sometimes diverge dramatically. We will see that these differences illuminate some important aspects of auditory expectation.

8 Evoked Response Potential (ERP)

The activity of neurons results in tiny electrical currents. When large numbers of neurons are active at the same time, the aggregate electrical current can often be detected through the scalp using suitably sensitive electrodes. The complicated hills and valleys of electroencephalographs have been studied for decades. Most of the activity remains an enigma; however, a consensus has slowly emerged about the interpretation of several specific features.

The most pertinent research related to expectation involves those electrical patterns that arise in response to a particular stimulus, like a tone. Since the recorded brain activity is in response to a stimulus, the ensuing electrical behavior is referred to as an *evoked response potential* or ERP.

Researchers still have difficulty interpreting individual ERP recordings. Typically, researchers average together many trials in which the same stimulus condition exists. It is the averaged data set—sometimes averaged across many subjects—that is able to tell a story. After the onset of the stimulus, a characteristic sequence of peaks and troughs can be observed in the ERP data. For convenience, successive peaks are designated P1, P2, P3, and so on, while successive troughs are designated N1, N2, and so on.

Suppose that a repeated sequence of identical sounds is interrupted occasionally by a deviant sound. Typically, this change is reflected in the listeners' response as an increased amplitude of the N2 waveform, which usually peaks around 100 to 250 milliseconds following the occurrence of the deviant sound. Because the electrical potential is negative and because it occurs in response to stimuli that fail to match the expected sound, the event is referred to as a *mismatch negativity* or MMN. MMNs can occur in response to changes of pitch, changes of loudness, and changes of timbre, among other things. MMNs also occur if any expected tone is replaced by a silent rest. Interestingly, MMNs occur if the listener *expects* to hear a change in the sound even if the sound remains unchanged.[24]

MMNs have been observed in listeners who are asleep and even in anesthetized rats, so the mismatch negativity can be evoked without conscious awareness of changes in sound. The magnitude of the effect is influenced, however, by attentiveness. For example, distracting tasks will attenuate the MMN.

The location in the brain of the evoked MMN is known to change depending on the type of stimulus change. For example, a change in frequency will evoke the largest MMN response along the sides of the head (temporal cortex). However, suppose an auditory pattern A–B–A–B–A–B is interrupted by the repetition of either A–A or B–B. In this case the site of maximum MMN response shifts toward the top of the head.[25]

ERP methods have a number of advantages and disadvantages. Like the head-turning paradigm and the bradycardic response method, ERP can be useful for studying the expectations of nonhuman animals, and especially for studying preverbal infants. No verbal responses are necessary, and no conscious awareness or thought is required. Unfortunately, many trials must be averaged together in order to infer anything.

Subjective Probability and Uncertainty

Having described the major experimental techniques for characterizing a listener's expectation, let us now turn to a second question: How do we express the strength of conviction or certainty of an expectation?

Suppose you are a participant in a betting paradigm experiment. There might be ten possible notes, and your task is to place your bets according to how likely you think each outcome is. There are lots of possible circumstances. If you are completely clueless, the best strategy would be to simply spread your poker chips evenly across all ten possibilities. Conversely, if you are completely certain of the outcome, the best strategy would be to place all of your chips on the expected note. In the first case, the bets reflect that you are completely uncertain of the outcome, whereas in the second case the bets indicate you are absolutely certain. Of course, there are many intermediate situations. You might be pretty sure that a certain outcome will happen, but you are less than 100 percent certain. In this case, you might place a small bet on all ten outcomes, but place most of your chips on the one outcome you think is most likely. Alternatively, you might be absolutely certain that one of the ten outcomes will *not* happen. In this case, you might spread your bets evenly across the remaining nine possibilities—placing no chips on the outcome you are certain won't occur. A more complicated situation might arise if you are moderately certain that only three outcomes are likely. Here you might split the majority of the poker chips between the three most likely outcomes, while placing small bets on the remaining possibilities.

From a research perspective, it would be convenient if we could distill any complex arrangement of bets into a single number representing the overall degree of certainty or uncertainty. Such a summary measure is provided by information theory. Using the so-called Shannon–Weaver equation, any arrangement of probabilities can be summarized by a single value that represents the aggregate uncertainty—measured in *bits*. When the number of bits is high, it means that the bets (probabilities) represent a high degree of overall uncertainty. Conversely, a low number of bits is indicative of high certainty.[26]

We won't bother to explain the equation here. However, the flavor of using bits to characterize uncertainty can be conveyed by some examples. Suppose we want to characterize the uncertainty of tossing a fair coin. With two possibilities (heads or tails) the amount of uncertainty is precisely 1 bit. Think of a bit as equivalent to one "yes-or-no" question: Did the coin come up heads? A single yes-or-no answer to that question is all we need to know to resolve the uncertainty. Similarly, if there are four equally likely outcomes, the number of bits is 2. If we have to choose from 8 equally likely pitches, then the uncertainty is 3 bits ($2 \times 2 \times 2$).

In the case of the von Hippel, Huron, and Harnish experiment described earlier, the *peng ugal* instrument provided ten possible tones. If all ten tones were equally

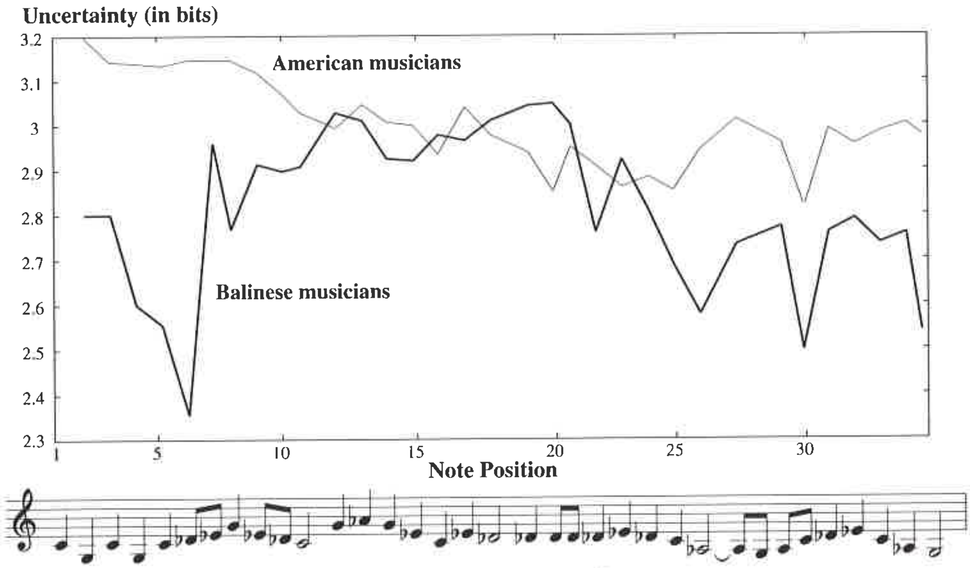

Figure 3.1
Average moment-to-moment uncertainty for Balinese and American musicians listening to an unfamiliar traditional Balinese melody. Uncertainty is plotted as entropy, measured in bits. In general, Balinese listeners show less average uncertainty. Note positions correspond with underlying notational rendering. The pitch levels shown in the notation are only approximate.

probable, then the uncertainty would represent 3.32 bits. A completely clueless listener (acting rationally) would place equal bets on all ten notes, and the Shannon–Weaver equation applied to this arrangement of bets would result in 3.32 bits of uncertainty.

Figure 3.1 plots the average uncertainty (expressed in bits) for the Balinese and American listeners in our experiment. Notice that after the first note, the average uncertainty for the American musicians was 3.2 bits. Since maximum uncertainty for ten outcomes is 3.32 bits, this means that after hearing the first note, the American musicians were almost perfectly clueless about what might happen next. After hearing the first note, the average uncertainty for the Balinese listeners was 2.8 bits, which corresponds almost precisely with seven equally probable states. This advantage is equivalent to being able to eliminate three of the ten notes as possible successors. By the fifth note of the melody, the average uncertainty for the Balinese musicians was roughly 2.35 bits. This is equivalent to being able to exclude five of the ten notes as possible successors.

Notice that by about ten notes into the melody, the American musicians are now comparable in confidence to the Balinese musicians in placing their bets. However, the Balinese musicians continue to exhibit less uncertainty—especially as the end of the melody approaches.

It is important to understand that figure 3.1 portrays average *uncertainty*—not predictive *success*. One can be confidently wrong as well as confidently right. In this case the Balinese musicians were not only less uncertain than the American musicians; they were also more accurate in their bets. A simple summary measure of predictive success is to compare "winnings." We started our participants with a nominal grubstake of $1.50 (not real money—for reasons that will soon become apparent). We rewarded accurate bets tenfold, while inaccurate bets were lost. If a player simply distributed the bets equally across all ten notes on the *peng ugal*, and left them there throughout the melody, then the final winnings would be the same as the initial $1.50 grubstake.

With regard to predictive accuracy, the differences between the American and Balinese musicians were striking. By the end of the melody, the most successful Balinese musician had amassed a fortune of several million dollars. The most successful American musician failed to do as well as the least successful Balinese musician. Moreover, several American musicians went bankrupt during the game and had to be advanced a new grubstake in order to continue.

Not surprisingly, Balinese musicians do better than American musicians in forming accurate expectations related to a Balinese melody. Although the specific melody was unfamiliar to both the Balinese and American listeners, the Balinese were able to take advantage of their cultural familiarity in forming suitable melodic expectations. Familiarity with a musical genre leads to both more accurate expectations and less uncertainty. However, it would be wrong to conclude that the American musicians were utterly clueless when listening to Balinese music. On average, the American listeners performed much better than chance. Either the American musicians were able to adapt quickly to the unfamiliar music, or they were able to successfully apply intuitions formed by their extensive experience with Western music—or both.

Conditional Probabilities—The Role of Context

In casino gambling, there is no link between a previous outcome and a future outcome. Each time we roll a pair of dice, the number that is rolled is independent of numbers previously rolled (this is true even for loaded dice). But in many real-world phenomena, subsequent probabilities do depend on preceding states. The probability of the occurrence of the letter "u" in text increases considerably when the preceding letter is "q." Likewise, in tonal music, the probability of occurrence of the tonic degree increases when preceded by the leading tone (the seventh degree of the scale). When the probability of an event is dependent on some preexisting state, it is referred to as a *conditional probability*.

In describing conditional probabilities, two concerns are the *contextual distance* and *contextual size*. Some states are influenced only by their immediate neighbors (i.e., small contextual distance). Other states are influenced only by states that are far away in space or time (i.e., large contextual distance). At the same time, states might be influenced by just a few other states or by a large number of other states. The size of the context of probabilitistic influence is sometimes also called the *probability order*. When the probability of occurrence for elements is totally independent of preceding elements (as with fairly thrown dice) the probability order is called the *zeroth order*; context sizes that take into account a single preceding element are called *first order*; *second order* denotes the probability order in which two preceding elements are taken into account, and so on.

It is important to note that the contextual size or probability order is independent of the contextual distance. Some event or state might be constrained only by its immediate neighbor (near context, small order). If an event is constrained by many neighbors, it will have a near context and large order. If an event is constrained by the presence of a single distant event, then it will have a distant context and small order.

By way of illustration, consider the following four contrasting examples:

1. Far context, small order A worker who receives a bonus might decide some weeks later to go shopping for a new jacket. Here, the likelihood of a future event (purchasing a jacket) is constrained by a single, somewhat distant earlier event.
2. Near context, small order Hearing her child cry, a mother might pick up the child. Here the future event (picking up the child) is evoked principally by a single immediately preceding state.
3. Near context, large order At a bingo parlor, a winner shouts out "bingo!" This event is provoked only by many preceding events, each of which caused another number on the card to be marked or covered.
4. Far context, large order A talented scientist might carry out a number of experiments leading to a major discovery that many years later results in her receiving a Nobel prize. The prize arose from many activities that were carried out decades earlier.

As we will see in later chapters, music exhibits a complete range of such dependencies. Most of the time, the principal constraints are of low probability order and involve a near context (e.g., one note influences the next note). But music also exhibits distinctive patterns of organization where distant contexts are more influential than near contexts and the probability order is quite large.

Reprise

In this chapter we have covered some basic background that will help us in discussing some of the experimental research pertaining to expectation. In the first instance we have described eight experimental paradigms used to characterize listeners' expectations. None of the methods is without difficulties. Each method makes different assumptions and provides subtly different information.

In addition, we have shown how information theory provides useful tools for measuring the strength or uncertainty of an expectation. We have also provided some conceptual language that will help us describe how the occurrence of a particular event might be shaped by other neighboring or distant events.