

CHAPTER 3

Emotions in the Music

There has been a growing consensus among philosophers of music that, contrary to the skeptical claims of Hanslick, it makes perfect sense to describe music in expressive terms, and that, again contrary to Hanslick's skeptical claims, there is more or less general agreement, among qualified listeners, as to what the music is expressive of, in any given instance, if, that is, it is expressive of anything (which need not necessarily be the case). More specifically, there has been a growing consensus that music can be, and often is, expressive of the garden-variety emotions, such as sorrow, joy, fear, hope, and a few other basic emotions like these.

As well, the consensus generally is that, when we say a passage of music is sorrowful, or fearful, or the like, we are not describing a disposition of the music to arouse such an emotion in us, but ascribing such an emotion, as a perceived property, to the music itself. This way in which the musical emotions are perceived, as, rather, in the music, than in us, with the music as their cause, was well captured by the late American philosopher O. K. Bouwsma, when he quipped that the emotion is more like the redness to the apple than like the burp to the cider.

But this view, though initially appealing, is not without its problems, the most discussed of which is *how* the emotion can be *in* the music. We all understand how an emotion can be 'in' something as a 'disposition.' Sad news is news that makes people sad: the sadness is in the news merely in the sense of a tendency of such news to sadden people. There is no 'metaphysical' problem there: that is, no problem about the nature of the property we are ascribing to the news.

Likewise, we all understand quite well what we are saying when we assert that a person is sad: the sadness is a conscious state of that person; that feeling of depression and lassitude one experiences upon losing a loved one or suffering a great disappointment. Again, there is no 'metaphysical' problem about how a person can possess the 'property' of sadness; what its mode of existence is. Sadness is a conscious state and persons are capable of having conscious states.

But music is not capable of being in conscious states (needless to say), so it can't possess sadness in *that* way. And, since it is the consensus that it

doesn't possess the thing as a disposition to make us sad, or, as we have seen, as a representational property, that is, as a representation of sadness, how exactly *does* it—the music—possess the sadness? It is all very well to say, with Bouwsma, that it is as the redness to the apple rather than the burp to the cider. However, that does not really answer our question. For, although we have a pretty good idea of how redness 'inheres' in apples and other red things, what we *don't* have is a pretty good idea of how the emotions 'inhere' in the music. Indeed, some people think we don't have any idea at all.

One of the traditional ways philosophers have of dealing with such cases is to try to make an analogy between the problematic case and an unproblematic one that it relevantly resembles. In the present instance, what might make us less uncomfortable with the notion of emotions 'inhering' in music as perceptual properties would be to discover cases in our ordinary experience where we commonly accept as a matter of course the notion of perceptual emotive properties belonging to non-sentient 'objects.'

Such a general argument was advanced by the remarkable American philosopher Charles Hartshorne in his highly original book *The Philosophy and Psychology of Sensation* (1934), many years before the problem emerged decisively for contemporary philosophy of music. Hartshorne adduced the common phenomenon of the emotive tone of colors to illustrate the point that emotions can be part of our perceptual field in perfectly ordinary circumstances well known to us all. Thus, he pointed out, yellow is a 'cheerful' color, not because it makes us cheerful but because its cheerfulness just is a part of its perceived quality, inseparable from its yellowness. That's just how we perceive yellow. In the same vein, he instanced other colors, as well as sounds, and other visual aspects of our world. In so doing, Hartshorne drew our attention to the fact that music is not alone in possessing for us emotions as perceived qualities of our sensible experience. The phenomenon is ubiquitous to our perceptual world. It should make us less uncomfortable, then, with the phenomenon in music if we realize that it is not merely a *musical* phenomenon but a phenomenon of human perceptual experience in general.

Nevertheless, the skeptical may reply, the problem of how emotions can be possessed by music, as perceptual qualities, has hardly been solved by discovering that objects other than music possess them. Indeed, it could be argued that the problem has actually been exacerbated. For now we have the problem not only for music, but for the other objects as well. If it is mysterious how music can possess emotive qualities, it is equally mysterious, for example, how colors can possess them. First we had one problem; now

we have two.

Hartshorne's answer to how emotive properties constitute part of the perceptual world of insentient matter is unlikely to appeal to the contemporary analytic mind. Indeed, his philosophy is a kind of 'pampsychism': the view that so-called 'insentient' matter has itself at least a degree of sentience. In other words, Hartshorne does *not* explain how insentient matter can come to possess emotive properties, but, rather, blurs the distinction between sentient and insentient; which is to say, to a degree, it's sentience all the way down.

As I say, few will be willing to take such a drastic step with Hartshorne, although, I hasten to add, he is not some wild-eyed fanatic, in touch with astral forces, but a cool-headed philosopher of the first rank whose works abound in useful and penetrating insights. So for those of us who are not attracted to anything as beyond human experience as what Hartshorne offers, a more mundane answer must be sought for the question of how music can possess perceptual emotive properties. It is a comfort, indeed, to know that music shares this expressive aspect with such everyday objects as simple colors, and helps to assure us that there is nothing bizarre or seemingly outside the realm of the human to call a passage of music mournful or jolly or anguished. However, that does not replace the need for an explanation of how music can possess such emotions as perceived properties.

One thing to notice, straightaway, about the comparison of the cheerfulness of a musical passage and, say, the cheerfulness of the color yellow is that the cheerfulness of yellow, as yellow itself, is a 'simple' property whereas the cheerfulness of music is a 'complex' one. What this means is that, if I say, 'This kerchief is yellow,' and you say, 'No, it's orange,' there is nothing else I can point to in the kerchief to show or convince you that the kerchief really is yellow and not orange. I cannot respond to your denial by saying something like: 'But don't you see it is such-and-such and so-and-so, so it *must* be yellow.' This, of course, does not mean that there is *no* way I can try to show you that the kerchief is yellow, not orange. The usual procedure would be for us to view it in sunlight rather than artificial light, or make sure one of us is not sight-impaired in some way, and so on. What I *cannot* do, because yellow is a simple property, is to point to something *else* in the kerchief and say: 'It's yellow because of *that*.' Yellow is not a complex property, so there is no 'that' to point to.

The same seems to be true of the cheerfulness of yellow. Like the yellowness itself, it is a simple quality. The cheerfulness just is a quality of the yellowness. Yellowness is not cheerful in virtue of some other perceived quality or qualities that it has beyond the yellowness (which is simply what

it is).

But when we go from the cheerfulness of the color yellow to the cheerfulness of a musical passage, the situation is radically altered. When I say that a passage of music is cheerful or melancholy, I can *defend* my claim by pointing to *other* features of the music in virtue of which the music is cheerful or melancholy (or whatever). I can say: 'Notice the rapid, skipping tempo, the bright major tonalities, the generally loud level of sound, the leaping, galloping themes. That's what makes it so cheerful.' Or: 'Notice the slow, dragging tempo, the dark minor tonalities, the subdued, quiet dynamics, the faltering, drooping themes. That's what makes it so mournful.' In this sense, the cheerfulness, mournfulness, and other emotive qualities of music are 'complex' qualities, not simple ones like the cheerfulness of the color yellow. They are also the kinds of quality that are sometimes called 'emergent,' because they perceptually 'emerge' from the other qualities that make them up. The cheerfulness of the music is a new quality, so to speak, that is produced by the combined force of the bright major tonality, the rapid skipping tempo, the loud dynamics, the leaping, galloping themes.

Calling the emotive qualities of music 'complex' qualities emphasizes the fact that a passage of music is cheerful, or melancholy, or whatever in virtue of *other* musical features that make it so. Calling them 'emergent' qualities emphasizes the fact that they are perceived as distinct qualities in their own right, separate from the qualities that may 'produce' them.

That the emotive qualities of music are *complex* qualities should not be thought to imply that when someone is hearing, say, the melancholy quality of a musical passage, he or she is necessarily aware of the other qualities productive of the melancholy. A person may be hearing the melancholy quality of the music without being aware that the music is melancholy in virtue of its slow, halting tempo, subdued dynamics, dark minor tonalities, faltering, drooping themes. He or she may not even know that it is those kinds of musical features that are generally responsible for making music melancholy. But if someone *fails* to hear the melancholy in the music, he or she can, often, be helped to hear it by having attention drawn to those features: 'Listen to its slow, halting tempo, those dark minor tonalities, the subdued dynamics, the faltering, drooping themes. Now don't you hear the melancholy?' Frequently the answer will be 'yes.'

Among those who think that the garden-variety emotions belong to music as perceptual, heard qualities, there is little dissension from the belief that these qualities are complex, emergent qualities in the sense explained just now. Furthermore, there is general agreement about what features of music are associated with what emotive qualities. Indeed, this is not just a

'theoretical' agreement among philosophers but a 'practical' agreement among composers and ordinary listeners as well. We know the latter because for well over 300 years composers have consistently utilized just those musical features to set melancholy and cheerful and fearful texts that we, as listeners, perceive to be the features responsible for the melancholy, cheerful, fearful emotive qualities we hear in textless music, as well as the features that make the music appropriate to the emotive tone of the texts the composers set. This is not a philosopher's pipe dream but a basic fact of musical listening, and of the musical craft in the West for centuries.

However, because there is general agreement that this piece of music is melancholy in virtue of its dark, minor tonalities, subdued dynamics, slow, halting tempo, drooping, halting melody, that piece cheerful in virtue of its light, major harmonies, loud dynamics, fast, skipping tempo, leaping, galloping melodies, does not solve the problem of how or why the music is in the first case melancholy, and in the second cheerful. Why, after all, don't I just hear the dark, minor harmonies, the subdued dynamics, the slow, halting tempo, the drooping, halting melodies? Why do I *also* hear the *melancholy*? This is the problem.

The most tempting approach to the problem of how emotions get 'into' the music is one that I am intimately familiar with, as I have been tempted by it myself. This approach begins with the thought that sad music being in slow, halting tempo, subdued in dynamics, with drooping, faltering melodies, and sad people walking in slow halting gate, with drooping bodies, speaking in subdued halting voice, cannot be altogether coincidental. Nor can it be altogether coincidental that cheerful musical works and cheerful people move rapidly, speak loudly, and even leap about, melodically in the music, bodily in people. In other words, there seems to be a direct analogy between how people look and sound when they *express* the garden-variety emotions (at least *some* of them) and how music sounds or is described when it is perceived as *expressive* of those same emotions. The intuition is that there must be some causal explanation lurking in this analogy: that the way we customarily express the garden-variety emotions must somehow explain why we hear those emotions in the music.

In my 1980 book *The Corded Shell*, later reprinted, with major additions, as *Sound Sentiment*, I attempted to give an account of musical expressiveness based, in large part, on the analogy between musical expressiveness and human *expression*. To facilitate that project I substituted for Hartshorne's example of the cheerfulness of yellow my own example of a St Bernard's sad face.

The St Bernard's face is not *expressing* sadness. The face of the St Bernard

is sad even when the creature is happy, it being at the other end that she *expresses* her emotions. The face, rather, is *expressive* of sadness: the sadness is a quality of the face as the cheerfulness is a quality of the yellow color. But, unlike the color yellow, the face of the St Bernard is a *complex* perceptual object and, hence, makes a far better analogue to the expressive musical object, which is also a *complex* one. What does the analogy have to teach us?

It seems fairly clear that the sadness of the St Bernard's face is there in virtue of our seeing it as a kind of caricature—but a recognizable resemblance—of the human visage when expressing sadness. Furthermore, we can point to individual aspects of the canine face—the sad eyes, the wrinkled brow, the drooping mouth and ears, the dewlap—that are exaggerated reflections of just such sadness-expressing features of a human face. The theory of musical expressiveness now under discussion is that expressive music is to human expressive features, overall, as the St Bernard's face is to the 'expressing' human countenance. Can this claim be substantiated?

We might usefully distinguish among the following three kinds of expressive features of music that the theory under discussion must deal with. First, there are the features of music that might be claimed to 'sound like' the sounds human beings make in expressing their emotions: the most obvious being speech. Second, there are the features of music that are said to resemble, in their sound, visible aspects of human expression behavior: for example, human gesture and bodily movement. Third, there are certain musical features, notably the major, minor, and diminished chords (to be explained in a moment), that have, for most people, the emotive tones of cheerfulness, melancholy, and anguish, respectively, but that, because they are, like yellow and its cheerfulness, simple perceptual qualities, do not seem to resemble either the sound of human expression, or its visible aspect. The chords by themselves, in other words, do not have a linear (but only a vertical) structure, and so do not seem to be able to be construed as resembling human expression behavior, which *does* exhibit complex structure: there is no structure in the chords to resemble the structure of human utterance and behavior.

The resemblance of the sounds of music to the sounds of human beings expressing their emotions is something that, as we have seen, has been claimed ever since antiquity, and was the motivating force behind the endeavors of the Florentine *Camerata*, in the sixteenth century, to revive, as they saw it, ancient sung drama. But whereas Plato and the *Camerata* utilized the supposed resemblance to explain how, they thought, music can arouse the emotions, the present project is to utilize it to explain how music can

embody them as heard qualities.

The first question for such a theory would appear to be whether there really is a perceived resemblance between the sounds of music and the sounds of human expression. Let us take, for example, music that is heard as melancholy and music that is heard as cheerful.

Certainly this much can be said. Melancholy music and melancholy speech and utterance have some obvious sound qualities in common. Melancholy people tend to express themselves in soft, subdued tones of voice; and melancholy music tends to be soft and subdued. Melancholy people tend to speak slowly and haltingly; and melancholy music tends to be in slow tempi and halting rhythm. Melancholy people's voices tend to 'sink,' and tend to remain in the low vocal register; and melancholy music too exhibits the same characteristics.

In contrast, cheerful people express themselves in bright, loud, sometimes even raucous—certainly not subdued—tones; and cheerful music tends to be bright, loud, and in the high register. Cheerful people are not slow or halting in speech and utterance but bright and sprightly; and cheerful music, likewise, is quick and sprightly. Cheerful people's voices rise energetically into the high register; and so too do the melodies of cheerful music.

In all of this particular attention should be paid to melody. For there is no aspect of Western music that is more amenable to analogy with the rise and fall in pitch of the human speaking voice than the rise and fall in pitch of music's melodic line. Furthermore, melody is that aspect of music that, historically, has been singled out most frequently as the primary *expressive* aspect. This perceived analogy between melody and speech is a leitmotif in writings about music's expressiveness from Plato to the *Cameraia* to the eighteenth-century philosophers to the present day. The cheerful melodic line, like the cheerful speaking voice, is high, loud, fast, 'running,' and 'leaping.' The anguished melody, like the anguished speaking voice, shrieks and cries, leaps in dissonant intervals, and proceeds in 'jerks,' with irregular pauses.

But beyond the phenomenon of music's 'sounding like' the vocal expressions of melancholy or cheerful people, many listeners perceive an analogy between the heard properties of music and visible human behavior as well. Music is customarily described in terms very similar to those we use to describe the motion of the human body under the influence of such emotions as melancholy and cheerfulness. Thus a musical phrase may leap joyously, or droop, or falter, like a person in motion. To put it more generally, music is customarily described in terms of motion; and so the same descriptions we use to characterize it are frequently the ones we use to

describe the visible motions of the human body in the expression of the garden-variety emotions.

I called this theory of musical expressiveness, in *The Corded Shell*, the 'contour theory,' because, to put it somewhat figuratively, the 'contour' of music, its sonic 'shape,' bears a structural analogy to the heard and seen manifestations of human emotive expression. One thing to be noticed straightaway about the contour theory is that it is going to have trouble with our third kind of expressive musical feature, the major, minor, and diminished chords: for *they* do not have any contour at all, so find no analogue, apparently, in the contours of human expression behavior. I shall deal with these seemingly anomalous expressive features in a moment. But before I do I want to complete this account of the contour theory.

The contour theory of musical expressiveness faces immediate problems. To begin with, it must not become a representational theory: it must not, that is to say, be construed as the theory that music 'represents' the voice and gesture of human expression, the way paint on canvas represents the visible features of the world. For representation does not capture the way we experience the emotive qualities of music. We do not, that is to say, hear sounds as representations of melancholy and cheerful behavior; the way we see paint on canvas as a representation of melancholy and cheerful men and women, and *then* hear the music, in virtue of these representations, as melancholy and cheerful. We hear the melancholy and cheerfulness of the music immediately, in the music, and can be quite unaware of the features of the music in virtue of which it is melancholy or cheerful. And even if we are consciously aware of the expressive-making features of the music, which we may frequently be, we do not perceive them as representations of anything.

Furthermore, there must be some explanation produced, in defense of the contour theory, for why it is the similarity in structure between music and *expression behavior* that plays so important a role in the listening experience. After all, the contour of music is probably similar in structure to inanimate sounds and natural objects, as well as to human expression behavior. What's so special about expression behavior that it should be singled out for mention above those other things?

Finally, does the contour theory do any better a job of capturing our experience of music's expressive qualities than a representational theory would do? Can it capture the way we experience the emotions in music, namely, as directly perceived perceptual qualities?

In order for the analogy between musical contour and the contour of human expression behavior to work non-representationally, it must work

subliminally: that is to say, we must not be fully aware of what is going on; we must not be aware of the analogy. Let us, for the moment, assume that this is what is happening. But why should we hear *emotions* in the music because of this subliminal perception, and not something else?

I believe one possible answer to this question can be found in a well-known perceptual phenomenon. When presented with ambiguous figures, we tend to see them as animate rather than inanimate forms: as living rather than non-living entities. We tend to see living forms in clouds, in stains on walls, in the shadowy things lurking in the woods. We see the stick as a snake. Why? Because, perhaps, we are hard-wired by evolution—by natural selection—to do so. Evolution says: 'Better safe than sorry. Better wrong than eaten.' Living things can be dangers to you. It is better to see the stick, immediately, incorrectly, as a snake, than to be snake bit, in pondering the question, if it turns out to be a snake after all.

Now, if this be true of sights, might it not be true of sounds as well? Might it not be the case that we are hard-wired by natural selection to hear sounds, where possible, as animate: where possible, as in music, as utterance and 'behavior'? I advance that as a possible, perhaps plausible, hypothesis. If it is true, it tells us why we hear the analogue of musical contour to human behavior and not to the other things it might resemble.

But if we hear sound as animate, why do we hear it as *expressively* animate? Well, if you want to carry through the 'survival' idea, it is to our advantage to know what emotive attitude is being evinced towards us by the living creature we may be encountering. If it is anger, we must prepare for fight or flight to be safe. If it is a benign emotion, other behavior is appropriate.

One might also add to these 'evolutionary' considerations the fact that just those 'simple' emotions we tend to hear in music are emotions whose expression behavior in human beings has direct analogues in the expression behavior of the higher primates and other mammals. In other words, their modes of expression seem hard-wired and deep in the inherited nature not only of us, but of other animals as well. This can be seen as corroborating evidence for the evolutionary story already told.

But now a further question presses itself upon us. In the case of ambiguous visual phenomena, we are conscious of what we are seeing (or think we are seeing). I take the stick for a snake and *run*. However, that does not seem to be what is happening in music. We are conscious of the expressive property, the emotion; we are not conscious of taking the musical contour for human utterance or behavior. Is there any plausible reason for

this to be the case? We can't just *assume* it to make our theory work.

Well, consider this. The sense of sight is the primary 'survival' sense for human beings (and other of the higher primates). The sense of hearing is *not*, although it may very well have been for our ancestors way down the evolutionary chain. Thus there is no need for us *consciously* to hear things as threatening the way we consciously see things that way. So it is not completely unreasonable to suppose that what may very well have been a propensity consciously to hear ambiguous sounds as animate and (potentially) threatening has atrophied in us, like the appendix, and remains a vestigial relic of a more sound-oriented past. To put it another way, it is not completely unreasonable, on evolutionary grounds, to think that, while the seeing of ambiguous forms as animate remains a conscious phenomenon of human perception, the hearing of sounds that way has sunk back into semi-consciousness as a kind of 'background noise.'

Here, then, is *one* theory of how music comes to embody expressive qualities like melancholy and cheerfulness. It is agreed on all hands that music is melancholy and cheerful, and so on, in virtue of certain standardly accepted features. It is perennially remarked on that these features bear analogy to the expression behavior, bodily, gestural, vocal, linguistic, of human beings. One can construct an evolutionary story of how and why we might be subconsciously, subliminally aware of this analogy and that this should cause us to perceive the music as melancholy or cheerful or the like as we perceive the sadness of the St Bernard's face. I have named this theory the contour theory of musical expressiveness.

But we have yet to work one further element into the contour theory: that is the expressive chords, major, minor, and diminished. These chords are generally perceived as cheerful, melancholy, and anguished, respectively; and you can hear this for yourself by playing the three notes together on the piano: first, C-E-G (major); then C-E flat-G (minor); finally, C-E flat-G flat (diminished). The problem is that these individual chords, not having a contour, being experienced as simple qualities, do not seem to bear any analogy at all to human behavior—hence must be expressive of cheerfulness, melancholy, and anguish in some *other* way than that allowed for by the contour theory of musical expressiveness. So the contour theory cannot be the whole story.

At least this much can be said for the contour theory straightaway. It is *no* worse off than any other theory in this regard. There is *no* generally accepted explanation for the difference in emotive quality between the major and minor chords, over which much ink has been spilled in the last 300 years. So, that the contour theory cannot provide one is no great deficit. But

there is no harm, anyhow, in trying. Here are two suggestions.

The first suggestion is that we hear the vertical structure of the chords as a kind of contour. Compared to the major triad—that is, the major three-note chord, C-E-G—the minor triad has a lowered third: that is, the E is the third of the C-major chord, the E flat is the third lowered a half step, the smallest interval in the Western harmonic system. (The E is called the 'third' because it is the third note up from the C: that is, C (1), D (2), E (3). The G is called the 'fifth' because it is the fifth note up.) Now think of the lowered third, E flat as kind of sagging, or sinking, depressingly from E to E flat. Might that give a depressing, melancholy cast to the C-minor triad? There is a downward tending contour of the C-minor triad, as compared to the C-major one, like the downcast contour of the melancholy speaking voice or posture. And the diminished triad, C-E flat-G flat, is even more depressed: it has both a sinking third *and* a sinking fifth. Pretty far-fetched? Perhaps so.

Well, then, let's try this. Perhaps the chords should not be considered as isolated elements but in context: that is to say, functioning in musical structure. Perhaps it is from their function within musical compositions that they gain their emotive color.

If you go to the piano again, and play the major triad, C-E-G, the minor triad, C-E flat-G, and the diminished triad, C-E flat-G flat, in succession, I think you will perceive a big difference between the first two and the third. The diminished triad is an 'active' chord: that is to say, it sounds as if it must go somewhere; it must, as a musician would say, be 'resolved.' It can't stay at rest where it is. Or, another way of putting it: you couldn't hear it as the final chord of a musical composition. To hear what I mean, now lower the G flat to D flat, and raise the C to D flat. You have, in doing that, 'resolved' the diminished chord, and you can hear for yourself that things are now 'at rest.'

Might one suggest, then, that what gives the diminished chord its dark, anguished quality is its function, in musical structure, as an active, uncommuted, unresolved chord? It is restless, so to say, in its musical function; when it occurs in a compositional structure, at least until fairly recently in the history of the Western harmonic system, it imparts that restlessness to the contour of the melody it accompanies. From its 'syntactic' or 'grammatical' role in music it gains, by association, as it were, even when alone, its restless, 'anxious' emotive tone.

Furthermore, might this not be true, to a lesser degree, of the minor triad as well? It is, indeed, a matter of historical fact that the minor chord, until well into the nineteenth century, was considered more active than the major,

which is why compositions in the minor mode for a long time almost invariably ended on a major rather than a minor chord. And, although ending on a minor chord is now no big deal, I think we still feel the minor ending as more restless than the major. You can hear this for yourself by resolving the diminished triad, C-E flat-G flat, two different ways: first, by lowering the G flat one whole step to E (F flat) and raising the C one half step to D flat; and then doing what you did before, lowering the G flat to F and raising the C to D flat. Do this a few times in succession. Which resolution sounds more 'restful' to you. I would venture to predict that it is going to be the latter: the major resolution. For what you did first was to resolve the diminished chord to D flat minor, and then to resolve it to D flat major. I think you, like your musical ancestors for over three centuries, found the major resolution grammatically and emotionally the more fulfilling one: the one that sounds the most stationary; the one that sounds the most completely *final*.

I don't know if this attempt to accommodate the major, minor, and diminished chords to the contour theory of musical expressiveness is any more plausible than the first. In any event, even where the contour theory does best, with the larger structural elements of music, particularly melody, it is not without serious problems and many detractors.

In the version of the contour theory that I have given, *my own version*, there are numerous difficulties that even I find daunting. Here are some.

To begin with, how convincing is the claim that there are recognizable analogies or similarities between music and the 'shape' of human expression? Do melodies *really* much resemble human speech in any significant way? Many people think the similarities adduced pretty far-fetched. And when one goes to supposed analogies between how music sounds and how human expression looks, there is bound to be more skepticism still. Does it make any sense at all to say that a passage of music is melancholy in virtue of sounding the way a human being gestures or moves when he or she is melancholy? Can music *sound* like a gesture or bodily posture? Can sense modalities be crossed that way? There is certainly plenty of room for doubt about it.

Furthermore, the whole psychological apparatus that the contour theory requires is highly conjectural, to say the least. What evidence, if any, is there for the claim that listeners subliminally hear the analogy, if indeed it exists, between the contour of music and human expression? And even if they do hear it, does that adequately explain our experience of hearing emotions in the music as perceptual qualities. As well, does the phenomenon of seeing things in ambiguous figures—seeing the stick as a snake, or the faces and

figures in clouds—transfer to sounds and what we hear (if anything) in them?

Finally, what of the evolutionary explanation offered for, first, the tendency to ‘animate’ ambiguous figures and, second, the difference between this tendency, as evinced in visual perception, and as evinced in hearing? Are they plausible? Is there evidence for *either*?

The biologist Stephen J. Gould scorns such armchair evolutionary explanations as I have given, labeling them, contemptuously, ‘just-so stories,’ the point being that a natural selection story, just like Kipling’s fanciful ‘explanations’ for how the leopard got its spots, or the elephant its trunk, can be made up by almost anyone, including an amateur like myself, for any trait you like. So it is probably wise not to place much faith in these exercises. The skeptical, therefore, will be wary of the evolutionary support I have adduced for the claims that, all things being equal, we will see living forms rather than non-living ones in ambiguous perceptual arrays, and, second, that, because of the primacy of sight over hearing as a ‘survival sense,’ in human beings, our perception of ‘animate’ forms in sounds, as opposed to sights, will be dim or subliminal, as the contour theory requires.

Add all of these difficulties together and the contour theory begins to look pretty shaky. Indeed, it looks shaky not only to its detractors but to at least one of its supporters as well: *me*. Having vigorously defended the contour theory on two separate occasions, I can no longer say that I am not without serious qualms.

But the funny thing about the contour theory, or, in general, the theory that music is expressive of the garden-variety emotions by virtue of analogy to human expression behavior, both vocal and bodily, is its perennial attraction. It simply refuses to die, in spite of its numerous difficulties. There doesn’t seem to be another, more plausible alternative.

That there isn’t another game in town, of course, does not constitute much of an argument for the contour theory or its relations. Tomorrow is another day, and may well bring another, more convincing account.

What, then, should we do here and now? When a problem remains unsolved, one obvious strategy is to work at it until it gives up its secret, and only *then* go on to the problem that logically follows next. But that is not the strategy I think will have the best results. I think it a mistake for us to remain bogged down with the question of *how* music comes to embody the garden-variety emotions as perceptual qualities. Rather, since there is a consensus, more or less, that music does exhibit the expressive qualities in this way, let us now go on to examine what role, as so understood, they play

in musical structure and the musical experience.

Let us, then, treat music, in this regard, as what the scientists call a ‘black box’: that is to say, a machine whose inner workings are unknown to us. We know what goes in, and what comes out, but what causes what goes in to produce what comes out—of that we are ignorant. With regard to how music comes to exhibit the garden-variety emotions as perceptual qualities, it is to us a black box. We know what goes in: the musical features that, for three centuries, have been associated with the particular emotions music is expressive of. And we know what goes out: the expressive qualities the music is heard to be expressive of. And rather than becoming obsessed with penetrating this black box, we should, or at least *some* of us should, go on to see what implications this new way of looking at music’s expressive qualities (for it is a new way) has for our understanding of music as a whole. That is the project I propose to pursue in Chapters 6 and 7. But before I do that I must convey to you some general idea of how music, and our experience of it, are to be construed. Only then will we be able, so to speak, to add the garden-variety emotions, as perceptual qualities, into the mix, and determine the part they play. So we will put aside, temporarily, the question of music and the emotions, in the following two chapters, only to return to it after the necessary groundwork has been laid.