UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

AUDITORY AND MUSICAL EXPERIENCE: PERCEPTION, EMOTION, AND EXPRESSION

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of
Doctor of Philosophy

By
Babak M. Khoshroo
Norman, Oklahoma
2023

AUDITORY AND MUSICAL EXPERIENCE: PERCEPTION, EMOTION, AND EXPRESSION

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF PHILOSOPHY

BY THE COMMITTEE CONSISTING OF

Dr. Sherri Irvin, Chair

Dr. Sarah Ellis

Dr. Adam Green

Dr. Martin Montminy

Dr. Amy Olberding

To Shailyn

ABSTRACT

In this dissertation, I address two questions about auditory and musical experience: (1) How can music (a set of organized sounds) be heard as expressive of emotions such as sadness or happiness? (2) How are sounds (not limited to musical sounds) and their sources related to each other? Regarding the first question, my focus is on the *experience* of musical expression: what it is like to experience music as cheerful or anxious, for example. I argue that our so experiencing music is based on experiencing a *feeling* while attending to the relevant qualities of an expressive piece or passage of music. I suggest that to hear the flow and continuity of music, we feel an inner movement, and this feeling sometimes leads us to attribute emotions to music. This is a version of a so-called arousal theory of musical expression, but a merit of my proposal compared with previously defended arousal and cognitive views is that it can explain better how attending to musical properties results in hearing music as expressive. The second question of the dissertation relates to the problem of how when we hear sounds, we hear them as being *of* things and events in our surroundings (we can hear, for example, the sound of a metal bar being struck, a wooden door being knocked, a violin being bowed, and so on). How should we explain this relation between a sound and the thing it is a sound *of*? I argue that a sound is a property of an object *and* the event it participates in to make that sound. In other words, a sound is a *complex property* instantiated by an object and the relevant (sonic) event. For example, when I hear the sound of a violin being bowed, I hear properties of the instrument and properties of the bowing of the instrument's strings *in* that sound. An advantage my theory has, particularly over the event and property theories of sound in the literature, is that it can explain more convincingly how sounds give us perceptual access to the things and events they are sounds of. I will also sketch a future research direction based on applying my theory of sound to explore the content of musical experience, what we hear by hearing musical sounds in a piece of music.

Table of Contents

Acknowledgments

The germ of my work in this dissertation formed when I discovered philosophy through music and decided to change my academic field, leave my career in my home country of Iran, and move to the United States. However, I could not have been brave enough to realize this change and pursue my dream without the support and encouragement of the love of my life and my always friend Sara. Moreover, as she is a visual artist, her artistic insights and ideas helped me develop many of the ideas presented in this dissertation—in fact, it seems I can never form a new idea without discussing it with her in one way or another on several occasions.

Moving to and living in a different culture, particularly when, because of different social and political circumstances, you do not have the opportunity to go to and return from your home country with ease of mind, means experiencing a form of loneliness and being far from your family and many of your friends, the people with whom you share a part of your identity. However, I have been lucky to have a family whose love and support go far beyond geographical and political hurdles. Particularly without the support and encouragement of my parents I could not have focused on my work and tolerated being far from them for several years.

I have also been very fortunate that I have had the chance to work with Dr. Irvin, who did not merely guide me philosophically but kindly understood my life situation and generously supported me during these years. She showed me how doing philosophy as a scholar can be another aspect of living life humanely.

I am also indebted to all the members of my committee—Dr. Ellis, Dr. Green, Dr. Montminy, and Dr. Olberding—for their engagement with my work and insightful comments. I am indebted to Dr. Ellis who patiently and carefully helped me particularly with the musical

parts of this document. Moreover, I was honored to have Dr. Linda Zagzebski on my committee and benefit from her thoughtful feedback on the first two chapters of the dissertation before her retirement.

Doing philosophy is specifically enjoyable when you have the opportunity to be a part of a close-knit group with whom you can share and discuss your thoughts from their embryonic stages. I did have that opportunity. As a member of an active group working on different aesthetic subjects, I am indebted to Atabak Akhlaghi, Cheryl Frazier, Jeremy Fried, and Stephanie Holt for fruitful discussions as I progressively developed the chapters of this dissertation.

And at the end, I should thank my lifelong friends Hamed Hassanzadeh and Hamed Moghaddam, from whose musical insights and knowledge I benefitted at different points of this dissertation, and Behnam Akbari, a philosopher and musician in Persian classical music, with whom I had fruitful discussions on culture and music.

Introduction

Research Question and Approach: One

Music has an indispensable place in many human cultures. Singing, performing, and composing are three main musical practices shaping a musical culture. Central to such practices are sounds (combined with words in the case of songs) organized by people while they sing, play musical instruments, or only imagine and hear sounds (and words) in their heads to create music. However, music is not only how humans organize and produce sounds but also how they listen to or *experience* them. As with our mother tongue, we learn how to listen to and appreciate the music produced in our culture when we are exposed to different types of music from childhood. This is how our musicality is developed as a basic human capacity, a capacity that is at the core of all musical activities. Timothy Rice (2014) has defined this capacity thus:

> "Musical" […] does not refer to musical talent or ability; rather it refers to the capacity of humans to create, perform, organize cognitively, react physically and emotionally to, and interpret the meanings of humanly organized sounds. The definition assumes that all humans, not just those we call musicians, are musical to some degree, and that musicality (the capacity to make and make sense of music) defines our humanity and provides one of the touchstones of human experience.[1]

A part of this dissertation focuses on an aspect of musicality that is related, in Rice's terms, to our *emotional reaction* to music. It is a commonplace phenomenon that, on many occasions, our experience of music gets related to our feelings and emotional life in one way or another. For example, a piece might make us sad or cheerful. Or we might describe some music or performance as melancholy or energetic. Or, during listening to a piece or song, we might have feelings of different sorts, which even we cannot describe, at different points. It has been a

[1] p. 1

riddle how to explain this phenomenon for many years. Philosophy and psychology have been the two main fields to grapple with this problem, getting insights from musicology and (sometimes) ethnomusicology.[2]

In the philosophical literature, this problem is typically called the problem of *musical expression*: *how can a set of organized sounds be heard as expressive of emotions such as sadness or happiness?* A puzzling facet of this problem is that music is not a sentient being like humans in such a way that it can have any emotional experience and express it. So how can a set of organized sounds be expressive of any emotion? But if, and this is another puzzling facet, the emotional experience in question is somehow what *we listeners* feel in response to music, why should hearing a set of organized sounds arouse any emotional feeling in us (for example, the feeling of cheerfulness or nervousness)?

The approach taken in this dissertation to tackle the above problem is to delineate how *attending to* organized sounds during music listening leads to emotional feelings experienced by the listener. That is, I will argue that we must sometimes give a feeling response to music in order to make sense of how sounds are organized in a musical piece. I will argue that this process is at the core of hearing music as expressive of emotions. The contribution of my approach is that it can integrate two main intuitions defended in two opposing views of musical expression in the literature: (a) expression is a musical property that we perceive in *music*, and (b) expression is a feature of music *experience*. The main argument I will develop and argue for in the first three

[2] For example, in his *Republic*, Plato pondered the arousal power of music and its ethical impacts on human souls. Different modern thinkers also raised or tried to deal with the issue in different contexts. A few examples of philosophical studies that initiated approaching this matter in more substantive way are Langer (1956), Kivy (1981), and Davies (1980, 1994). For a survey of the main psychological attempts to approach the problem, see Juslin (2019). An example of the importance of this matter in a non-Western musical culture is sadness in Persian classical music (see Majid Kiani (1999)).

chapters of this dissertation integrates (a) and (b) by showing that musical expression is the result of how we make sense of music as organized sounds by *attending to musical properties*.

Research Question and Approach: Two

My research on musical expression led me to identify and focus on a different but related problem regarding the nature of sounds. What I found an interesting problem concerning our music experience was whether and how, in hearing musical sounds, we hear the instruments and people making those sounds. In a musical piece, I seem to hear *violins*, *flutes*, *trumpets*, etc., in the sounds they are making (if listening to a live performance) or they made (if listening to a recording). When I hear a person singing, recorded in an audio file, I seem to hear *that person* and not merely a set of sounds. The point is that in musical sounds, we seem to hear objects and people in or through the sounds they make. But how is that possible? Is it the case that hearing the sound of, say, a violin is also hearing the instrument itself? If so, how should we explain this?

What made this problem specifically interesting for me is related to the cultural roots of our music experience. Musical entities (works, performances, etc.) have strong ties to the cultural contexts in which they are created.[3] These entities result from concrete events, objects or instruments, and people's actions or decisions in specific socio-historical contexts. It seems—or at least it seemed to me when I was pondering this problem—that one thing that can explain this cultural aspect is that musical sounds are essentially connected to the concrete cultural and historical events, artifacts, and people that make them. An acculturated listener seems somehow connected to such socio-historical artifacts by hearing sounds humanly organized in a specific

[3] For how this idea can be philosophically corroborated from the perspective of music ontology, see Lydia Goehr (2007) and Jerrold Levinson (2001).

culture. My initial hypothesis was that in hearing musical sounds, we *hear* their (cultural) sources.

However, to examine my hypothesis, I found that I needed to consider the problem in a broader context not restricted to music: *how are sounds and their sources related to each other?* This problem has been explored in the literature concerning auditory perception and the metaphysics of sound. In approaching the broader problem, there are two key considerations: (a) What is a sound? (b) What do we hear in sounds? While the former is a metaphysical question about the nature of sounds, the latter concerns the phenomenology of hearing. To understand whether and how we hear sound sources in or through the sounds they make, I explore (a) and (b) in relation to each other: I try to develop a theory on the nature of sounds through phenomenological considerations on our hearing of sounds and things in our surroundings. The view that I found plausible and defend in Chapter 4 is that sounds are complex properties instantiated by both objects and events they participate in. This is a metaphysical view on sounds that, hopefully, explains how hearing sounds is hearing their sources. This view sets a basis for exploring the musical problem I formulated at the beginning of this section as regards the cultural roots of music experience. I will sketch some aesthetic and musical implications of my view in the Conclusion, but due to the scope of this dissertation, this will be a subject for my future research and not the focus of any chapter in the dissertation.

Structure

The first three chapters of the dissertation are devoted to the first research question and Chapter 4 to the second one explained above.

In Chapter 1, I first put the problem of musical expression in the context of the problem of musical meaning. I then suggest that in understanding musical expression, we must distinguish two dimensions of the problem: phenomenological and causal. Based on this distinction, I critically examine a theory of musical expression according to which expression is a *property* of a musical work. I argue that this theory is phenomenologically unconvincing. My arguments in Chapter 1 set up an analytical framework for exploring musical expression from phenomenological and causal viewpoints.

In Chapter 2, I focus on three ideas central to some theories of musical expression that try to explain what it is like to experience musical expression. Those ideas are: (1) hearing musical expression is *imagining* musical personae expressing emotions; (2) hearing musical expression is *metaphorically* attributing emotions to music; (3) experiencing musical expression is being *aroused* by music and *feeling* emotions. I examine these ideas and show why none can convincingly capture the nature of the intended experience. A problem common to all is that they ignore how *attending* to musical properties is a key component of hearing musical expression.

Chapter 3 presents my view on the problem of musical expression. My intended argument is developed based on analyzing the notion of *musical movement*. I argue that the phenomenological observation that we perceive a movement, flow, or continuity among musical events (or ingredients) in a piece of music is the basis of hearing musical expression. The idea is that in order to make sense of musical movement, flow, and continuity, the listener gives a feeling response to music – or, as I put it, moves with the music. An upshot of this argument is that perceiving musical movement has a feeling component, and this feeling contributes to our attribution of emotions to musical works or passages.

In the last chapter, Chapter 4, I examine how sounds and their sources are related to each other. By appealing to different considerations about the phenomenology of hearing and empirical studies on sounds and sonic properties, I argue that the best candidate for the above problem is a version of the property view: that sounds are properties of things in the world. However, I propose that sounds are not properties of objects—as colors are—but are instead complex properties constituted by properties of objects *and* properties of *events* those objects participate in.

Bibliography

Davies, S. (1980). The Expression of Emotion in Music. *Mind*, *89*(353), 67–86.

Davies, S. (1994). *Musical Meaning and Expression*. Cornell University Press.

Goehr, L. (2007). *The Imaginary Museum of Musical Works: An Essay in the Philosophy of Music*. Oxford University Press.

Juslin, P. N. (2019). *Musical Emotions Explained*. Oxford University Press.

Kivy, P. (1981). *The Corded Shell: Reflections on Musical Expression*. Princeton University Press.

Kiani, M. (1999). Sadness in Iranian Music. *Monthly Book of Art*, *7*, 28–28.

Langer, S. K. (1954). *Philosophy in a New Key: A Study in the Symbolism of Reason, Rite, and Art* (3rd Edition). Harvard University Press.

Levinson, J. (2011). What a Musical Work Is. In *Music, Art, and Metaphysics:Essays in Philosophical Aesthetics* (pp. 63–88). Oxford University Press.

Rice, T. (2014). *Ethnomusicology: A Very Short Introduction*. Oxford University Press.

Chapter 1: Musical Aboutness and the Case of Musical Emotions

1. Musical Aboutness: A Demarcation of the Problem

Can music be *about* anything? To answer this question, we need first to specify what kind of music we are talking about. A general distinction can be made between two kinds of music. The first, *representationally-aided music*, is accompanied by words or some sort of representational content such as a descriptive title, a narrative text, or a scene. The other kind is *instrumental music*, music which is unaccompanied by any sort of representational content.[4]

Let us consider an example of each kind. Consider Antonio Vivaldi's *Four Seasons*. What is this work about? "The four seasons," one might quickly respond. And this seems obvious: according to the titles given by the composer for the work and each movement, the first movement is about the spring, the second about the summer, and so on.[5] Now, consider a work of instrumental music: Franz Joseph Haydn's Symphony No. 93 in D Major. There is no title or text provided by Haydn that could relate the symphony to an extra-musical thing—something in nature, a person's name, a place, etc. What is this work about? How could we relate the sonic patterns of the work to concepts or things in the world? Unlike the previous example, this time we cannot answer these questions with certainty or might even be inclined to say that this symphony is not about anything. It is just beautiful music; we do not need to listen to it by thinking of anything extra-musical and somehow relating that to the symphony. One might even

[4] In the philosophical literature, the second kind has also been called *pure* or *absolute* music, for example, in Kivy (2009). However, considering that the term absolute and its historical and social connotations can be deeply controversial (see, for example, Bonds (2014), and Pederson (2009)). I use the term 'instrumental' for this kind and limit it only to a kind of musical work that is historically not connected to any descriptive title or narrative text by its composer. Defining it in this limited sense is not contrary to interpreting and analyzing this type of music in the relevant historical contexts and with the help of extra-musical connotations connected to such works. (See, as an example of this sort of musical analysis, Bonds (2006).)

[5] For simplicity's sake, I am only considering the title of this work and not the sonnets Vivaldi included with each movement. (I am thankful to Sarah Ellis for pointing this out to me.)

think that listening in this way would be inattentive and non-musical, distracting us from the work itself.

The way we talked about the above two examples shows one sense of musical aboutness: musical association or reference, that is, whether and how music can refer, or become associated with, things in the world. As regards this sense, I will hereafter use *musical association* and *musical reference* interchangeably. (I should clarify that by "association" I do not mean those private associations a piece could have for one listener, the things the piece reminds the listener of in her private life, or what are also called the "subjective meanings" of the piece (Koopman and Davies (2001); Kivy (2007); and Higgins (1997)). I will return to this matter in Part 2 of the dissertation.) Through our examples we saw that in the case of representationally-aided music like *The Four Seasons*, talking about a work's association(s) is quite straightforward compared with instrumental music, like Haydn's Symphony No. 93.

A second sense of musical aboutness is the following. Consider again *The Four Seasons*. Although with the help of the titles we can securely associate, for instance, the first movement with spring, there remains one important question to be answered: what exactly does the first movement *say* about the spring? Consider this linguistic example: "This year we had a rainy spring." This sentence refers in part to the spring of a particular, intended year; but beyond that it also says something about it: that that spring was rainy. In this example, the reference comes with or is put into a linguistic content (or propositional content). We can, analogously, ask a similar question about our musical example, the first movement of *The Four Seasons*: does the music provide some *content* for the association it makes with spring? Here, giving an answer becomes very difficult or even, some might say, impossible. How can we know what exactly the movement is supposed to say about spring? Although as regards musical association there is a

title, given by the composer himself, which justifies our associating each movement with a season, we do not have any texts for each movement of the work. So, any answer to the above question would be a matter of "musical interpretation," that is, translating or "decoding" a music unaccompanied by any detailed narrative text into a linguistic interpretation—and this is a highly controversial matter. In comparison, it is not surprising that getting to such a musical interpretation for a work of instrumental music would be even more challenging and, indeed, controversial[6].

I want to set aside for the moment the deep issues regarding musical aboutness and instead focus on the demarcations I provided in this regard. My adumbration of the subject shows that we need to think of two important senses of whether music can be about anything: musical association (or reference); and musical content, that is, what music can say about the associations it can make.[7] Moreover, the considerations may differ depending on whether we are considering instrumental music or representationally-aided music. So four areas can be demarcated for the problem of musical aboutness, each of which has its own specific issues.

For example, concerning musical association for the case of representationally-aided music, an interesting question is how knowledge of a composer's intended association shapes or influences the listener's musical appreciation. Or, with regard to musical content for this kind of music, an issue can be whether and how music could *add*, content-wise, anything to the representational content that accompanies a work. (For example, concerning a work of program

[6] For example, Kivy (2009) has elaborately argued against such interpretations of absolute music. (See also Zangwill (2014)). For a response to a part of Kivy's arguments, see Puolakka (2017). Historically, the arguments against attributing "non-musical" meanings to music is rooted in Hanslick (1986).

[7] This distinction is similar to the one used in the philosophy of language between what a sentence/utterance is about (its topic or subject matter) and what it says (its content).

music like Franz Liszt's *Faust Symphony*, what content exactly does music add to the accompanying program?)

Delving into musical aboutness for representationally-aided music could lead one to give the primary role in a work's being about things in the world to, say, the text accompanying the work and consider the music itself as having a secondary, or supporting, role in that relation. Taking *The Four Seasons* again as an example, it is very tempting to say that the burden of reference-making is primarily on the piece's descriptive titles. Or, for a Johann Sebastian Bach cantata it might reasonably be said that the associations and (semantic) content are primarily given by the accompanying religious texts; the contribution of music is, inter alia, to make those associations and (linguistic) content more affective or psychologically influential, one might say.

Examining musical aboutness may be more illuminating and musically revealing if we focus on instrumental music. In this way, we can get to the core of the problem: whether *music* is able to be about anything at all. If we reach a negative answer to this question, we will be more certain in holding that as regards representationally-aided music, it is an accompanying representational content that underlies a work's being about extra-musical phenomena, with the music itself having a secondary and affectively supporting role. But showing that instrumental music can be about extra-musical phenomena in its own musically specific way would also show that in the case of representationally-aided music, a part of a work's being about extra-musical subjects could be musical and not, say, primarily linguistic.

Accordingly, in this dissertation my focus will be on instrumental music when it gets to how music can be about any extra-musical matter. Regarding the two senses of the problem I considered above (musical association and musical content), my main focus will be on musical

associations in order to explore the mechanisms through which music can go beyond itself and refer to things in the world. So, on most occasions when I talk about musical aboutness or a work's being about some extra-musical subject, I will specifically mean how music can be associated with or refer to the subject. Although I will not focus on musical content, I will try, at points, to show how my arguments bear on this matter as well. I believe that approaching the problem of musical content requires a convincing solution to the problem of musical associations. Moreover, my focus will particularly be on whether and how music can be about human emotions, to which I want to turn now.

2. Emotions and Music

As a few examples of the kinds of associations music might be able to make, one can think of the religious aspects of some of Bach's music (Marissen (2016)), some music's political associations (Goehr (1998) and Taruskin (2019)), the religious and mystical aspect of some Persian classical music (During (1991)), or whether music can have moral dimensions (Kivy (2009, chap. 10). However, the musical association that has been explored far more than any other is with emotions, or our emotional life connected with music. This can, at least in Western thought, be traced back to Plato's work and the ancient, Pythagorean ways of thinking about music.

There are psychological factors behind this matter, that is, our tendencies to relate music, in one way or another, to our feelings and mental states. As psychologists have mentioned, "[o]ne of the main reasons people engage with music is to experience emotions" (Peltola (2017, 82)). It is commonplace to attribute different emotional and psychological qualities to music ("This piece is cheerful," "This part is anguished," etc.), or to describe music in terms of how it makes

us feel emotions or puts us into different moods ("This is a sad piece," "This work really cheers me up," etc.). One thing all such emotional descriptions of musical work and passages reveals is that while listening to expressive music, we relate it to our emotional life in one way or another. The central philosophical question with regard to this matter, in Stephen Davies's terms, is: "how can [pure/instrumental] music express emotion when it is abstract and insentient?" (2011a, 8).

In the three chapters of this dissertation focusing on musical emotions, I zero in on how we experience the relationship between music and emotions, or psychological states more generally. I examine the related philosophical theories that try to explain *musical expression* or *musical expressiveness* (there is a subtle difference between these phrases which will be explained later). This examination will enable us to delve into the most important issues surrounding musical association and experience. In particular, it is a deep and, of course, controversial matter whether or not we should construe the relation between music and emotions as an instance of musical association, that is, as an example of how music can refer to or be associated with an extra-musical phenomenon: our emotional life.

2.1. Emotions *in* musical works

In this section I take up the relation between emotions and music and examine an influential theory that construes musical emotions as properties of musical works. In the next chapter, I will examine views that explain the relation between emotions and music by focusing on musical experience. But two matters first. By "relation" I do not mean reference; I do not want to assume that the relation between music and emotion is necessarily one of reference. Any conclusion about this should be illuminated by examination of the relevant theories. I also want to mention a terminological matter: by "musical emotion" I mean those emotional qualities that are sometimes

attributed to music; and by the *experience* of musical emotions I mean the experience through which a listener attributes emotional qualities to music.[8] All the theories we will consider are attempts to explain such attributions.

According to the *appearance theory*, first defended by Peter Kivy and Stephen Davies,[9] musical emotions are *properties* of musical works. In other words, an expressive passage of music *possesses* an emotional property, as a red apple possesses the property of redness. This theory was proposed mainly as a response to two other theories of musical emotions: the arousal theory and the expression theory. According to the expression theory, the emotion a musical passage is expressive of is the felt emotion of the composer while composing the work. And according to the arousal theory, a musical passage is expressive of an emotion when it arouses that emotion in the listener. Based on both of these theories musical emotion is a *felt* emotion: the expression theory claims it is the felt emotion of a composer, and the arousal theory claims it is the felt emotion of the listener. A problem with the expression theory is that it is not obvious that a composer must feel all the emotions she wants her work to be expressive of. And a problem with the arousal theory is that it is not obvious that to recognize that a passage is, say, sad, the listener must feel sadness.[10] A problem common to both the theories is that neither is able to explain expression in virtue of *musical* features: they merely appeal to the felt emotion of either the composer or the listener. The appearance theory holds, in contrast, that the attribution

[8] It is important to notice that by the experience of musical emotions I do not necessarily mean *feeling* such emotions.

[9] Kivy (1981) and Davies (1980, 1994). (At the same time but independently of each other – their versions are a bit different.)

[10] The more sophisticated versions of the arousal theory (such as Matravers's (2003)) have tried to solve the mentioned problem. The examination of these theories will be the subject of the next chapter.

of an emotional quality to music is a *recognition* or appreciation of some qualities *in* the music; and because music is not a psychological being, the qualities recognized cannot be mental states but some perceptual properties in or possessed by music. I am going to evaluate how the appearance theory corroborates its core idea that musical emotions are perceptual properties of musical objects/works, considering that we intuitively tend to regard emotions as psychological, mental states *felt* by beings having a psychology.

The theory's main approach is to distinguish between two cases in which we attribute emotional qualities to things in the world: (a) sometimes such attributions refer to a mental state of a psychological being, like when one says, "Sina is sad today," or "What a cheerful child!"; and (b) sometimes emotional attributions are references to or descriptions of the perceptual features presented in the appearance of an object, like when one says, "A Saint Bernard has a sad face," or "I am seeing a happy face in the cloud." The appearance theory describes the phenomenon in case (a) as *expression*, when a psychological being expresses a felt emotion, and the phenomenon in case (b) is called *expressiveness*, when an insentient being is expressive of an emotion via its perceptual characteristics (Kivy 1981, chap. 2). The theory purports to show that musical emotions should be understood as a case of (b) and not (a); in other words, they are cases of expressiveness and not expression. Kivy (1981, 12) says:

> The Saint Bernard has a sad face [...]. We do not mean to say by this that the Saint Bernard's face *expresses* sadness. For certainly the Saint Bernard is not *always* sad. [...] When, therefore, we describe the Saint Bernard's face as a sad face, we are not saying that it *expresses* sadness, but, rather, that it is *expressive* of sadness.

In Kivy's view, the sadness of a Saint Bernard's face is not a felt emotion but a perceptual quality of the dog's face; and, accordingly, the word "sad" in describing the dog's face is a description of the face as an object and not of any mental state or felt emotion. This is a case in

which (the usage of) an emotion term is disconnected from any psychological state and shows how an object through its perceptual features can present an emotional quality. The perceptual features are related to a contour or configuration in the dog's face that makes an onlooker perceive the dog's face similar to a sad human face. Such perceptual features are what Davies calls "emotion characteristics in appearance":

> My point is that emotion words have a common, licit use that does not involve even implicit reference to the occurrence of feelings. Although "He is a sad-looking person" may be used to refer to a person's (present or predicted) feelings, I distinguish the no-reference-to-feeling use as a distinct one. It is in this latter case that emotion words refer to emotion characteristics in appearances. (1994, 224)

And:

> [...] the appearances that display emotion characteristics are not always human appearances. [...] Basset hounds and Saint Bernards are sad-looking dogs. Presumably they do not feel sad more often than any other breed of dog. The sadness lies in their appearance and not in their feelings. (1994, 227)

The upshot is that expressiveness, unlike expression, can be attributed in the absence of mental states. So a possibility remains that although musical works cannot be in any mental state, they might be expressive of emotions. And the appearance theory claims that music, like a Saint Bernard's face, can be expressive of emotion by presenting features that bear perceptual similarities to human emotional appearances. According to the appearance theory, attributing emotions to music—such as "The beginning of Gustav Mahler's Second symphony is anxious"—is not completely but basically similar to emotional attributions to (non-musical) objects—in descriptions like "a sad willow tree," "See those two happy faces in the clouds," or

"Basset hounds have sad faces." The suggestion is that all such descriptions refer to perceptual qualities of their intended objects and not to any mental state or psychology.[11]

But what perceptual qualities of music are appealed to when an emotional description of it is given? According to the appearance theory, musical elements have the potential to present the appearance of human emotional behavior or contour. For example, a slow tempo could be a presentation of the slow movement of sad people; the timbre of instruments like the cello could sometimes mimic a melancholy (low) human voice; an energetic rhythm could be a presentation of the typical behavior of cheerful people; and so on and so forth. Importantly, this expressive potential is independent of the feelings of either the listener or the composer-musician. We *hear* emotional qualities in music as one of its audible properties (like melody, rhythm, timbre, harmony, etc.).[12] And now, I want to examine a few consequences of such a viewpoint.

2.2. How Is It *Experienced*?

If musical emotions are to be understood as perceptual properties of musical works as claimed in the appearance theory, we need to consider what this entails for the experiential or phenomenological side of the problem of musical emotions. I am in particular thinking of these questions: what is experiencing emotional qualities *in* music like? What is the nature of the appreciation of such qualities of music? And what exactly in our musical experience makes us describe music in emotional terms? I intend to see whether these questions can be fully answered based on the core idea of the appearance theory. (Through this consideration we will identify an

[11] In a more recent defense of his theory, Davies (2011b) argues that an emotion term can have polysemous meanings: sometimes they are used to describe the appearance of objects and sometimes they refer to mental states.

[12] To Kivy, musical expressiveness is a formal part or property of a work and has formal functions in the work's whole formal structure (like the ways in which a musical motif or the harmony of a part have formal functions). See Kivy (2002, chap. 6).

evaluative criterion for the other views on musical emotions we will discuss in the next chapter.) But I first need to clarify what I mean by the experiential or phenomenological dimension of musical expression.

I use "phenomenological dimension" as opposed to, what I will call, the *causal dimension* of musical emotions. The distinction I make is between what it is like to experience a musical emotion, let us say musical sadness, and what musical and non-musical properties that experience *depends on*.[13] This dependence relationship can metaphysically be understood in different ways. We can say some musical and non-musical properties *cause* the experience of musical emotions; or, to say it more neutrally and without assuming that the dependence is necessarily causal, it should be said that what it is like to experience musical emotions supervenes on the related musical and non-musical properties. Because in the following discussions I am not concerned with the metaphysical side of the dependence relationship, I interchangeably use supervenience and causality to describe it and use the term "causal" in a broad sense including different forms of dependency. So, *the causal dimension* of the problem of musical emotions is related to the musical and non-musical properties the experience of musical emotions depends on, and the *experiential* or *phenomenological dimension* of the problem to the experience itself or what having that experience is like for listeners.

Two things about this distinction. First, with regard to the causal dimension, I say musical *and* non-musical properties, because it is quite indisputable that besides musical properties (melodic pattern, harmonic quality, timbral quality, etc.) there are non-musical factors and properties involved in our general experience and cognition of musical qualities; and it is

[13] The former is what is called *qualia* in the philosophy of mind, in "the sense of there being something *it is like* to undergo an experience, something one can normally introspect" (Balog (2009), p. 293).

plausible to say that our experience of musical emotions is a part or aspect of our musical experience generally construed. Three main categories for such non-musical properties are: (1) listeners' biological and neurological features connected with processing musical sounds;[14] (2) the higher cognitive features of listeners connected with music listening (for example, whether and how they can exercise their imagination while listening to music);[15] and (3) the characteristics of the cultural context in which music is listened to and appreciated.[16] The first category contains those deep-seated properties that influence music listening totally unconsciously. Although the characteristics in the second one are rooted in those deep-seated properties, they can also be formed and developed in different ways. For example, a person's cognitive ability to recognize melodies in musical objects might be developed in different ways depending on how she has musically been raised and what musical idioms she has been exposed to from childhood. And the characteristics in the third category can be arguably vast. For example, it might be the case that a person's different non-musical experiences in her culture influence and shape how she experiences musical emotions. I will later return to these specific categories, particularly the second and third ones, and how they can be related to each other and to the experience of musical emotions. For now, what is important to keep in mind is that when we talk about the causal dimension of musical emotions, we mean both musical and non-musical properties responsible for how musical emotions are experienced.

The second point about the distinction I made between the causal and phenomenological dimension is as follows. The way I construe this distinction is in line with the distinction drawn

[14] For example, there are studies on neurological factors involved in musical imagery; see, Zatorre, and Halpern (2005).
[15] See, for example, Jackendoff, and Lerdahl (2006).
[16] See, for example, Morrison, Demorest, and Pearce (2019).

by the philosophers of mind between mental/epiphenomenal *qualia* and their related physical properties. In other words, the distinction between what the experience of a musical emotion is like and all the musical and non-musical properties that experience supervenes on is seen to be basically, even if not in all respects, similar to the distinction between what the experience of, for example, pain is like and all the physical properties and facts pain supervenes on (say, some specific chemical processes in the brain). In fact, my major concern with the causal-phenomenological distinction as regards musical emotions is that each dimension needs to be reflected on *distinctly*, though they are not unrelated. I think in understanding musical emotions we need, at the outset, to keep in mind that our intended problem has two dimensions each of which might have different characteristics. It is one thing to focus on characterizing what the experience of musical emotions itself is, and the other to discover the causes of that experience (or what that experience supervenes on).

Distinctions fundamentally akin to what I have portrayed have been pointed out by others as well. For example, Christopher Peacocke says:

> The question of what it is for emotion or other states to be heard in music is not the causal or computational question of how it comes to be heard. It is not the question of the social influences on how we hear music. Nor is it the question of the evolutionary explanation, if such there be, of the existence of such perceptions. It is the constitutive question, the 'what-is-it?' question [...]. (2009, 257)

Also, concerning hearing tension in music, Kendall Walton has remarked:

> It is one thing to identify the sources of tension in music, and quite another to say what musical tension is. What is it that harmonic dissonances in tonal music or Stravinsky's syncopations or Beethoven's driving rhythms are responsible for? (1999, 409)

In these passages, a distinction can be seen between what *causes* a listener to hear an emotional or psychological quality in music and what such a hearing experience *is* or what constitutes this experience. Peacocke distinguishes between *how* a musical emotion comes about

and *what* hearing it is; and Walton distinguishes between what is responsible for musical tension and what hearing that tension itself is like.

And at the end, I want to highlight a point about this distinction that will be important in our future discussions: the fact that the causal dimension can mostly be *hidden* to listeners while they are experiencing musical emotions. Remember that the causal dimension contains both musical and non-musical properties. Regarding non-musical properties it is not hard to see how such properties responsible for musical emotions can be totally hidden during the listening process. For example, many cognitive processes involved in listening to music, say processing or responding to rhythm, are hidden to a listener. What I want to highlight is that this also applies (at least in part) to *musical* properties involved in experiencing musical expression. For instance, in Western tonal music, dissonant chords in the harmonic fabric of a work brings about tension or can, if properly used, give a tense quality to music. Nonetheless, when a listener attends to a musical passage containing chords with dissonant intervals, she might experience a tension without knowing anything about the cause of it, that is, without knowing that the tension is due to the dissonant chords. Although the causal dimension of her experience is hidden from her, she is aware, phenomenologically speaking, that she is experiencing tension or a tense musical passage. As an example, in the second movement of Ludwig van Beethoven's Piano Trio No. 2 in G major, Op. 1 no. 2, the dissonant chords in mm. 41-46 (Example A) imbue the music with a tense quality. A listener might be unaware that the cause of musical tension is in part these chords while being able to appreciate the (musical) tension itself. I will refer to this aspect of the causal-phenomenological distinction as the Causal Opacity Thesis (COT), which can be formulated from two perspectives:

(COT): (a) the causes of experiencing musical emotions might or might not enter the phenomenology of the experience; or, from a different perspective: (b) the phenomenal character of the experience itself might or might not hinge upon an awareness of any particular cause of the experience, either musical or non-musical.

The Causal Opacity Thesis proposes that there is some level of independence between the causal and phenomenal dimensions of the experience of musical expression. This will be important in evaluating a theory of musical expression. Whether a theory can explain musical expression in a cogent manner depends on how it treats the causal-phenomenological distinction specified in COT. But we need to see it particularly in the examinations of some of the main theories of musical expression developed in the philosophical literature. Let's return now to the appearance theory whose central thesis we explained in the previous section.

2.3. Is the Resemblance Experienced?

Remember that according to the appearance theory, musical emotions are considered to be properties of musical works. Let us focus on an important aspect of the theory we did not examine earlier: that the central claim of the appearance theory is based on the notion of *similarity* or *resemblance*. The theory claims that the emotion a musical passage is expressive of is a property of music because the musical structure of the passage is capable of presenting an appearance or contour similar to the appearance of expressive human behavior (a sad face or an exuberant behavior, for example). Connected with this claim is the idea that a listener experiences musical expression when she *recognizes* such a property *in* music. In this section I want to examine what role similarity or resemblance—in the way suggested by the appearance theory—plays in a listener's experience of musical emotions.

Davies believes that the emotional property of music is *response-dependent*: "By calling it response-*dependent* I mean it is the concept it is in virtue of its ability to produce a certain characteristic response in creatures of an appropriate kind under suitable conditions. We might think of response-dependent properties as powers things have to produce certain kinds of responses, including experiences." (2011a, 8) The idea, in fact, is akin to philosophers' use of the phrase "secondary quality" for simple properties of objects, like color. Color perception is the result of how an object in suitable lighting conditions is perceived by a perceiver with a specific perceptual apparatus; but this does not prevent us from saying that color is a property *of* an object. After all, when we talk about colors of objects we are not referring to things in our head; we refer to objective (or inter-subjective) properties of objects themselves. A secondary quality of an object is response-dependent but, at the same time a property possessed by the object itself. Davies thinks that the same idea is applicable to the emotional or expressive properties of music: while they are response-dependent, they are properties of music.

We should, however, notice that there is an important difference between a secondary quality (or simple properties like color) and the emotional properties of music: detecting the latter is more complicated, requiring attention and, in many cases, relevant acculturation or trained cognitive abilities. Recognizing the expressive qualities of, say, a Mahler symphony is a matter not merely of hearing the sounds in the work but of attentive listening to different parts of the work with sufficient musical sensitivity, which can be attained either by musical training or by listening a few times to a particular work while having enough listening experience of the works of the same kind of music (say, a good number of works in Western classical music). To put it differently, perceiving a color does not require training or sophisticated perceptual and cognitive ability, but recognizing the emotional quality of music usually does.

So, the expressive qualities of music are response-dependent like color, but are more complex in terms of the kind of response they need in order to be recognized. One reason for this complexity is that the expressive qualities of music are structural or holistic properties that are themselves based on more basic, audible properties of music. I think Davies (1994, 256) also agrees when he says: "[...] musical expressiveness is an aspect (or an emergent property, or a supervenient property) depending for its character on the structure of musical movement."

I think one important upshot of the above consideration regarding the response-dependence of musical expression is this: the emotional qualities music can be expressive of are not the result of individual musical elements (rhythm, harmony, melody, etc.) separately or independently, but dependent on all of them together in relation to each other. Emotional properties are holistic properties of musical objects.[17] Consider this example. The beginning moments of Gustav Mahler's Symphony No. 2 "Resurrection" (see Example B) have an anxious quality. What imbues this music with such a quality is revealed if we look at the basic musical components of the passage.[18] Let us pick first the dynamics. The passage opens with a *fortissimo* chord in the upper strings (the violins and the violas) which decrescendo to *piano* by the end of the measure, creating sudden change in the dynamics in the first measure. We have another change in the dynamics at the beginning of the second measure: the cellos and the double basses enter with the first theme of the movement at *fortissimo*. There are also more changes in the dynamics in the subsequent measures (noticeably in measures 8, 9, and 10). Now, let us investigate another

[17] We can say if S is the set of all the basic audible properties of an expressive passage P and E is the related expressive quality of P, then E supervenes on S if and only if any qualitative difference in E is not without some qualitative difference(s) in S and vice versa. (See Leuenberger (2008) for the metaphysical dimensions of such supervenience claims.)

[18] I say "in part" because I do not want to assume that the potential of musical elements to make basic similarities to human behavior is the only cause of experiencing musical emotions.

musical element. Starting in the second measure, the violins and the violas play *piano* tremolos over the primary theme in the basses and violas. This playing technique, as "a rapid reiteration of a single note or chord without regard to measured time values taken independently,"[19] can impart a trembling, expressive quality.

Combining all of the formal, audible properties of the passage together, we can recognize more clearly at least *a part* of the cause of the listener's attribution of the anxious quality to the music. The sudden changes of the dynamics at different moments along with the tremolo playing in the background imbue the music with a kind of instability, nervousness and uncertainty. A main point here is that none of these formal properties of the musical elements taken separately is responsible for the anxiety of the passage: tremolo playing could, for example, increase the cheerfulness of a passage, or sudden changes in dynamics could cause surprise. But in the context of this passage each of these basic properties is playing a role to make the music nervous, anxious, or uncertain. It is in this sense that the emotional quality of the passage is a holistic property (or supervenes, at least in part, on those basic properties).

An important implication is that recognizing emotional qualities in music typically requires *structural attention*, a skill acquired through musical acculturation. In most cases, we passively recognize the color of objects. But the recognition of expressive qualities, which are also response-dependent, requires careful attention to musical properties and how they together can result in holistic emotional qualities.

The appearance theory teaches us three key lessons: (1) the emotional qualities of music are response-dependent; (2) they are holistic properties which are dependent (at least in part) on more basic musical properties; and (3) recognizing them requires careful structural attention.

[19] Oxford Grove Music Online (https://doi.org/10.1093/gmo/9781561592630.article.28314)

These ideas are connected with what I call the *recognitional claim*: in experiencing musical expression a listener *recognizes* the expressively relevant properties *in* music.

The appearance theory also makes the *resemblance claim*: that recognizing emotional properties in music involves *perceiving a resemblance* to human expressive behaviors. In Kivy's (1981, 53) terms, "[...] we hear sadness in this complex musical line [a passage from Bach's *Mass in B minor*], we hear it as expressive of sadness, because we *hear it as a musical resemblance* of the gesture and carriage appropriate to the expression of our sadness" (emphasis added). Or, as Davies (1994, 229) says: "I believe that the expressiveness of music depends mainly on a resemblance we perceive between the dynamic character of music and human movement, gait, bearing, or carriage." Let us zero in on the resemblance claim.

I will argue that however illuminating and plausible the three points extracted from the recognition claim of the theory are, when the resemblance claim comes into play, the main idea of the theory becomes untenable. My main concern pertains to the causal-phenomenological distinction I draw earlier with regard to musical emotions. As a causal claim the recognitional claim is plausible. It seems plausible to say that there are properties in expressive music the recognition of which causes (or at least partly causes) a listener to attribute emotional qualities to it; and it also seems a plausible idea that such musical properties have resemblances to human emotional behavior in one way or another (for example, a resemblance between a fast rhythm and a jaunty gait). But this does not necessarily and automatically entail the resemblance claim: that in hearing musical expression the listener *perceives a similarity* between musical structure and the appearance of human expressive behavior. This is a phenomenological claim which,

considering the causal-phenomenological distinction, needs to be taken independently of the causal claim of the theory. The problem with the theory is then specifically phenomenological.[20]

If there is a perceived similarity between musical appearance or contour and human expressive behavior, then, as Jerrold Levinson has pointed out about Davies's theory, "since everything is similar to everything else to some degree, the issue [...] becomes one of *how* similar such an appearance must be to one presented by human behavior in order to constitute an emotion-characteristic-in-sound of the emotion in question." (1996, 97)[21] What Levinson rightly remarks here seems to me to be itself a twofold issue: (a) why does the listener compare the appearance of music to *human* behavior instead of something else ("since everything is similar to everything else," as Levinson says)? (b) Even if the comparison is made to human behavior, in which respects does the listener perceive them to be similar to each other ("*how* similar" are they)?

Kivy has a response to (a). To him, the reason why we identify a similarity between musical contour and human behavior (or voice) is our deep-seated biological tendency to animate non-sentient objects ("Tie a piece of cloth around the handle of a wooden spoon and a child will accept it as a doll" (1981, 57)). We respond to natural objects in this way on an evolutionary basis "because," holds Kivy, "this way of seeing has survival value." (1981, 57-58) And Kivy thinks, in a basically similar fashion, we animate sounds in music and perceive musical contours as human expression (behavior or voice).

[20] Although not based on a distinction between the phenomenological and causal dimension of musical expression, in the same vein, Aaron Ridley (1995) has remarked that recognizing a resemblance between basic musical properties (what he calls "millesma") and human behavior is not tantamount to experiencing expressiveness in music.

[21] For a similar criticism, see also Stecker (1999).

Kivy's idea is mostly an empirical speculation whose truth depends on facts about human biology and evolutionary history. It might in fact be true. But my worry is that even if it was, I am not sure it could solve the problem raised in (a) against the theory. How exactly should we take Kivy's suggestion? Does it say that we listeners are biologically disposed to animate a musical object? Animation is typically experiencing an inanimate object *as* an animate being. Taking Kivy's example, when we see a doll we experience (imagine, think of, consider, etc.) the object as a person. Do we in the same fashion animate a musical object, that is, experience it as a psychological being? We might. But even if such an idea could serve as the basis of a theory of musical expression,[22] that theory would not obviously be the appearance theory. According to the appearance theory, the expressiveness is a perceptual property of musical object and a listener perceives that property in music. The theory does not construe a listener, in her experience of music, as going beyond the perception of the (perceptual) appearance of a musical object. If it did, it would need to give us more explanation of how that "going beyond" happens: is it imagination, imaginative perception, perceiving-as, or some other mode of experience? So, the issue is that it is not clear how we should incorporate animation in the framework of the appearance theory. (And if we wanted to flesh out animation with notions like imagination, that would not be the appearance theory.)

The second issue (b) related to the resemblance claim was this: even if there is a similarity between expressive musical contour and expressive human behavior, how exactly is this similarity experienced by the listener? It seems that to recognize such a resemblance intended by the theory, some form of cross-modal comparison is required. Standard cases of perceptual resemblance are related to one sense modality. I see a new person and I recognize how his face is

[22] For example, Saam Trivedi's (2011) theory is based on such an idea. He thinks that in hearing musical emotions we imagine the musical object itself as a psychological being.

similar to the face of a friend of mine. This is a visual comparison. Or I hear a sound produced by a machine and I recognize how it is similar to the sound of a particular bird chirping every morning near my apartment. This is an auditory comparison. But comparing the taste of a food to visual characteristics of a particular scene requires comparing different sensory modalities. The same applies to expressive music. Musical experience is basically auditory. Cross-modally comparing the auditory properties of music with human expressive posture, gesture or behavior, which we *see* (or *feel* from the inside, so to speak), cannot be merely perceptual like recognizing perceptual resemblance happening in one sense modality.[23] It's not that we cannot find structural isomorphism in cross-modal cases. But recognizing such isomorphism does not appear to be mere sensory perception—it seems to involve other cognitive faculties as well. (One might say here that it requires imagination. It might, but this is not what is claimed by the theory we have been examining. A theory appealing to imagination becomes a different type of theory a few main versions of which will be examined in the next chapter.)

Let me recap our discussion of the appearance theory. The theory involves a recognition claim and a resemblance claim. According to the former, experiencing musical emotions requires recognition, and this recognition is in fact the result of a listener's perception of the expressively relevant musical properties. According to the latter, the experience of musical emotions is in fact perceiving a similarity between musical structures and expressive human behavior. I argued that the trecognition claim is plausible in the sense that experiencing musical emotions needs a listener to attend to musical structures and to the musical properties that are responsible for musical expression. Also, the claim that musical structures have a capacity to resemble or

[23] Here, I am only considering those expressive human behaviors that are basically visual and not those that are audible like the voice of a sad person. The resemblance that music can have to the latter of course does not require cross-modal comparison, to which the issue I am raising is not applicable.

suggest analogies with human expressive behavior does seem plausible. We saw that how musical properties like tempo can suggest such analogies, and this may reveal a cause behind our experience of musical emotions. But I argued that this does not entail that experiencing musical emotions—or hearing emotions in music—is perceiving similarities between musical structures and human expressive behavior. So, the problem, which is phenomenological, is with the resemblance claim. The theory cannot explain why one perceives a resemblance between sonic patterns and human behavior, instead of many other things or phenomena such patterns can be similar to, and it cannot explain how exactly one cross-modally compares such patterns, which are audible, with expressive behaviors many of which are visual. These concerns make the idea of perceiving resemblances in music phenomenologically nebulous.

3. Relevance to Musical Association

How is our discussion of the appearance theory related to musical aboutness? Let us first think of the core idea of the theory from the perspective of meaning. The idea is that the emotions music can be expressive of are properties of music (or properties of musical works). Can an object be *about* a property it possesses? Is a red apple about redness? It seems not. Typically, we do not consider objects to mean something about (or refer to) the different properties they possess. We do not think of a whiteboard to be about (or refer to) being white or rectangular; it is just white and rectangular. We might use an object as a sample to refer to and talk about a particular property it possesses. For example, a teacher can point to the rectangular shape of a whiteboard to explain rectangularity in general—to refer to the rectangular form in

general and not just the one possessed by that whiteboard.[24] But the point is that typically we do not think of the properties of objects in this way. Kivy holds this is the case with expressive music too. Considering that he construes expressiveness as a property of a musical work, he believes an expressive musical passage is not about and does not refer to the emotional property it possesses: "The Saint Bernard's face does not mean 'sadness' but possesses it. A sad theme does not mean 'sadness'; it is expressive of it, and the sadness is a 'sensual' property of the music. To describe music in emotive terms is to attribute properties to it." (1981, 118)

Davies thinks, however, that music's capacity to have an emotional property could involve reference to emotions. His main reason is that a musical work is not merely a natural object having different properties, like an apple being red, but it is an artifact made by people. As he says: "[...] Saint Bernards do not mean anything by their appearance, but a sculptor who chooses to make a statue of a dog may intend that expressive appearance to refer to the emotion it presents." (1994, 265) To give a non-artistic example, the situation is like when I use a red apple to say something about the color of red; in this case although the apple's redness, outside any context of use, does not refer to redness, in the *context* of my intentional usage of the object, the apple's redness can make a reference to redness in general. That is, the reference made here is due to my intentional use. Similarly, Davies thinks that in the case of music, composers sometimes *intentionally use* the perceptual properties of music to refer to emotions. "Even if the composer does not," says Davies (ibid.), "manufacture the expressive potential of the materials she uses, she does put the expressive potential to work." To him, in such cases, the perceptual

[24] Goodman (1976) calls this *exemplification*. In our example the whiteboard used in that way exemplifies rectangularity. He gives the example of a tailor's swatch. Although a small piece of cloth out of any context of use does not refer to the color it possesses, when it is used as a tailor's swatch, its color can be said to refer to (to exemplify) that color.

similarities of musical structures to expressive human behavior can be used to secure reference or aboutness.[25]

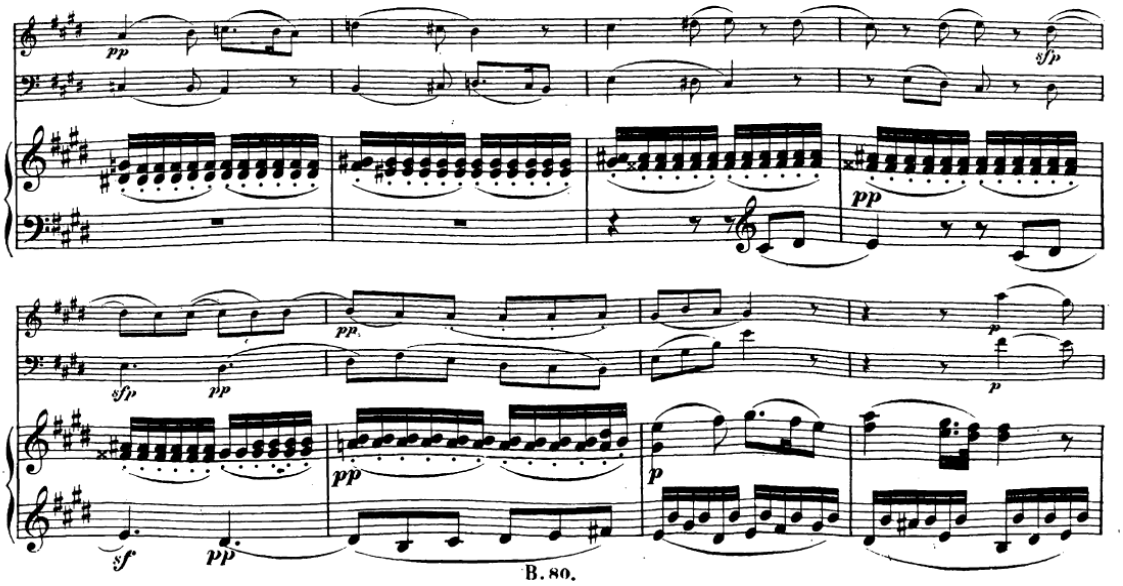
But there is an issue with Davies's idea. He does not say clearly how the expressiveness of a musical passage is dependent on the composer's intention. Is it always or sometimes the case that the emotional quality of an expressive passage is the result of the composer's intentional use of musical elements? It seems implausible to say that it is always, for two reasons. (1) We cannot always be certain about a composer's intention, even by using all the contextual evidence about a work. (2) Based on the central idea of the appearance theory itself it seems that a musical passage can sometimes be expressive regardless of the intention of the composer—for example, an energetic rhythm can be experienced as being similar to the appearance of a jolly, energetic behavior independently of the composer's intention. But if Davies's idea is that it is *sometimes* the case that the expressiveness is the result of a composer's intention, how can we distinguish when it is and when it is not? This is not explained in Davies's appearance theory. So, the main problem is that if we accept Davies's suggestion that expressive music (at least sometimes) refers to, or is about, emotions because musical expressiveness is usually the result of how a composer uses (the expressive potential of) musical elements, we need an account of when and how exactly expressiveness is linked with a composer's intention; but we are left with no explanation on this matter. Davies is then not able to give us a persuasive account of how music can be about emotions if musical emotions are, according to the appearance theory, properties of musical works.

[25] Davies's view on the referential capacity of musical expressiveness is somehow akin to Goodman's idea of *exemplification*.

In this chapter, after delineating the specific problem of this dissertation regarding musical aboutness, I focused on the relation between music and emotions to see whether we can construe this relation as a case of musical reference or not. Specifically, I examined the appearance theory according to which the emotional qualities we attribute to expressive music are perceptual properties possessed by music. The core idea of this theory is that expressive music has perceptual similarities with human expressive behavior. I argued that this theory cannot convincingly capture the phenomenology of experiencing musical emotions. My argument was based on a distinction I made between the causal and phenomenological dimensions of musical expression. Although considered as a causal claim the theory seems to be able to unravel a cause of musical expression, it cannot explain, in a phenomenologically convincing way, how a listener *experiences* expressive music as being similar to expressive human behavior. Also, we saw that in Kivy's version of the appearance theory it is denied that expressive music is about emotions while in Davies's version it is allowed, although not in a persuasive way, that expressive music can be about emotions. However, I want to leave a deeper examination of expressive music's being about, or referring to, emotions until after consideration of another class of theories on musical expression. In the next chapter, I will examine those theories that try to explain musical emotions by focusing not on perceptual properties of music but on the nature of musical experience.

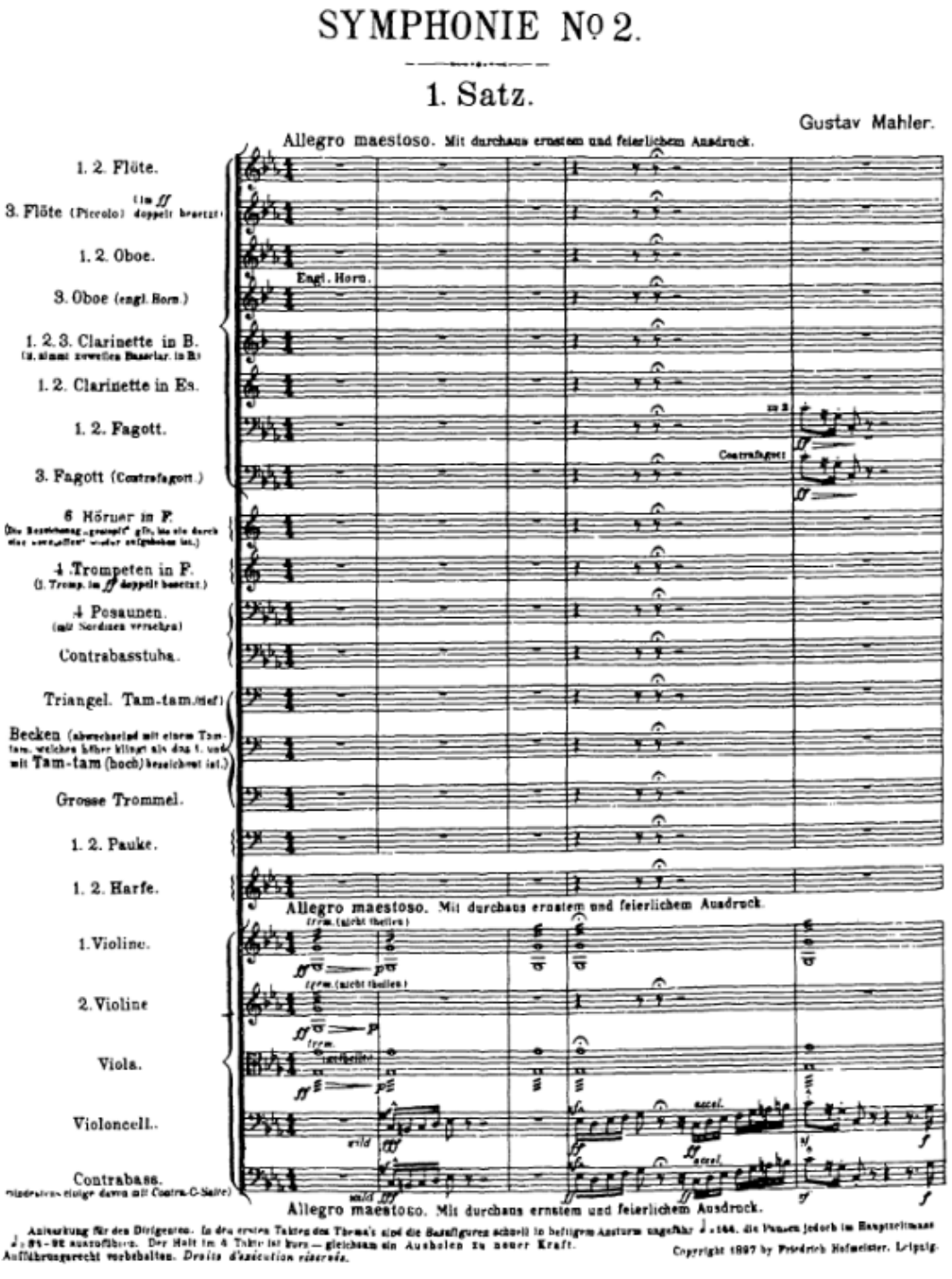
Musical Examples

Example A. Ludwig van Beethoven's Piano Trio, Op. 1, no. 2, second movement, mm. 41-48[26]

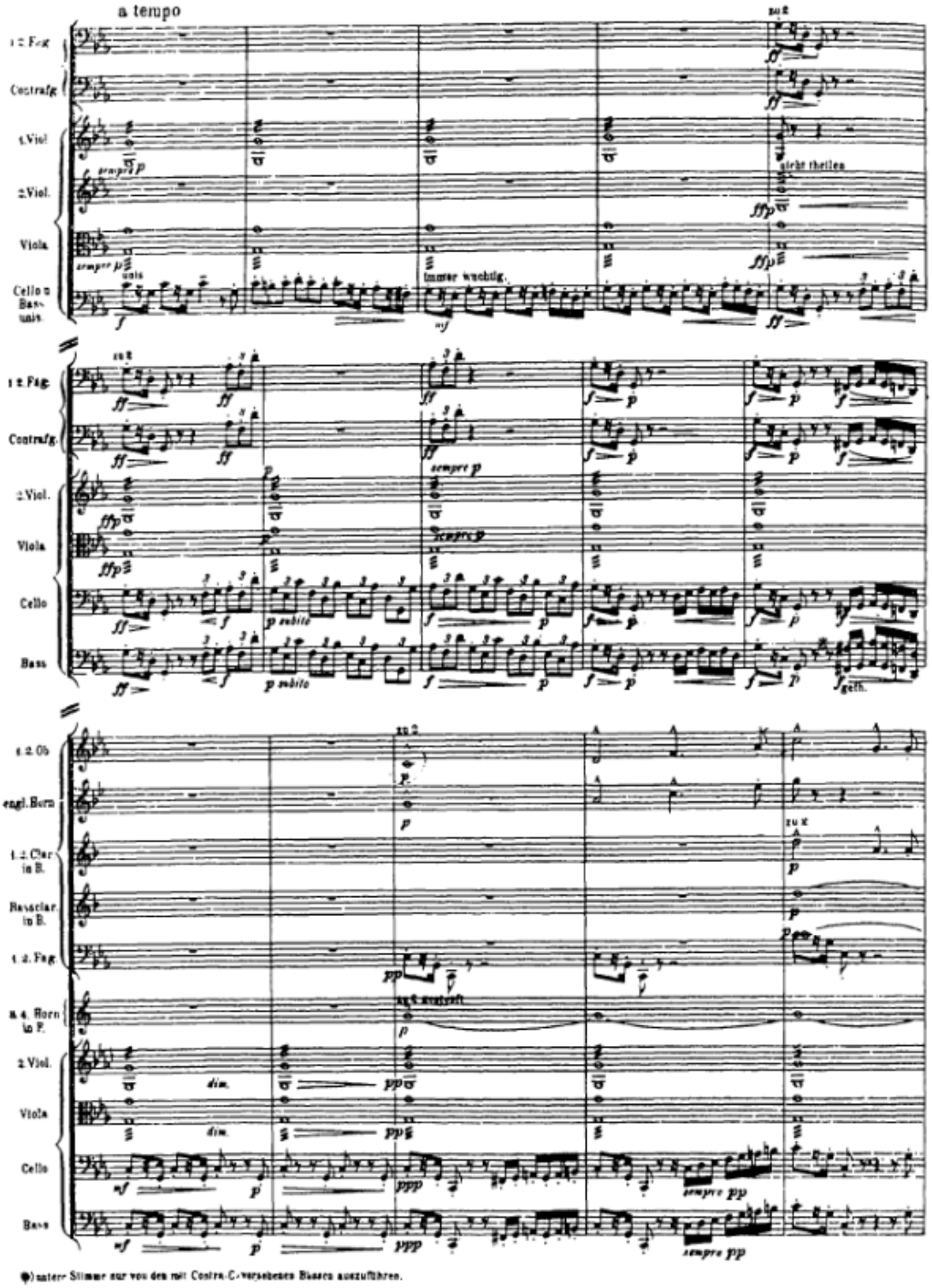


[26] In this performance of the work, https://www.youtube.com/watch?v=JeJhKWXkE7Q, the second movement begins in 12':20" and the intended passage is in the following time slot: 15':16" – 15':50"

Example B. Gustav Mahler, Symphony No. 2, "Resurrection", mm. 1-17



Example B. (Cont.)



*) unter Stimmen nur von den mit Contra-C versehenen Bässen auszuführen.

Bibliography

Balog, K. (2009). Phenomenal Concepts. In B. P. McLaughlin & A. Beckermann (Eds.), The Oxford Handbook of Philosophy of Mind (pp. 292–312). Oxford University Press.

Bonds, M. E. 2006. *Music as Thought: Listening to the Symphony in the Age of Beethoven*. Princeton University Press.

——. 2014. *Absolute Music: The History of an Idea*. Oxford University Press.

Budd, Malcolm. 2003. "Musical Movements and Aesthetic Metaphors." *British Journal of Aesthetics* 43 (3): 209–23.

Davies, Stephen. 1980. "The Expression of Emotion in Music." *Mind* 89 (353): 67–86.

——. 1994. *Musical Meaning and Expression*. Cornell University Press.

——. 2011a. "Artistic Expression and the Hard Case of Pure Music." In *Musical Understandings and Other Essays on the Philosophy of Music*, 7–20. Oxford University Press.

——. 2011b. "Music and Metaphor." In *Musical Understandings and Other Essays on the Philosophy of Music*, 21–33. Oxford University Press.

During, Jean. 1991. "Music, Mysticism, and Metaphysics." In *The Art of Persian Music*, by Jean During, Zia Mirabdolbaghi, and Dariush Safvat, 167–76. Washington, DC: Mage Publishers.

Goodman, N. (1976). *Languages of Art: An Approach to a Theory of Symbols* (2nd ed.). Hackett Publishing Company.

Goehr, Lydia. 1998. "Political Music and the Politics of Music." In *Musical Worlds: New Directions in the Philosophy of Music*, edited by Philip Alperson, 131–44. University Park, PA: The Pennsylvania State University Press.

Hanslick, E. (1986). *On the Musically Beautiful: A Contribution Towards the Revision of the Aesthetics of Music* (G. Payzant, Ed.). Hackett Publishing Company.

Higgins, Kathleen Marie. 1997. "Musical Idiosyncrasy and Perspectival Listening." In *Music & Meaning*, edited by Jenefer Robinson, 83–102. Ithaca, NY: Cornell University Press.

Jackendoff, Ray, and Fred Lerdahl. 2006. "The Capacity for Music: What Is It, and What's Special about It?" *Cognition* 100 (1): 33–72.

Kania, Andrew. 2015. "An Imaginative Theory of Musical Space and Movement." *British Journal of Aesthetics* 55 (2): 157–72.

Kivy, Peter. 1981. *The Corded Shell: Reflections on Musical Expression*. Princeton, N.J.: Princeton University Press.

——. 2002. *Introduction to A Philosophy of Music*. Oxford: Oxford University Press.

——. 2007. "Another Go at the Meaning of Music: Koopman, Davies, and the Meanings of 'Meaning.'" In *Music, Language, and Cognition*, 137–53. New York: Oxford University Press.

——. 2009. *Antithetical Arts: On the Ancient Quarrel Between Literature and Music*. Oxford: Clarendon Press.

Koopman, Constantijn, and Stephen Davies. 2001. “Musical Meaning in a Broader Perspective.” *The Journal of Aesthetics and Art Criticism* 59 (3): 261–73.

Leuenberger, Stephan. 2008. “Supervenience in Metaphysics.” *Philosophy Compass* 3 (4): 749–762.

Levinson, Jerrold. 2006. “Musical Expressiveness as Hearability-as-Expression.” In *Contemplating Art: Essays in Aesthetics*, 91–108. Oxford University Press.

Marissen, Michael. 2016. *Bach & God*. New York: Oxford University Press.

Matravers, Derek. 2003. “The Experience of Emotion in Music.” *The Journal of Aesthetics and Art Criticism* 61 (4): 353–63.

Morrison, Steven J., Steven M. Demorest, and Marcus T. Pearce. 2019. “Cultural Distance: A Computational Approach to Exploring Cultural Influences on Music Cognition.” In *The Oxford Handbook of Music and the Brain*, edited by Michael H. Thaut and Donald A. Hodges, 42–63. Oxford University Press.

Peacocke, Christopher. 2009. “The Perception of Music: Sources of Signifi Cance.” *British Journal of Aesthetics* 49 (3): 257–75.

Pederson, S. (2009). Defining the Term “Absolute Music” Historically. *Music & Letters*, *90*(2), 240–262.

Peltola, Henna-Riikka. 2017. “Sharing Experienced Sadness: Negotiating Meanings of Self-Defined Sad Music within a Group Interview Session.” *Psychology of Music* 45 (1): 82–98.

Puolakka, Kalle. 2017. “Musical Quotations and Shostakovich’s Secret: A Response to Kivy.” *British Journal of Aesthetics* 57 (1): 37–50.

Ridley, Aaron. 1995. “Musical Sympathies: The Experience of Expressive Music.” *The Journal of Aesthetics and Art Criticism* 53 (1): 49–57.

Searle, John R. 2004. *Mind: A Brief Introduction*. Oxford University Press.

Scruton, Roger. 1974. *Art and Imagination: A Study in the Philosophy of Mind*. Methuen & Co Ltd.

——. 2004. “Musical Movement: A Reply to Budd.” *British Journal of Aesthetics* 44 (2): 184–87

Stecker, Robert. 1999. “Davies on Musical Expression of Emotion.” *British Journal of Aesthetics* 39 (3): 273–81.

Taruskin, Richard. 2019. “How to Win a Stalin Prize: Shostakovich and His Piano Quintet in G-Minor (1940).” *International Review of the Aesthetics and Sociology of Music* 50 (1 & 2): 19–46.

Trivedi, Saam. 2011. “Music and Imagination.” In *The Routledge Companion to Philosophy and Music*, edited by Theodore Gracyk and Andrew Kania, 114–22. Routledge.

Walton, Kendall. 1999. “Projectivism, Empathy, and Musical Tension.” *Philosophical Topics* 26 (1 & 2): 407–40.

Zangwill, N. (2014). Re-Centring Musicology and Philosophy of Music. *Journal of Aesthetics and Phenomenology*, *2*, 231–240.

Zatorre, Robert J., and Andrea R. Halpern. 2005. “Mental Concerts: Musical Imagery and Auditory Cortex.” *Neuron* 47: 9–12.

Chapter 2: The Experience of Musical Expression

In this chapter I intend to examine those theories of musical expression that try to explain the phenomenon by focusing on the *experience* of expression rather than any particular *cause* of the experience. Though these theories are not totally indifferent to the causes of the experience and sometimes rely on a causal consideration tacitly or explicitly, what they *primarily* try to do is to explain what the experience itself is and not why one can have it. By referring to the causal-phenomenological distinction drawn in the former chapter, we can say these theories are concerned primarily with the phenomenological rather than the causal dimension of musical emotions.

I will consider three main ideas such theories typically appeal to:

a. Imaginative Idea (II): appreciating musical expression requires imagination.

b. Metaphorical Idea (MI): emotions are metaphorically attributed to music.

c. Dispositional Idea (DI): musical expression should be explained in terms of the dispositional capacity of expressive music.

A theory that appeals to II claims that (one form of) imagination is an essential component of the experience of musical emotions. The idea is that when a listener has this experience, or, as it is sometimes articulated, when a listener hears emotions in music, the listener needs to imagine something or imagine music in a particular way. MI usually leads to an anti-realist view concerning musical emotions. By appealing to it a theory (unlike the appearance theory) regards an emotional quality not as a property of a work but as what is metaphorically attributed to music as a result of the way in which some aesthetic properties of music are experienced. And according to DI expressive music causes or disposes a listener to *experience an emotion.* Yet, a theory that uses DI as its central idea does not necessarily consider "experiencing an emotion" to

be the same thing as "feeling an emotion"; so, such a theory can distance itself from the arousal theory whose main problem we saw in the previous chapter.

It is important to note that these ideas need not be appealed to by a theory in a mutually exclusive way. So, for example, a theory can be both imaginative and metaphorical, using II and MI together, or both dispositional and imaginative by appealing to both DI and II. In the theories that I will examine in this chapter at least one of these three ideas plays the central role. In what follows, I am going to explore each of these ideas, starting with II. Through examining the main theories in the literature that have used them in explaining musical expression, my aim is to investigate whether and how each of the three ideas can shed light on the phenomenological dimension of musical expression. In the next chapter, we will consider what our investigations in this chapter can add to our discussions of the appearance theory in Chapter 1. This will, I hope, get us to a better understanding of the relation between music and emotions, and musical reference more generally.

1. Musical Emotions and Imagination

One idea resulting from my critical evaluation of the appearance theory in Chapter 1 was that to recognize a similarity between (expressive) music and human behavior seemed to require a little more than mere perception: imagination might be needed to perceive musical movement as human behavior. But, we saw, this is not the claim of the appearance theory. The theory's main thesis is that musical expression is the result of the listener's *perceiving a similarity* between musical structures and human expressive behavior; and this is a perceptual and not an imaginative, or a perceptual-cum-imaginative recognition.

However, some hold that hearing musical emotions requires a listener to *imagine* music in a particular way. They construe, in other words, experiencing musical expression as imaginative perception. A main theory of this sort is the *persona theory.* My aim is to examine one of the main formulations of the theory by Jerrold Levinson.[27] This is how he articulates the central characteristics of his theory:

> [M]usic expresses an emotion only to the extent that we are disposed to hear it as the expression of an emotion, although in a non-standard manner, by a person or personlike entity. […] Since expressing requires an expresser, this means that in so hearing the music the listener is in effect committed to hearing an agent in the music—what we can call the music's persona—or to at least imagining such an agent in a backgrounded manner. But this agent or persona, it must be stressed, is almost entirely indefinite, a sort of minimal person, characterized only by the emotion we hear it to be expressing and the musical gesture through which it does so. (93)

Remember the distinction between expressiveness and expression we talked about in the previous chapter. Expression is the behavior of a psychological being caused by her feeling. Levinson does not, obviously, think music is a psychological being capable of expressing any feeling; but he believes a listener hears expressive music *as if* it is expression—that is, as if a "personlike" agent is expressing an emotion (through musical sounds). So, Levinson regards musical expressiveness as an *as-if-expression*. He thinks that expressive music *disposes* a listener to hear it as an expression. It is worth noticing that the dispositional articulation of Levinson's theory is not related to the Dispositional Idea (DI). Levinson does not think music disposes a listener to feel anything. His idea is that expressive music disposes a listener to *imagine* that a musical persona is expressing an emotion.

One characteristic of Levinson's theory that is important for my purposes in this chapter is that Levinson's primary focus is not on *why* expressive music disposes a listener to hear it as an

[27] Karl and Robinson (2015) have used this notion of "musical persona" to account for how music can present psychological dramas in which musical personae can express different emotions.

expression (that is, to imagine that a persona is expressing an emotion). The cause of musical expression is not his primary concern: his aim is to explain what the listener's experience is like when she hears a musical emotion or an emotion *in* music. His theory then provides an explanation for the experience of musical expression which is not reliant on a listener's awareness of any cause of the experience.

To clarify the point, let us compare the persona theory with the appearance theory. According to the appearance theory, as we saw, the main cause of musical expression is the resemblance musical structures can have to the appearance of human expressive behavior, *and* the experience of the listener is the *perception* of this resemblance. But according to the persona theory, experiencing musical expression is not perceiving any particular property of music or awareness of any cause of musical expression; it is imagining a musical persona expressing an emotion. In other words, based on the appearance theory the content of the experience of musical expression is "perceiving a resemblance [between music and human behavior]", and a part of this content—the intended resemblance—is the main cause of musical expression. However, based on the persona theory, the content of musical expression is "imagining a persona expressing some emotion [through musical sounds]," where no part of this content is any musical property, or any conventional, cultural, or other kind of the cause of musical expression. That is to say, according to this theory no constituent of the phenomenology of musical expression is necessarily knowledge of or awareness of any cause of musical expression.

A concern about the persona theory has been expressed by Paul Boghossian (2007, p. 126): "Levinson's listener is like the experienced observer, moving swiftly from his hearing to the attribution of expression. But he cannot be said to *know*, purely on the basis of hearing, that a particular passage has a certain expressive property" (emphasis added). Boghossian thinks it is a

problem that there is no justificatory consideration in Levinson's persona theory: it is not clear, according to the theory, how a listener should justify herself in hearing a passage as expressive based only on what she hears in the music. We will later examine whether this can be a problem with the theory. But here, I want to argue that this can be a merit to the theory.

The positive aspect is that Levinson's theory aims to offer a unitary characterization of the experience of musical expression over and above the experience's different causes. The theory, in other words, tries to give an account of the experience in question independently of different probable causes it might have. Whether the cause of the experience is merely conventional/cultural, or specific perceptual properties of music, or is a combination of the other two, the theory provides a unitary characterization: that experiencing a cheerful passage, for example, is imagining a musical persona expressing cheerfulness. I think this is a plus for Levinson's theory because it seems that a listener can hear or experience a passage as, say, sad and be unaware of what causes that experience. This is what I, in the previous chapter, referred to as a distinction between the causal and the phenomenological dimensions of the experience of musical expression. And now, I argue that Levinson tries to capture this distinction in his theory. To clarify this point, it is worth making a comparison between this theory and the appearance theory.

According to the appearance theory, you will remember, a main cause of musical expression is the resemblances a listener perceives between perceptual properties of music and the appearance of expressive behavior. That theory is not, however, dismissive of the conventional or cultural factors that in many cases play the key role in musical expression. I have not examined the conventional component of the appearance theory so far; but now it is time to do so, particularly because it will enable us to see the positive side of Levinson's theory

mentioned above and why this theory, in comparison with the appearance theory, is more convincing when it comes to the conventional dimension of musical expression. Hence our digression from the persona theory for the moment.

1.1 The Appearance Theory: Conventional Component

We can disentangle two ways in which Davies's and Kivy's appearance theory incorporates the conventional cause(s) of musical expression: (1) that convention can play a causal role in musical expression *independently* of the resemblance perception; and (2) that convention shows a listener *how to perceive* a resemblance between music and human expressive behavior. I will refer to (1) as *conventional association* and to (2) as *conventional framing*.

Conventional association is the idea that sometimes attributing an emotional quality to music is merely the result of how in a musical convention some ideas have been associated with some musical qualities. For example, in Western musical convention the timbre of trumpet has militaristic and sometime heroic associations. Or, in Persian musical convention the timbre of *Ney*, a woodwind instrument, has some associations with mystic melancholy rooted in Persian classical literature. Such associations are mostly, if not totally, the result of how these instruments have been *used* in their musical culture and not the result merely of the very sound quality they produce. It would, for instance, be very unlikely that a person unfamiliar with Western musical culture could experience the sound of a trumpet similarly to a Westerner who has all the conventional associations attached to the instrument at the back of her mind. What is important to note about conventional association is that it does not function based on a resemblance perception proposed in the appearance theory. That is, conventional associations are not necessarily based on perceiving a similarity between a musical quality (like a trumpet's

sound quality) to human expressive behavior (a person's expressive voice or gesture, for example). Appreciating expressive qualities of music coming from such conventional associations is grounded on a listener's living in and being exposed to a musical culture and musically internalizing such associations, consciously or unconsciously. This is one way the conventional/cultural aspect of musical expressiveness has been construed in the appearance theory.[28] Related to the core proposal of the theory—the resemblance claim—Kivy and Davies have, however, considered another way musical expression can be cultural as well, what I have called conventional framing.

The idea of conventional framing is as follows. Most of the time, the recognition and perception of how music can be similar to human expressive behavior requires a listener to be acquainted with (or even musically trained in) the related musical convention. In Davies's (1994, 246) terms: "Music's expressiveness might be rooted in resemblances between the ways we experience musical movement and human behaviors presenting emotion characteristics in appearances, but one can be sure about the expressive character of any musical work only if one is familiar with the musical conventions that structure the musical practice or style in question." And concerning his resemblance claim—that is, the claim that music can present perceptual properties bearing resemblances to human expressive contour or behavior—Kivy (1980, pp. 85-86) says:

> Let us suppose a non-Western listener, completely uninitiated into Western musical culture, "hears" a performance of *X*, some musical composition. He does not hear *X*, in a certain obvious sense, but he hears a collection of sounds that has quite a different significance for him from the significance it has for the Western ear. These sounds may, indeed, be heard by him as expressive behavior, or

[28] Kivy (1983, 76) in fact uses the idea of conventional associations to base another theory of musical expressiveness, what he calls the "'convention' theory," which "explains the expressiveness of music as a function, simply, of the customary association of certain musical features with certain emotive ones, quite apart from any structural analogy between them." Davies considers such associations as non-natural meanings of music; see Davies (1994, p. 42).

> happen, by chance, to conform to some of the expressive conventions of his musical culture. But he has not heard *X*; so he has not heard the expressive properties of *X*. And we are interested here only in the expressive properties that the musical composition has, not the ones that the collection of sounds involved in its performance might have, if heard by someone completely untrained in Western musical listening.

As we can see, for both Davies and Kivy familiarity with a musical convention is required for a listener to appreciate the expressive properties of musical objects of that convention. We can roughly say that the intended familiarity is a familiarity with the "language" of music in a musical convention. They believe that without such a familiarity a listener cannot experience/engage in appropriate aesthetic appreciation of music. In fact, such a familiarity *musically frames* the ways in which a listener experiences the music of a culture. As Kivy says, without it musical experience becomes merely listening to a "collection of sounds." (And I believe, or at least assume, that both agree that this kind of familiarity basically and naturally is gained and internalized when one is exposed to the music of her own culture.)

Kivy also talks about another way convention is related to musical expressiveness: in perceiving expressive music as resembling expressive human behavior, the listener also needs to be acquainted with the related expressive conventions. For example, a non-Westerner from a culture in which sadness is outwardly expressed quite differently from the emotion's expression in Western culture is unable to perceive sadness in a work of Western music—at least the musically expressed sadness that is aesthetically relevant in Western musical culture. (As is clear in the above quoted passage, Kivy does not rule out that such a listener can perceive some resemblances between that music and human expressive behavior based on the expressive conventions of her own culture, but he seems to think that such resemblances would not be those that were aesthetically relevant according to Western musical conventions.) So, conventional framing is based on two kinds of familiarity a listener needs in order to perceive the

expressiveness of music from a particular culture: familiarity with the musical conventions of that musical culture, and familiarity with the conventions related to expressive behavior in that culture. Either way, conventional knowledge (acquired whether consciously or just by living in a culture) *frames* how a listener perceives resemblances between expressive music and human expressive behavior. It is in this way, at least to my reading, that the resemblance perception intended in the appearance theory is related to convention or has a cultural dimension.

Now, after understanding the conventional component of the appearance theory, we can see more comprehensively how the *experience* of musical expressiveness is characterized by that theory. Experiencing musical expressiveness is considered either as recognizing how music can sometimes make conventional associations or perceiving how music can sometimes resemble expressive human behavior.[29] That is, according to the appearance theory we either *are reminded of* an emotional association by a musical quality (like the timbre of an instrument) or *perceive* a resemblance between musical structures and the appearance of human expressive behavior. In fact, the theory gives us a divided description of how we experience musical expression.

It is divided because the two experiences mentioned are phenomenologically different. An example can be clarifying. Compare the following two cases: seeing a symbol which is conventionally associated with some religious or mythical concept in a specific culture, call it C, and seeing a picture of a sad face.[30] In the first case, by seeing the symbol one entertains C. But in the second case, one perceives the visual characteristics of the picture as being similar to a sad face. In the first case, entertaining C is not the result of how one perceives the structural parts of

[29] It should be noted that the recognition of cultural association is not necessarily a conscious process—or at least the theory does not claim such a thing. Such associations can be made unconsciously in the listener's mind. However, to claim that perceiving a similarity between music and human behavior can be conscious or unconscious is, I think, a more complicated matter as discussed later in this chapter.

[30] Suppose further that the symbol does not have any similarity to anything related to C; the two are merely conventionally associated to each other.

the icon. But in the second case, one's seeing sadness in the picture is the result of *how she perceives* the picture's visual characteristics. These two experiences are then different phenomenologically.

Similarly, experiencing musical expressiveness through conventional associations is different phenomenologically from experiencing it through resemblance perception. The former is like seeing a symbol and the latter seeing, say, the picture of a sad face. The sound of, say, trumpet is not experientially similar to being triumphant but conventionally associated with it—in the same fashion, a visual symbol may not be perceptually similar to anything related to, say, a religious concept but conventionally associated with it. However, according to the appearance theory when music is expressive based on its perceptual properties and not based on conventional associations, its *experiential features* make it expressive and not its conventional association with emotions. This is why I believe the appearance theory gives us a divided account of the *experience* of musical expressiveness. And I want to argue why such an approach is problematic in explaining that experience.

I do not deny that musical expression may have two or more *causes*. As the appearance theory claims, one cause of musical expression might be conventional associations, and another might be the resemblances musical structures can make with emotional behaviors. What does not seem to be the case, however, is that we have two different kinds of experience of musical expression while listening to music—one whose phenomenology is structural perception and the other whose phenomenology is making associations between an experience and some emotions (or emotional concepts). Hearing expressive music by conventionally associating it with emotions is similar to seeing a cultural icon (mythical or religious) and conventionally associating it with some concepts in the sense that both involve *conventional associations*. On

the other hand, hearing expressive music by perceptually recognizing a resemblance between the music and some human expressive behavior is like seeing a picture of an expressive (say, a sad) face in the sense that both experiences involve *perceptual engagement*. And these are, as argued, experiences that are phenomenologically different. The experience of musical expressiveness does not seem to be divided into these two sorts of experiences when one hears emotions in a piece. As argued in the previous chapter, one reason is that hearing musical expression is not necessarily dependent on the listener's *awareness* of why music is expressive.

An example can make the point clear. The beginning moments of the Allegro of Johannes Brahms's Clarinet Trio in A minor is melancholic. A part of the reason can be found in conventional associations the music makes to minor mode or to the ways in which the cello has been used to bring about such qualities in the Western musical convention. However, the reason also can be related to similarities the music can have to sad and melancholic behavior: for one, we have the sonority of the cello that could, aside from its conventional associations, remind us of a melancholy, low voice. Also, the main theme played by the cello first soars to D and then sinks to the tonic an octave below. The clarinet continues the theme with a similar pattern: soaring and then sinking, this time with a diminuendo, to the tonic (again more than an octave below where it had picked up the theme). This thematic pattern can be reminiscent of a person who prefers to sink into her inner feelings and memories.

Let us assume that in this example the only main two causes of the musical melancholy are conventional associations and resemblance perception. It seems that our experiencing that melancholy is neither merely an experience of the first nor merely of the second cause. In other words, our hearing the melancholic quality of the passage is not an experience fractured by two sorts of other experiences: making conventional associations and perceiving similarities. Our

hearing music in that way is an experience over and above any awareness of such causal factors: it seems to be a *unitary, holistic* experience. Now, the question is what that unitary, holistic experience is like.[31] The appearance theory does not enable us to approach this question by its appealing to similarity perception and conventional association making. Focusing on and dealing with the nature of this unitary experience of musical expression appears to be what Levinson is after through his persona theory.

1.2 The Persona Theory and Musical Expression as a Unitary Experience

Levinson says:

> [M]usic's *being* expressive of α just means, once again, that it is readily and aptly heard as sui generis expression of α by the class of appropriate listeners. Whether it is direct resemblance to human expressive behavior, or conventional associations attaching to certain musical features, or resemblance to expressive natural phenomena or even partial isomorphism between the flow of the music and the inner flow of feelings, which we identify in a particular case as undergirding the expressiveness we take to be present, such factors do not, I maintain, figure in the analytic explication of the concept of musical expressiveness. (Levinson 2011, 339n7)

Levinson, as we can see, seeks an account of musical expressiveness that can be characterized independently of any specific cause of the phenomenon—whether the cause is resemblance to human behavior, conventional association, or anything else. His theory is an attempt to give, in his words, an "analytic explication of the concept of musical expressiveness." Whatever his conceptual analysis of expression amounts to, I think we can consider the result of Levinson's theory as an account of the nature of the experience of musical expression. As we saw earlier, the thrust of the theory is that the experience of a listener while listening to an

[31] By "unitary" I do not mean that all listeners should necessarily have the same experience of musical expression. What I mean is that the experience of *one* listener seems to be a unitary experience and not divided into different causes of musical expression.

expressive passage is an imaginative experience. The listener imagines that an indefinite musical persona is expressing an emotion. Even if the theory should be taken as an analysis of what we *mean* when we say a listener hears a passage as expressive of an emotion, it is obvious that the theory (also) provides a description of the *experience* of a listener when she hears musical expression. This experiential or phenomenological claim of the theory is what I am concerned with.

In my understanding, the theory construes the experience in question as a unitary experience whose phenomenology can be characterized independently of any particular cause. In hearing an emotion in an expressive passage a listener, it is claimed, does not perceive a resemblance (between music and human behavior), nor does she recognize a conventional association, nor is she aware of any other cause behind the passage's expressiveness. Whatever the cause(s) might be, the listener undergoes the imaginative experience described earlier. This is why I believe that the persona theory gives us a phenomenological account of musical expressiveness that does not hinge upon any cause of the related experience.

This approach, at least in my reading of it, has an important advantage over theories like the appearance theory. As I argued in the first chapter, there is some level of independence between what the experience of musical expression is like—or what the experience *is*—and what causes it. Our experience of musical expression is not awareness of any particular causal element of the expression (like conventional associations) but an experience *caused by* such elements. It may be the case that conventional associations and resemblances that music can have to human expressive behavior are the two main (or even the only—this would not affect the current argument) causal elements behind musical expression. This, however, does not entail that in experiencing an expressive passage we think of a conventional association or notice a

resemblance. The plausible view appears to be that our experience is somehow *brought about by* such causal elements functioning together in a musical piece or passage. What we need phenomenologically is to explore and explain what this experience is—an experience that is, metaphorically speaking, "over and above" what causes it. Levinson's approach is directed towards such an explanation, which itself is a merit of the theory methodologically. Nonetheless, it faces at least two serious problems to which I want to turn now.

One problem is the vagueness of the object of imagination in the theory's imaginative proposal. The problem is related to the notion of musical persona. As Stephan Davies (2011) has pointed out, it is not obvious whether in listening to expressive music we really imagine such an agent indefinitely musically presented to the listener. The issue is that however indefinite or minimal this imaginative agent is taken[32] to be it is *something* that Levinson thinks we imagine while listening to expressive music. Now, it is phenomenologically dubious that *something* is imagined in hearing expression.

A form of response to this challenge can, I reckon, be found in Levinson's theory. Levinson does not say one *should* imagine a persona in the case of expressive music. He thinks a musically competent, culturally backgrounded listener naturally or automatically imagines such a persona. Expressive music, in his terms, is "*mostly spontaneously and readily* imagined" to be the expression of an emotion. This can entail the following. It might be that in many cases a competent listener *unconsciously* imagines a persona. Or such a listener need not be aware of *what* she imagines or that she imagines at all; her competence enables her to attend to expressively relevant features of music in a culturally suitable way and, as a result, she hears an

[32] As Levinson says: "[persona] is almost entirely indefinite, a sort of minimal person, characterized only by the emotion we hear it to be expressing and the musical gesture through which it does so." (2006, 93)

as-if-expression through music. I think this can remedy the above issue to some extent: Levinson does not claim that imagining a persona in music is a necessarily conscious process.

This might be the case. Or it might solve the issue to some extent, but not completely. In my opinion, the idea that a competent listener performs Levinson's intended imagination spontaneously and even unconsciously does not remedy the issue because the kind of imagination Levinson is appealing to needs an object that is definite to some extent. I might exercise my imagination in encountering artworks spontaneously and unconsciously. But if I was reminded of the fact that in, say, reading a novel I do a lot of imagination or imagine different things mostly unconsciously, I would be able to understand the suggestion, to understand both *what* I imagine on such an occasion and why the work causes me to imagine them (different descriptions in the novel of scenes, characters, events, etc.). But the object of Levinson's proposed imagination does not have this characteristic. Musical persona is too sui generis to be informative. Even the minimal definition given in the theory is not very helpful and is to some extent question-begging.

The intentional object of Levinson's intended imagination involved in hearing emotions in music is described as follows: a persona about which/whom the only thing we know or need to know is that she/he/it can express an emotion at a particular musical moment *through music*. (A sad passage is, for example, the expression of sadness by such a minimally defined persona.) But when we say "through music" our description becomes uninformative and question-begging in a sense. The theory tells us that experiencing musical expressiveness is imagining a persona that can *musically* express. Now, it can be asked what experiencing musical expressiveness of an imaginative agent/a persona is. The theory does not provide a phenomenologically informative

exploration of (the experience of) expressiveness; it only pushes what is phenomenologically in question from music to an imaginative agent.

A second problem with the account is related to the notion of musical competence. As claimed, the musical persona is imagined by a competent or qualified listener. But who is a competent listener? Levinson (2006, p. 102) explains:

> […] a qualified listener who hears the expression of some emotion in music normally acquires, without reflection, the conviction that the music is expressive of that emotion, that is, would be readily hearable as such by other qualified listeners. […] Qualified listeners arguably tacitly assume, while listening to music, that they are qualified listeners, and are listening appropriately.

What Levinson says about musical competence seems to be applicable to language competence as well. Many people "tacitly assume" themselves to be qualified users of their mother tongue and normally this does not rely, say, on any reflection on their knowledge about that language. Many of us are not aware of and do not know all the grammatical subtleties behind the ways we naturally use our mother tongue; but this does not prevent us from regarding ourselves as qualified and competent users of that language. I think Levinson is right that something like this happens in our musical experience, specifically in our experience of a musical idiom we have enough familiarity with. That is, I agree with Levinson that musical competence connected with expression can be unreflective and a natural result of someone's sufficient familiarity with a musical idiom. But I believe something important is missing in Levinson's description: a connection between a listener's experience of musical expression and how she attends to music.

Remember the main lesson we took from the appearance theory in the previous chapter. It was that musical expressiveness is a holistic quality somehow based on (supervenient on or emerging from) musical elements, and the recognition of this quality requires structural attention.

Although I tried to show that this recognition is not resemblance perception, as claimed in the appearance theory, I argued that it does require *attending to* musical features involved in an expressive passage. How exactly this attention is to be fleshed out is one of the main questions we will get to in the next chapter. But it suffices for the current argument to say that in accounting for how musical expressiveness is experienced by a (competent) listener a theory needs to specify structural attention in one way or another. This is altogether missing in what Levinson says about the imaginative way in which a competent listener experiences musical expression.

That is, Levinson does not provide how exactly emotions are heard (imaginatively) in music *as a result of* attending to the relevant musical features.[33] Notice that this is not to say that a qualified listener should be aware of any particular cause of expressiveness or know why the music she is listening to is expressive. Structural attention is not identical with such an awareness or knowledge. Yet, I doubt that a competent listener is *totally* unaware of her way of listening to music.[34] My worry is that Levinson's expressive theory makes experiencing musical expression too unreflective. As regards musical persona he seems to claim that imagining it is mostly unconscious, and regarding musical competence he says this competence is mostly unreflective.

---

[33] In his *Music in the Moment*, Levinson (2007) gives a theory (*concatenationism*) of what kind of attention is needed in order for a listener to have basic musical understanding. There, Levinson also claims that appreciating musical expression is dependent on that kind of attention. But in my understanding, neither in that book nor in his specific work on musical expression does Levinson explain how exactly a concatenationist musical attention leads a listener to experience musical expression in an imaginative way, that is, through imagining a musical persona. This is I think a gap between Levinson's theory of musical expression we are examining here and his theory of musical understanding or attention. From a different perspective, I will discuss Levinson's concatenationist theory in the next chapter. But what is important for my discussion in this chapter is that even considering Levinson's theory of musical understanding does not remedy the problem I am here considering related to the notion of competent listener.

[34] In the psychological study conducted by Peltola (2017) people in a study group try to describe to, and also persuade, each other—not necessarily in technical terms—how they hear sadness in their chosen pieces. This can provide psychological evidence for the claim that people have some awareness of how they attend to musical features in a piece as a result of which they might attribute emotions to it.

Let me recap our evaluation of the persona theory. I argued that the strength of the theory is that it gives us a unitary account of the experience of musical expression—that is, the account is not divided into conventional and musical causes of expression. This has been achieved through the imaginative approach of the theory. Thus, appealing to imagination to explain how musical expressiveness is experienced seems to be a plausible approach. However, I argued that there are two downsides to the theory. The first one is related to musical persona. As the intentional object of the sort of imagination suggested in the theory, it is vague and described uninformatively. Second, the persona theory cannot tell us how a competent listener's imaginative hearing of emotions in music is the result of her attending to the relevant musical features. Now, let us consider the theories that deal with musical expression basically through *metaphor*.

2. Musical Emotions as Metaphor

As mentioned at the beginning of this chapter, another main idea some theories appeal to in explaining musical expression is the Metaphorical Idea (MI) according to which emotions are metaphorically attributed to music. I will examine two theories of this kind in what follows. But before getting to the claims of these theories in detail, I want to say something about the initial appeal of the idea itself.

Remember that we formulated the core philosophical problem of musical expression in this way. Musical works (or performances) are not the sort of things psychologically able to express any emotions; so how do we attribute emotions to them when we describe a passage as, say, sad or cheerful? One response here can be that in such descriptions we do not *actually* think that musical objects are psychological beings. What happens is that we borrow an emotional-

psychological term from the domain in which it is primarily used, that is, describing psychological beings, and we use that term in a secondary domain, that is, describing inanimate, musical objects. Now, the notion of metaphor seems to be immediately applicable here since such uses of terms typically happen in figurative speech—for example, when one says: "His heart is an ocean." This can show why MI has an initial appeal with regard to musical expression: it seems to open a promising route to explain why an emotional-psychological term is used to describe a musical work which cannot actually be in any psychological state.

Related to this appeal, the metaphorical idea appears to have another interesting aspect: it seems to be able to show that musical emotions have a relation to emotions themselves as mental states and not merely to the appearance of emotional behavior. What a metaphorical approach tries to show is that the emotions (sometimes) attributed to music are metaphorically related to those mental states that are expressed by people. This is not what, for example, the appearance theory claims. According to that theory, as we saw, musical emotions are primarily related to the *appearance* of emotional behavior or posture (like the perceptual configuration of a Saint Bernard's face) and not any mental state. And I think a metaphorical approach has at least a prima facie advantage over the appearance theory, because it is aimed at explaining emotional attributions to music through the primary and not the secondary usage of emotional terms. The primary usage is the usage of emotional terms for referring to emotions as mental states and the secondary usage the one for referring to emotional appearances. The latter is secondary because it is always parasitic on the former. (For example, we say a Saint Bernard's face is sad in its appearance because it *looks like* the face of a person when she feels sad and expresses the emotion through her face.) By saving the relation between musical emotions and the primary usage of emotional terms (or emotions as mental states), a theory might better get to the

psychological depth of our musical experience, to the fact that there are occasions in which musical emotions can move us, change our moods or how we feel. This prima facie advantage notwithstanding, there is a real challenge for any theory of musical expression that appeals to MI: how should the metaphorical relation between music and emotions be pinned down?

The challenge is that our attribution of an emotional quality to a musical passage is not arbitrary; there must be a musical basis for such an attribution. Suppose we attribute, say, agitation to a passage and this is the quality the attribution of which to that music most competent listeners consider appropriate. Now, it would not be enough for a theory to tell us that this attribution is metaphorical because the music itself is unable to feel agitation or any emotion. We also need to know why that particular emotional quality, rather than another such as happiness, is attributed to that music. In other words, we need to know why using that specific metaphor, instead of some other one, is musically and aesthetically appropriate.

In what follows, I do not want to go into or examine the philosophical problems around the notion of metaphor and its application in artistic expression. My focus will be on the musical side of the problem by examining a theory of musical expression which is rooted in metaphorical considerations of musical experience. I do not mean that the examination of this theory cannot be affected by those broader considerations related to aesthetic and artistic applications of metaphor. But what I am first concerned with is whether and how a metaphorical view of musical expression can remedy the challenge I mentioned above, that is, whether and how it can tell us what basis can make the (metaphorical) attribution of emotions to music *musically* justifiable.

The theory I am going to examine is Roger Scruton's which is developed in his *Aesthetics of Music*.[35] The work provides an elaborate and insightful investigation into different aspects of musical understanding. Examining that investigation fully is beyond the scope of my purposes here. My concentration will be on Scruton's proposal for musical expression and particularly how that proposal is based on his metaphorical explanation of musical elements. The following is Scruton's specific view on musical expression:

> [W]e could understand expression only if we could understand our response to it. The response to expression is a sympathetic response […]. (1999, 354)
> Our response to music is a sympathetic response: a response to human life, imagined in the sounds we hear. However, in the absence of representation there is no precise object of sympathy—neither an imaginary human subject, nor a situation perceived through his eyes. The life in music belongs in the musical process, abstract, indeterminate, unowned except through the act whereby we listeners possess it. (ibid., 355)

Unpacking this passage will get us to the metaphorical roots of Scruton's expression theory. Scruton thinks that to understand musical expression we primarily need to understand how we listeners *respond to* music instead of what is *in* music. His main reason is that music cannot literally possess any quality that can make it resemble emotions. To corroborate this, Scruton undermines the idea that emotions can be *in* music, as Kivy's and Davies's appearance theory claims. According to that theory, as we saw, what enables music to resemble the appearance of emotional behavior are its more basic elements and the perceptual qualities they can present. For example, we saw in the previous chapter that the shakiness of the tremolos played by the strings is one musical feature that imbues the passage in Mahler's Second an anxious quality. So, one can argue, following the appearance theory, that the passage's anxiety is grounded on the shaky movement the music presents; and if that shakiness is *in* the music, so is

[35] Some of the arguments of this work, particularly the ones on musical expression, are based on Scruton's aesthetic views defended in *Art and Imagination*. When I use his ideas specifically in the latter, I will make it explicit.

the related emotion the passage is expressive of. However, Scruton denies that even such more basic musical qualities, like shakiness, can be *in* music. He thinks they are the result of how one *hears* musical sounds; we hear the tremolos played by the strings *as* a shaky movement in Mahler's passage. As he says:

> Music does not languish in the way that people languish; it does not droop, revive, palpitate, and strain in the fashion of a human gesture, but in its own fashion. It is only through the musical experience that these features can be grasped. You have to hear the emotion *in* the music, and that means attending to the music, in an act of aesthetic interest.

And in criticizing Kivy's view, he says:

> We call music sad, [Kivy] suggests, because it moves slowly, stumbles, droops, and so on—just as sad people do. But these are not real analogies, since music does not literally droop or stumble, any more than it is literally sad. The 'analogy' is simply the substitution of one metaphor for another.

I want to disentangle two of Scruton's ideas here: (a) that musical expression (or experiencing/appreciating musical emotions) is response-dependent; and (b) that musical emotions are *metaphorically in* music. And it seems that Scruton makes the following argument from (a) to (b): because appreciating emotions in music depends on how one experiences and attends to music or musical sounds (we appreciate them "through the musical experience," (a), then such emotions are metaphorically and not literally in music, (b). In other words, in my understanding of him, Scruton thinks musical emotions are essentially related to *how we experience music* and not *what is in music*. But a problem in this argument is that accepting (a) does not necessarily lead us to accepting (b).[36] As discussed in the previous chapter, even according to the appearance theory—which claims musical emotions are possessed by music as properties of it—it is not

[36] A similar criticism of Scruton's metaphorical view on musical experience has been raised by Budd (in Tanner and Budd (1985)).

denied that musical expression is response-dependent, that is, to hear, say, sadness in music, one has to listen to music in a particular way. This is denied neither by Kivy nor by Davies. For example, as we discussed earlier in this chapter, for both Kivy and Davies culture or musical conventions play a defining role in how one appreciates musical expressiveness.

There are then two issues with what Scruton says: (1) the fact that the expressiveness of music depends on how music is listened to does not entail that emotions are metaphorically attributed to expressive music; and (2) in criticizing (Kivy's) appearance theory, Scruton is mistaken to think that that theory does not acknowledge the response-dependence of musical expressiveness. We will, in a minute, discuss Scruton's more developed and stronger argument for the essential role of metaphor in experiencing musical expression. But before that, I think it is worth hammering out Scruton's criticism of the appearance theory by developing it a little.

To me, there is an important aspect in Scruton's criticism of the appearance theory especially if we consider his own suggestion for musical expression. His suggestion, as reflected in the first passage quoted from him, is that appreciating musical expression requires *sympathetic* response. We will see how this response should be understood based on Scruton's theory of musical understanding. But what I want to first highlight is that Scruton defends a *specific kind* of response to musical expressiveness because he seems to think the kind of response the appearance theory suggests is faulty. Considering the nature of his theory, I think Scruton's main problem with theories like the appearance theory is that they claim our response to musical emotions is recognitional—call it *recognitional response.*

The problem towards which Scruton's criticism of the appearance theory is directed is the following—to my interpretation of his view at least. Scruton believes that musical emotion *forms*, or comes to existence, in the way we listen, or respond dynamically, to music and not that it gets merely recognized through a response to music. An example can be clarifying here. Consider reading an X-ray image in order, say, to find a medical problem. This requires a specific kind of response by the viewer to the image, call it an *expert response*. Only a person with enough medical knowledge, experience, and expertise can identify the related problem by viewing the image. That is, only an expert can respond visually appropriately to the image in order to recognize those (visual) properties that show a medical problem. However, although perceiving those properties of the image requires *a specific response* to the image, that response is nothing but recognizing what is *in* the image. The related visual properties do not form or get created as a result of the expert response; the expert response merely *recognizes* what is already in the image. Now, Scruton thinks that our response to musical expression is not by nature such a recognitional response (however much that recognition is considered to require knowledge or expertise).

But the kind of response to music that the appearance theory suggests as regards musical expression is a recognitional response like the expert response to an X-ray image. It is claimed that a listener must be acquainted with the relevant musical conventions and, probably in some cases, have enough musical expertise to recognize the emotional qualities of music. But in the end, all such requirements for appropriately responding to music in order to experience/perceive its expressiveness facilitate our recognition of what is *in* music. And this is what Scruton criticizes; to Scruton, according to my reading of

him, musical emotions form in the (musical) *experience* and are not *in music*. So, even if we consider that the appearance theory acknowledges that musical expression is response-dependent, it does not completely shield the theory from Scruton's criticism, or the Scrutonian criticism I have formulated. The problem is that considering our response to musical emotions as a recognitional response cannot fully capture that musical expressiveness is not a matter of how more basic musical properties *are* but how they *are heard*. And the proposal of the appearance theory is slim in this regard.

Scruton's criticism of the appearance theory seems plausible then based on the reading of it given above. We need though to see how the problem he identifies in that theory can be dealt with in his own proposal. In fact, we need to examine how his theory can capture the phenomenology of the way in which we respond to/experience musical emotions. As quoted earlier, his proposal is that: "Our response to music is a sympathetic response: a response to human life, imagined in the sounds we hear." To understand what Scruton means by a sympathetic response to music we need to refer to his metaphorical view on how we experience musical elements.

To Scruton, metaphor plays an essential role in our experience of the building block of musical structures: tone. He thinks musical tones are not heard in isolation. A pitch heard as a sound with a specific frequency is not a musical tone; it is merely a sound (what he calls a secondary quality). To hear that sound/pitch as a tone, like D, is to hear it *in relation to* other pitches in a musical scale (or the other tones heard in a musical work). In other words, hearing the frequency basing, say, the tone D is not enough to hear it as a musical tone; it becomes a musical tone (a D) if it gets heard in relation to the other tones in the context of a musical scale, say in relation to C or G in the C major scale.

Scruton believes that this relational aspect of tone-hearing is connected with *musical space*. The tones in a musical scale are heard as *higher* or *lower* than each other and as having tendencies to *move* towards each other. For example, in listening to the nursery rhyme "Row, Row, Row Your Boat" (Fig. 1), the tones in the ending part ("life is but a dream") seem to be heard as going from G (dominant) to C (tonic) and reaching a resting point. It seems that we hear a moving tendency in that part, as if the melody moves in that way to naturally end or reach completion. Or we seem to hear tones going upwards (or downwards) in a melody or musical scale. All such characteristics—being higher or lower than, moving towards, ups and downs in a melody, and so on—are spatial descriptions of musical elements, descriptions that seem to be indispensable in how we experience music. And Scruton believes all such descriptions related to musical space are metaphorical because it is not the case that tones in music as discrete sonic events actually move towards or are higher or lower than each other; it is just how we hear them as moving or being in a "space." To Scruton, although the temporal dimension of music is real—for example, the duration of each tone—the spatial dimension is metaphorical. What in fact happens in our musical experience is "spatializing of the temporal order" we hear in music (1999, 75).

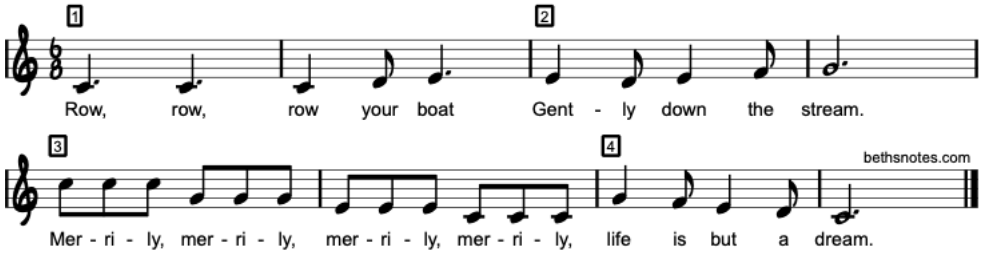


Fig. 1. The nursery rhyme "Row, Row, Row Your Boat"[37]

[37] https://www.bethsnotesplus.com/2013/08/row-row-row-your-boat.html, Accessed on February 1, 2023

It is important to note again that Scruton thinks this spatial metaphor is an indispensable and essential component of our musical experience. That is, we cannot but experience movement in music. It is also an essential ingredient of (Western classical) harmony; behind harmonic concepts like different sorts of cadences or harmonic resolution is the concept of movement—we cannot experience such harmonic elements without hearing movement in music. So, it seems Scruton is right about the indispensability of musical space in our musical experience. We saw that musical movement also plays an essential role in the appearance theory—the theory claims that musical expressiveness is the result of experiencing a similarity between musical movement and expressive movement in human behavior. What is, however, debatable is Scruton's claim that musical space is a metaphor and not a real property of music.

Contrary to Scruton's metaphorical claim, Davies (1994) has argued that our concept of movement has different, but related, meanings: sometimes it is used when a specific object goes from one place to another, and sometimes for non-physical changes, for instance, changes in stock markets or mathematical models. Although these are two different sorts of changes, we understand both as movement. Now, the argument goes, musical movement might be a case of the latter meaning of the term "movement"—movement in a mathematical model, for example. According to such a solution, the fact that we experience a movement in music even though nothing changes its physical place is explained by the general phenomenon of polysemous meanings in our language rather than by metaphor. However, through shifting the problem of musical movement to the linguistic domain, this non-metaphorical approach has its own issues.

It appears that our understanding of the sense of "movement" connected with entities like a geometrical or mathematical model is parasitic on our understanding of the sense of the term connected with physical movement—when a physical object like a ball moves to a different place. It is hard to imagine that one could use the notion of movement for changes, say, in the Dow Jones without an understanding of the sense of movement connected with physical objects. So, it seems that one of the two intended senses of "movement" is primary and the other secondary. And if we wanted to apply the polysemy view to the case of musical expression, we would need to explain how musical movement has a sense parasitic on movement's primary sense.

I think two options are available to a defender of the polysemy view. First, they can explain how these two senses of "movement" are related in the case of music. That is, they can explain why our perceiving movement in music is dependent on our perceiving the regular, physical sense. Without such an explanation, musical movement remains unexplained. Saying that musical movement is like movement in, say, the Dow Jones does not solve the problem. To my knowledge, the main theory of this sort by Davies does not provide any such explanation—it just appeals to the polysemous meanings in language. Or, as a second option, such a polysemy theorist as regards musical movement can deny that there is a primary-secondary relationship between the two senses of "movement" in question. But neither do we see this denial in Davies's theory, nor does it seem to be a plausible approach. (As I said, it would be hard to conceive that the two senses of "movement" are independent, that without having a sense of how a physical object moves one could understand movement in a graph or geometrical model.)

So, a central problem with the polysemy approach is that it does not provide any explanation of how musical movement is related to our perceiving movement in physical, non-musical cases. Particularly if musical movement is going to be compared with movements in human behavior (a case of physical movements) to explain musical emotions—a comparison that grounds, for example, the appearance theory—that relationship cannot be left unexplained. At least a prima facie advantage of the metaphorical approach is that it tries to explain the relationship. The explanation given is through the musical movement's metaphorical relationship with, or dependence on, the regular, physical sense of "movement." We need, though, to examine how successful the approach is.

To understand Scruton's view on how our experience of musical movement, which grounds our experience of musical emotions, is metaphorical, we need to consider his view on the notion of metaphor itself (Scruton 1999, Ch. 3). The gist of the view is that some metaphorical descriptions are the result of a perceiver's imaginative engagement with her object of experience. I said *some* metaphorical descriptions because to my understanding Scruton appears to be primarily concerned with metaphorical descriptions connected with our experience of art objects.

Two related theses can be identified in his view: that metaphorical descriptions are *untranslatable*; and that they are *indispensable*. Let's begin with untranslatability. Scruton believes that metaphorical descriptions (and let me repeat the caveat: successful, aesthetic metaphors) cannot be reduced to, or paraphrased by, a similarity comparison or a likeness. When a predicate *P* is metaphorically attributed to an object *O*, we cannot fully translate "the point" of the metaphor through some comparison, such as "*O* is like *P*

in such and such respects." With such a comparison the metaphor loses its point, holds Scruton.

I think we can agree with Scruton here particularly if we consider (successful) literary metaphors. (Scruton's own examples to corroborate this idea are from literature). For instance, think of a poem replete with metaphors and try to paraphrase all those metaphors—with the help of similarity relationships between the two things that are metaphorically associated. Could we in this way get to a "translation" of the poem into non-metaphorical prose? This is highly unlikely and against the spirit of our response to literature. It would ruin all the joy one could have through imaginative engagement with the poem and its metaphors. In other words, such a translation, even if possible, would eliminate a significant and I believe essential aspect of the poem (its literariness). So, it seems Scruton's untranslatability thesis points to a very important feature of artistic metaphors, or metaphors used in the art domain. Nonetheless, we need to see how exactly Scruton thinks this idea can be applied to *musical* experience and the case of musical emotions, specifically. Before that, let's take up Scruton's indispensability thesis.

We can regard the indispensability and the untranslatability theses as two sides of the same coin. The untranslatability is focused on an artistic object—that its metaphorical description cannot be fully paraphrased by a mere description of its features —while the indispensability is focused on artistic experience. Scruton believes some artistic experiences are indispensably metaphorical:

> [T]here are contexts in which metaphors seem indispensable: not merely because they are part of some unique literary experience […] [But] in particular because we are attempting to describe how the world *seems*, from the point of view of the active imagination. (Scruton 1999, 91)

To Scruton, artistic experiences of the world, particularly those containing metaphors, are not experiences of how the world *is* but how it *seems* to the experiencer. The metaphorical aspect of an artistic experience is connected with this *seeming* aspect. Sometimes the subject has to describe how the world seems to her by appealing to metaphors. The indispensability thesis, to my understanding, comes from such an argument: if an artistic metaphor is not connected with how the object of experience (the art object) is but with how that object seems to the subject or is experienced by her, it is *only* through experiencing the object in a particular way that one could get the point of the metaphor. In other words, the indispensability thesis claims that to get the point of a metaphor one must experience the object in question through the lens of the metaphor.

This can show how the indispensability and untranslatability theses are necessarily connected. If a metaphorical description of an object is about how that object is experienced, or appears to a subject, and not about the object itself, then any description that tells us only about the properties of the object cannot free us from appealing to metaphor or experiencing the object metaphorically. For example, if there is a metaphor related to the sun in a poem, that metaphor primarily is about how the sun seems to a subject and not merely how the sun *is*—that is, what properties it has. Any description of the sun itself would not then enable us to fully appreciate the point of the metaphor. To have that appreciation we need to think of the sun from the lens of that metaphorical description in a particular way. And according to Scruton, that particular way is *imaginative engagement*.

Scruton believes that appreciating a metaphor in general is through imaginative engagement with the object that is to be experienced metaphorically. When we perceive

or experience an object metaphorically as X (for example, some lines and colors in a painting as a tree) we exercise imagination. (Scruton thinks our perceiving a tree in a painting representing a tree is metaphorical. Construing representational art in this way can be contentious, but to clarify his view on metaphor I want to assume its truth.)

Imagination seems to serve as a bridge in Scruton's account: it bridges, in our aesthetic experience, perceiving an object which is not literally X, and experiencing X through it. And this imagination has a special feature for Scruton: our perceiving the (art) object and our imagining X are intertwined. That is, what is perceived and what is imagined are inseparable in our aesthetic experience. For example, in the case of seeing a tree in a painting, we do not perceive some lines and colors and then, in a *separate* experiential episode, imagine a tree: our perceiving the lines and colors in the painting and imagining a tree are united in our experience. (Scruton (1974) calls it "double intentionality"—our experience is doubly intentional in such cases: in the painting example, the experience is both about some lines and colors *and* about an imagined tree in an inseparable manner.)

I do not want here to examine all the details of this aesthetic theory and particularly how it can be applied to different art forms. Scruton's metaphorical-imaginative theory aims to explain a major part of our aesthetic experiences and our descriptions of art objects. Examining it is beyond my aims here. So, let's see how this theory specifically is applicable to the case of musical emotions, to the case of attributing emotions to musical passages or pieces. To recap, we saw the following main lines of thought in Scruton's theory:

(a) Our attribution of emotions to music is metaphorical. The main idea is that music cannot obviously be in any mental state. Nonetheless, there is a significant relation between the emotions attributed to music and the ones attributed to humans when they are in emotional, mental states. What can best explain this relationship is metaphor.

(b) The primary metaphor that is attributed to music and can explain musical emotions is musical space and movement. It is due to (metaphorically) experiencing movement in music that we experience music as being apt for emotional descriptions.

(c) The metaphorical descriptions used are both untranslatable and indispensable. Being applied to the case of music, untranslatability means that metaphorical descriptions (of music) cannot be translated fully to any description of how musical sounds are. And indispensability means that appreciating musical emotions is a matter of how we *hear/experience* musical sounds and not how musical objects are.

(d) Metaphors are the results of a subject's imaginative engagement with an object. In the case of music, it means that hearing emotions in music is the result of how the listener imaginatively responds to music. And the main feature of this imaginative engagement is that exercising imagination and hearing sounds are intertwined in musical experience (the idea of double intentionality).

Among these ideas constitutive of Scruton's account for musical emotions, I basically agree with (c) but I think there are serious concerns with (b) and (d). In my opinion, we can grant Scruton (a) in so far as his theory—particularly its imaginative

idea—can explain convincingly what the relationship between musical emotions and emotions as mental, psychological states is. In fact, I do have some reservations about Scruton's metaphorical claim in (a). For one, Scruton does not provide persuasive reasons why the fact that music cannot be literally in any mental state makes any emotional description of music metaphorical. Also, the fact that imagination seems to play some role in our hearing a space in music, in relation to which all musical elements like melody and harmony are experienced, does not turn musical space into a metaphorical concept. As Levinson (2000, 610) himself has mentioned:

> […] that such imaginative hearing of sound is illuminated by appeal to the basically *linguistic* idea of metaphor seems insufficiently justified. It seems we can think of musical properties—for example, the downward movement of a bass line—as emerging for imaginative perception without thinking of them as the upshot of some unobserved process of metaphor making.

I believe, however, that as a starting point Scruton's metaphorical thesis is not ungrounded. Besides its philosophical appeal, the metaphorical idea has been used in explaining musical emotions from the perspective of music cognition as well.[38] The same intuition seems to be used in such approaches: music cannot be literally in an emotional state and it must then be connected with emotions in some other way. I do not think this intuition itself is unwarranted. What is important is the explanation that comes after it. If through the concept of metaphor a theory can plausibly explain how exactly the way(s) in which we hear emotions in music is/are related to the way(s) in which we experience the expression of the same emotions by humans, it is, of course, a great success. But the issue is that because there is not a unified, universally accepted theory of metaphor, what a

[38] For example, with the help of some theories in cognitive linguistics, Steve Larson (2012) has argued that there is a cognitive basis for the idea that our understanding musical movement is the result of how we cognize musical time metaphorically in terms of space.

metaphorical theory of musical expression should answer is *how* metaphor figures in our experience of musical expression. In other words, it should provide a plausible theory of metaphor suitable for the case of music or, specifically, musical experience. With regard to Scruton's view, this implies that the burden of explanation is on his imaginative account of metaphor, (d) above; that is, we need to turn to that account to see if his explanation of musical expression (based on metaphor) enjoys plausibility.

We saw earlier that the gist of Scruton's view on metaphor is that a metaphorical experience of something requires an imaginative response to or engagement with it. So, if X is metaphorically Y, one needs to imaginatively engage with X to experience it as Y. Now, the question is: what is this "imaginative engagement" in the case of hearing emotions in music? To Scruton, such imaginative engagement seems to be the listener's sympathetic response to expressive music (from the passage quoted in the beginning of this section): "Our response to music is a sympathetic response: a response to human life, imagined in the sounds we hear." So, Scruton thinks that our hearing emotions in music is the result of our imagining *something* "in the sounds we hear": what is imagined is "human life." But he does not seem to think that what is imagined in the sounds we hear—"human life"—is something separate from the sounds heard, or something separate from music itself. This gets illuminated, I think, in how he qualifies what he says about sympathetic-imaginative response to music: "However, in the absence of representation there is no precise object of sympathy—neither an imaginary human subject, nor a situation perceived through his eyes. The life in music belongs in the musical process, abstract, indeterminate, unowned except through the act whereby we listeners possess it."

It seems then that Scruton wants to reconcile two apparently opposing ideas: that something is imagined in the sounds of music, which causes the listener to have a sympathetic response to music; and that the listener does not sympathize with something imaginary beyond or except for music itself. Scruton then believes that imagining something in music (human life) is not but perceiving music itself. But how is this possible considering that music (itself) obviously is not a human life? It seems that if we are to imagine human life in music, we have to undergo an imaginative experience distinct from, while in parallel and concurrent with, perceiving music itself. But it is doubtful that we necessarily engage in such imagination while listening to music.

Scruton might be able to remedy this challenge through his metaphorical-imaginative theory of musical experience explained earlier. As we saw, he thinks the perception of music is itself imaginative; that is, in perceiving music as music a listener *imagines* musical sounds in a musical space. To Scruton, this imagination is in fact a prerequisite to perceiving music as music. It then appears that Scruton does not face any problem—at least in the framework of his own theory of musical experience—when he considers imaginative engagement with and perceiving music as united. But can it really justify how he solves the problem of musical expression via imagination? I believe not.

The problem is that even if we grant Scruton his imaginative-metaphorical theory of musical experience[39], it does not entail that in experiencing musical emotions the listener imagines *human life* in music. Scruton's theory tells us that our perceiving music is infused by our imagining a musical space; but this imagination is not identical with the imagination of human life in (expressive) music. There is an explanatory gap in Scruton's

[39] For a critical examination of Scruton's metaphorical thesis, see Boghossian (2002). I am here particularly concerned with Scruton's imaginative proposal for musical expression.

suggestion for musical expression. How exactly does imagining musical space lead a listener to imagining human life? If the idea is that imagining musical space is in fact imagining movement in music and it is through imagining this movement that we imagine the movement of human life in music, the explanation is still incomplete. Although movement is an essential component of human life, it is not identical with it; obviously, not any movement is the movement of human life. Scruton still owes us an explanation of how human life is imagined in music. Here we can see that Scruton's imaginative project for musical experience, though insightful in other respects, is toothless with regard to the problem of musical emotions. It cannot convincingly explain how we experience musical expression.

However, although Scruton's imaginative proposal is unable to explain the experience of musical expression, it has a merit compared with Levinson's imaginative view discussed earlier. As discussed, Scruton argues that the imagination involved in musical expression does not have any representational character: it is not imagining a thing represented by music. Scruton tries to propose a kind of imagination that is totally infused with and inseparable from perceiving music. Whether or not he is able to elucidate this kind of imagination successfully, the mere idea seems to be more promising than Levinson's intended imagination.

We saw that a main problem with Levinson's imaginative account is that it is phenomenologically doubtful that in experiencing expression we always imagine an agent—however indefinite—who expresses that emotion through music. The phenomenology of our musical experience hardly suggests such a thing. I think Scruton wants to avoid this problem through his intended kind of imagination. He wants to appeal

to a kind of imagination involved in musical experience that is infused with the perception of music in such a way that they become phenomenologically indistinguishable. As an imaginative approach to musical experience this seems to be an advantage over Levinson's for it does not require a listener to imagine anything besides and beyond perceiving music itself. I believe if imagination were to play any role in experiencing musical expression, it should be construed somehow in line with Scruton's imaginative idea rather than Levinson's for the reasons given.

We now are at a point to see our evaluations of Scruton's view on musical expression from a more general perspective. Scruton gives a metaphorical theory of musical emotions and to explain this metaphorical basis, he appeals to imagination. So, at its root Scruton's theory can count as an imaginative theory and its merits or disadvantages should be found there. He appeals to a kind of imagination, infused with musical perception, which is itself more promising than imaginative approaches like Levinson's. However, the main problem is that it is not clear how this way of imagination in the fabric of musical perception leads to sympathetic reaction to music, a sympathy that grounds Scruton's explanation of musical expression.

3 Musical Emotions as Music's Dispositional Capacity

I want now to turn to another idea central to some theories in explaining musical expression—the third and last idea we are going to consider in this chapter. The Dispositional Idea (DI), as mentioned, is this: musical expression should be explained in terms of the dispositional capacity of expressive music. The theories that appeal to this idea mainly consider a listener's experience of musical expression the result of music's

capacity to dispose the listener to experience a feeling, or feeling-like mental state, connected with the emotion music is expressive of.

The dispositional theories are in fact the sophisticated versions of the traditional arousal theory. Remember that according to the arousal theory, the crude version, a musical passage or work is expressive of, say, happiness when it arouses happiness in a listener; that is, musical happiness is seen as being identical with a listener's feeling of happiness. The sophisticated versions of the view are different from the crude version in three main respects: (1) they consider the mental state a listener goes into due to listening to expressive music as being intentional and directed toward music; (2) they do not necessarily consider the mental state a listener goes into due to listening to expressive music as a full-blooded, emotional state; and (3) they do not necessarily consider the mental state a listener goes into due to listening to expressive music as being identical with what music is expressive of (a melancholy passage does not necessarily dispose a listener to have a mental state akin to melancholy). In what follows, we will see how these three features come up in the theories we will discuss.

I intend to consider two main dispositional theories proposed by Kendall Walton (1988, 1990, 1997, 1999), and Derek Matravers (1998). In line with the two previous sections, the aim is to understand whether and how through the dispositional idea the intended theories can shed light on the nature of our experience of musical expression. I begin with Walton's view.

3.1 Empathizing with Music

Walton's view on musical expression can be traced in his discussions of the semantic dimension of music (Walton 1988), musical tension (Walton 1999), imaginative listening (1997), and musical depiction (Walton 1990, chap. 8). Particularly connected with musical tension, we can see his developed theory of musical expression, as he (1999, 436) thinks musical tension and relaxation "have a lot to do with music's expression of emotions […]."[40]

His focus consistently is on how musical emotions are experienced rather than what in music, or what non-musical properties, bring that experience about. As I quoted him in the first chapter, to him, "[i]t is one thing to identify the sources of tension in music, and quite another to say what musical tension is." Regarding the sources of musical emotion he mentions different factors. For example, he points out that sometimes the source of musical emotion is music's "imitating behavioral [or vocal] expressions of feeling" (1988, 359). But in all his explorations of musical expression, he tries to shed light on is the *experience* through which one hears emotions in music. In what follows I give my interpretation and understanding of Walton's view by referring to what he says about musical experience generally and experiencing musical emotions more specifically.

The gist of his view is that a listener hears emotions in music due to her *empathetic* response to it. In the core of this empathetic account is the idea that a listener hears emotions in music through an awareness of *her own* mental state, a state which is basically akin to feeling an emotion. So, for example, when I hear tension in a musical

[40] He also thinks that musical tension and relaxation "contribute enormously to the representational functions music often serves." (436) But the representational capacity of music is not our primary concern in this section.

passage, I first become aware of a mental state like tension in me and then, due to that introspection, I become aware of the tension in music (or the source of my tension). The notion of *introspection* is then central in Walton's account: experiencing musical expression is grounded in an introspective awareness.

Walton thinks that one's introspective awareness while encountering expressive music is basically similar to her introspective awareness when she empathizes with other people. He argues that our empathy with others is based on *mental simulation* through which *"*we attribute psychological states to others on the basis of our experience of the same psychological states, or at least similar or corresponding or analogous ones." (1999, 418) The psychological state one experiences through this process of simulation informs one about the mental state of the other one empathizes with. That is, through empathy one forms a judgement that has a "demonstrative element": "the other person is in a mental state like *this*" in which *this* refers to one's own mental state. Not that one consciously forms such a judgment in empathizing with someone else[41]; what seems to be important to Walton is that one's becoming *aware of* the other person's psychology is through introspection.

Walton also believes we can have empathy with inanimate objects. For example, in seeing a pillar bearing a lot of load, one might empathize with it through simulating the situation the pillar is in and imagining the pressure it is under. In fact, she can become aware of the pressure the pillar is under by her imagining being in the situation the pillar is in. Walton does not think that in such cases one mistakenly thinks that she is empathizing with a psychological being. He wants to show that even in our encounter

[41] He in fact thinks that this awareness is nonconceptual (1999, 426).

with insentient things, sometimes we become aware of some of their properties via introspective awareness. He says: "Our empathetic responses to inanimate objects are often grounds for attributing properties to them. And there may be something like projecting features of one's experience onto inanimate objects." (1999, 429) In the pillar case, the person knows about the pressure the pillar is under through introspecting her own mental state: through imaginative feeling of pressure when she puts herself "in the pillar's shoes." She would in fact know about the physical state of the pillar—say, being physically under pressure—via her own psychological state.

But how is this outwardly directed, introspective awareness, which Walton thinks underlies empathy, applicable to the case of musical expression? For us to have empathy with someone (or something), there is typically a *ground* for empathy. That is, in an empathic response there is a ground on which one can put herself into someone or something else's shoes. When I empathize with my friend who has just lost her job, my imagining myself in her *situation* is the ground of my empathy (that she has no job, she has no savings, she has to pay her rent this month, etc.). Or, if we wanted to accept Walton's suggestion that sometimes we empathize with, say, a pillar under a heavy load, again there is a ground for our empathy: the situation the object is in, a situation that we perceive when we encounter it.

Is there a ground for our empathetic response to expressive music and, as a result, becoming aware of its emotional quality through introspection? In the case of music, we seem to only hear sounds. How or on which ground then can this hearing lead us to have empathy and become introspectively aware of what music is expressive of? Regarding the non-musical cases we considered above, there is something to imagine, an

imagination which then leads to an introspective awareness of some quality in the object or person one encounters. But what in hearing musical sounds leads a listener to have such introspection? This is a challenge that Walton has to deal with in applying his view on empathy to the case of expressive music. He himself, in my understanding, puts the challenge in this way: "Hearing *sounds* may differ too much from introspecting for us comfortably to imagine of our *hearing the music* that it is an experience of being aware of our states of mind." (1988, 359) I posed this challenge before examining Walton's view on musical expression because understanding that view depends on how Walton responds. And in his related work, Walton has given us two responses.

In "Projectivism, Empathy, and Musical Tension" Walton argues that in having empathy our emotional response is *undeliberate* and *immediate*. He thinks we do not necessarily first judge someone/something's circumstances and then empathize with them/it: the person or that thing sometimes *infects* us emotionally. What happens in such cases is *motor mimicry*, "[t]he nearly universal tendency to mimic, more or less automatically, the expressive behavior of other people, their facial expressions, gestures, and bodily attitudes […]". (1999, 420) Or, sometimes we engage in imagining someone's situation automatically and without deliberation (ibid., 425). He thinks something similar happens in the case of expressive music, an idea that grounds his view on musical tension:

> Music induces emotional infection, often, without any indication of an otherwise identifiable person. So I have an (imaginative) experience as of there merely being one or more nervous [persons] or nervous crowds around. Since I use my own nervousness to represent the state I experience myself being infected by, I can characterize my experience as one of there being one or more persons (or groups) around, who are in *this* state, the state I am in. […]

> Tense music infects me with tension, but the music is likely not to suggest, at all clearly, either a tense person as opposed to a tense inanimate object or the reverse. […] *Musical tension* is the property of being apt to elicit an experience of this kind. (ibid., 433)

There are two important themes in what Walton says here: one related to the kind of imagination involved in experiencing musical tension, and the other is the idea that that experience is the result of being emotionally infected by music. I will return to the former theme (imagination) later. The latter is related to our aim to understand how Walton responds to the challenge I posed earlier: on which ground do we empathize with, or emotionally respond to, expressive music?

In my understanding, Walton thinks in our empathetic response to expressive music we do not need to be aware of the ground: the response is undeliberate and automatic, and the result of our being emotionally *infected* by music. In so infecting us, music "elicits" a kind of empathetic experience in us, an experience of feeling (imaginatively or actually) the emotion music is expressive of. For example, a tense passage elicits in us a feeling (imaginatively or actually) of tension. Now, we become introspectively aware of our own tension, but, as we saw, Walton believes that through this introspection we become aware of the tension *in* music. ("I use my own nervousness to represent the state I experience myself being infected by.") What I want to highlight here is not this outwardly directed introspection, but Walton's idea that the listener gets to such an introspection *automatically* and in an *undeliberate* way. That is, she does not need to know on which ground and how tense music elicits such an (empathetic) experience in her.

It is important to notice that Walton does not deny that there is ground for our empathetic response to music; he seems to think that a listener does not need to be aware

of this ground. He mentions at different points that expressiveness of music comes sometimes from music's capacity to resemble human behavior, and sometimes from the sonic qualities of musical elements such as dissonant or consonant chords (1999, 434; 1988, 359). But a listener does not need to be aware of any of these: her experience is of "there being one or more persons (or groups) around, who are in *this* state, the state I [the listener] am in". (That is, through listening to (expressive) music the listener feels that there are people around her being in the psychological state she is in.)

Now, I want to consider a different way Walton has tried to deal with the challenge we raised earlier, that is: how does hearing musical sounds get us to a kind of introspection that makes us aware of musical emotions (or emotions in music)? We just saw that Walton's first solution is based on the idea that music emotionally *infects* us and elicits an empathetic response in us. The central idea of his other solution, which is less developed compared to the first, is sketched in the following passage a part of which was quoted earlier:

> Hearing *sounds* may differ too much from introspecting for us comfortably to imagine of our *hearing the music* that it is an experience of being aware of our states of mind. My suggestion is that we imagine this of our actual introspective awareness of auditory sensations. (1988, 359)

To understand this suggestion, I think we should refer to how Walton construes auditory sensation as being peculiar and distinct in comparison with other kinds of sensations, particularly visual sensations. Let us reflect on what he says in this relation in his *Mimesis as Make-Believe*:

> […] in the case of much "expressive" music, it may be fictional not that one sees or hears or otherwise perceives external things but that one experiences or is aware of (one's own) feelings or emotions or sensations or sentiments or moods. The listener imagines experiencing excitement, passion, fervor, despair, conflict, feelings of

> exuberance, of striving, of determination, of well-being, of trepidation, of repose. Moreover, I suggest, it may be fictional of one's actual awareness of auditory sensations (not one's perception of sounds) that it is an awareness of such feelings. In place of fictional perception of external objects we have fictional introspection or self-awareness. […]
> (The spectator of a picture may also, fictionally, be aware of feelings. But it is not fictional of her experiencing of visual sensations that that is an experiencing of her feelings. It strikes me that auditory sensations are somehow more suitable for this role than visual ones are, perhaps because introspecting is in some way more like hearing than like seeing.) (1990, 335-6)

The range of experiences expressive music can cause one to have is considered to be very broad here (well-being, fervor, trepidation, …). But let us set that aside and focus on the very mechanism through which Walton thinks music can bring such experiences about or elicit them in a listener. I am specifically concerned with the contrast Walton makes between auditory and visual sensations, the contrast that I reckon plays the central role in his suggestion about expressive music. He thinks auditory sensations are somehow closer to our introspective life than visual sensations. It is fruitful to reflect a little on what exactly Walton's observation is with regard to the phenomenology of hearing sounds. It will then become clearer why Walton thinks hearing musical sounds causes emotional introspection.

Walton seems to think that there is the following peculiarity about auditory sensations compared with visual sensations: we "feel" them *in* us more than visual sensations. Let's take and compare two perceptual cases, one auditory and one visual. Seeing a red apple causes different visual sensations in a perceiver: a patch of red color, a specific shape, etc. Now, let's imagine an auditory perception: hearing rain pitter-pattering on a roof. Again, there are auditory sensations connected with this auditory experience: the quick succession of sounds, a sound (timbral) quality that is related to the

physical properties of the roof, etc. Now, one thing is quite obvious with regard to the visual case. When one sees an apple, typically one does not introspectively attend to the visual sensations she has: the attention is directed toward the apple itself—the *source* of the visual sensations—and not the sensations. Phenomenologically, in such visual cases the so-called visual sensations are experienced as the qualities *of* the object seen and not things that are felt. (Such qualities are experienced as "tethered to the surface of the object," figuratively speaking.) Walton appears to think that phenomenologically something different happens in our imagined auditory case: in hearing rain pitter-pattering on the roof, the hearer seems to (introspectively) attend more to her auditory sensations than to the source(s) of such sensations. To Walton, the hearing here is more like introspection than outward attention which happens in the case of seeing an apple, for example.

I will later examine this phenomenological observation about hearing; but before that it is worth considering how the observation grounds Walton's idea that hearing sounds of music is like "introspective awareness of auditory sensations." As we saw, Walton thinks that in auditory perception the perceiver has an introspectively intimate connection with the related auditory sensations. Based on this phenomenological assumption, he construes hearing musical sounds as an introspective act, particularly: introspective awareness of "(one's own) feelings or emotions or sensations or sentiments or moods." This is in fact Walton's suggestion concerning musical expression in the passage quoted above, from his *Mimesis as Make-Believe*.

At this point, we are in a good position to see more clearly how Walton has responded to the aforementioned problem: how does perceiving musical sounds cause

one to be introspectively aware of feelings? As explicated, one response to this question is given in the context of his empathetic account and the other based on his phenomenological observation about auditory perception. They are:

(1) Music *emotionally infects* us; and this emotional infection is automatic and undeliberate. Through such infection, music elicits empathetic (imaginative) response in us, it gets us to introspect feelings which we attribute, as emotional qualities, to music itself.

(2) Music is an auditory experience. There is something phenomenologically peculiar about such experiences: hearing sounds is like introspection. This is why hearing musical sounds can lead to introspective awareness of the auditory sensations it causes.

The core idea in both (1) and (2) is that music can elicit emotional introspection in us. It has, in other words, a dispositional power to make us be introspectively aware of some sensations and feelings. However, there are a few things that are not clear in what Walton says about musical emotions and one serious problem with his overall view (or the core idea of his suggestions).

To begin with, I am not sure how exactly we can unite Walton's empathetic account, (1), and what he says about auditory perception—particularly, perceiving musical sounds—in (2). He himself, to my knowledge, has not explicitly united these two in his work; and it is not my aim here either. It strikes me though that Walton's main account of musical expression is his empathetic account and what he says about the phenomenology of hearing somehow bolsters that account. (*Not only does music have this power to emotionally infect us and elicit empathetic responses in us, but the very*

perceiving of musical sounds is also akin to introspecting sensations or feelings.) But in what follows, I am going to examine each idea separately, beginning with (2).

One thing is quite clear with regard to (2): it is not a complete theory of musical expression. Since Walton himself does not seem to claim such a thing, I do not want to pose this as a problem with his suggestion but merely point out what it needs to become complete. The gap in the suggestion is this. Walton's phenomenological idea is that perceiving (musical) sounds is like being introspectively aware of auditory sensations. This, however, does not automatically lead to the idea that in musical perception such sensations are experienced (imaginatively) as *feelings/emo*tions. Not every sensation is emotional. For example, even if we suppose that in hearing rain pitter-pattering on a roof one becomes introspectively aware of, or feels, the related auditory sensations, those sensations are not experienced as feelings/emotions. Then, even if Walton's idea about the phenomenology of hearing is true, he still needs to show why the auditory sensations that are felt in hearing musical sounds are like experiencing (imaginatively or actually) emotions. Otherwise, the suggestion is incomplete as a theory of musical emotions.[42]

However, regardless of how complete Walton's suggestion is, my main worry is about his phenomenological observation itself. In fact, it is not phenomenologically clear to me why hearing is closer to introspection than seeing. The idea might be the following (remember the apple and rain examples). In hearing one attends more to the auditory sensations than the source of them (that is, the source that causes them), while in seeing it is the reverse: one's attention is toward the source and not the sensations—all the

[42] Although Walton does not give such an argument, the idea might be that it is movement in music that causes a listener to experience the (musical) auditory sensations as feelings or emotions. But this needs to be developed by more arguments related to musical movement and its being felt by a listener.

sensations are experienced as qualities of the source. I am not convinced, however, that it is phenomenologically the case. It does not phenomenologically hold to me that in our typical auditory experiences, we attend to sounds mostly independently of their sources. A sound is always a sound *of* something from somewhere, even though most of the time it does not give us information about its source or the source's location as accurately as visual sensations. For example, in the rain case, a person normally hears the sound(s) as connected with what is happening on the roof. And this is basically similar to what happens in seeing, say, a red apple: the relevant color sensation is experienced as connected with the source—as a property *of* the apple, specifically speaking.

So it seems that both visual and auditory perceptions are outwardly directed basically in a similar way. (As mentioned, it is beside the point which one results in more accurate information about the object of perception.) If this is the case, there is little reason to accept Walton's phenomenological observation that hearing is more introspective than seeing. And this undermines Walton's musical suggestion, (2), that perceiving musical sounds is similar to introspective awareness of sensations or feelings because of the peculiar phenomenological character of auditory perception.

Now, I want to consider (1), Walton's more developed account of musical expressiveness based on empathy. A main worry, I think, with regarding musical expressiveness as emotional infection is that it totally divorces the experience of musical emotions from musical understanding in the following sense. In line with some other philosophers, I identify musical understanding with *ways of listening to music.*[43] Music is not merely hearing a collection and succession of sounds; listening to it is essentially a

[43] As Levinson has pointed out: "Understanding music, as a number of philosophers have recently observed, is fundamentally a matter of hearing it a certain way." (*Music in the Moment*, p. 29) See also Tanner and Budd (1985).

matter of how these sounds are attended to and organized in one's mind.[44] (I do not intend here to examine this matter in detail—I will return to it in the next chapter.) With regard to musical expression, it seems that the process of attributing emotional qualities to music is dependent on this attention and organization, that is, the *manner* of listening even though a listener is not conscious of it.

In Walton's analysis of experiencing musical emotions, it is missing that the experience is connected with how music is attended to. If expressive music emotionally infects a listener and the process is undeliberate and automatic, what should we say about a case in which two people listen to the same music and attribute to it two different emotions because each feels she has been infected by the music differently? (And suppose both the listeners are familiar with the relevant musical idiom but one of the attributions are musically irrelevant.) According to Walton's account, I think we would have to say that both the listeners are justified in their attribution. The account provides neither for us, from outside, to judge which attribution is musically justified nor for the listeners themselves to try to persuade each other which one has a better musical justification for her attribution. Let us see how.

Suppose one wanted to adjudicate between the two emotional attributions by considering which listener has listened to music in a proper way (whether, say, she has attended to the right and expressively relevant parts or properties of the music). Although such a consideration can probably guide our judge to find which listener's emotional attribution to the intended music is musically justified, it is outside Walton's account. That is to say, the account does not regard such considerations as relevant to the

[44] See Davies (1994), Ch. 7

experience of musical emotions: as long as a musical passage can infect a listener emotionally and cause her to empathize with it, the listener's experience is (musically) justified.[45] Walton's suggestion does not go further than this.

Now, suppose our two listeners themselves wanted to persuade each other of why each is justified in hearing the emotion he hears in the music. Again, it appears that they would have to give considerations similar to the ones our imagined judge would do. Not that they would necessarily have to give technical or formal considerations: they might just draw each other's attention to some parts, qualities, or even associations music makes with non-musical matters. But again, such considerations would take us to territory beyond Walton's theory. As long as the two listeners are from the same musical culture and familiar with the kind of music in question, they are both justified.

Here, it might be said that attributing emotional qualities to music is not a matter of justification or deliberation. A listener does not need to (musically) reason or deliberate in order to hear emotions in music: it is something that naturally and automatically happens to a person familiar with the kind of music in question. This is a plausible observation, I concede. But this does not show that there is not any justificatory element in hearing emotions in music. The fact that a listener does not *first* deliberate and *then* attribute emotional qualities to music does not entail that there is not a justificatory element in the attribution. There are psychological studies that show music listeners—with or without technical/musical background—are inclined to justify or enter some form of deliberation about why they hear an emotion in the music they are familiar with or

[45] I do not mean that Walton's account is of *when* the listener's experience toward expressive music is justified. His theory, in my understanding, is about the *phenomenology* of an appropriate (justified) response to expressive music and does not reveal anything about that appropriateness or justification.

like.[46] So, a disregard of the justificatory aspect of the experience of musical expression is a demerit for a theory. And I tried to show why there is a serious worry about Walton's account from this perspective.

It might also be objected that the problem I have raised against Walton's suggestion is based on an unfair reading of it. The theory does not claim that expressive music emotionally infects *any* listener, even the ones who are inattentive and are not culturally backgrounded as regards the music in question. The true reading of the view is something like this: "When (if) a listener is culturally competent, and she is attentive enough, she can emotionally become infected by the music and empathize with it." I agree and in fact think the spirit of Walton's theory should be read in this conditional manner. But still it does not alleviate the worry I have raised. For an important part of the problem (the problem of musical expression) lies in and is related to the *condition* described ("When (if) a listener..."). As I tried to illuminate above, *how* an attentive (culturally competent) listener hears sounds in music as a result of which she attributes emotions to the music cannot be separated from an understanding of the phenomenon, our experience of musical expressiveness. This matter cannot be merely left as an assumption or general condition in a theory. And this occurs in Walton's account, at least in the conditional reading given above. The condition in that reading is too broad and too general to be informative. It cannot spell out what kind of attention to expressive music or what way of listening to it results in empathizing with it.

However, setting the above problem aside, there is an insightful aspect in Walton's musical analyses in general: his drawing our attention towards the introspective element

[46] See, for example, Peltola (2017).

of musical experience. I think Walton is successful in highlighting what is interestingly peculiar about that experience in comparison with experiencing other kinds of art, particularly the representational ones. In reading literature, watching a movie, or seeing a painting, our imagination is typically involved through being directed towards what exists (imaginatively or fictionally) in the world of the work: characters, events, different relationships between characters, visual elements of a scene, etc. But in experiencing music (particularly instrumental music), there is no determinate musical world to attend to; and it does not seem to be phenomenologically the case that all our attention is directed to *musical* mechanisms in a work (the kinds of chords involved, the harmonic progression, the specific characterization of orchestration, modulation, etc.). To my understanding, this is why Walton tries to unravel and characterize an introspective element in musical experience, to develop the idea that an essential part of our engagement with music is our attending to our *own* experience and its ebbs and flows. And I think this idea in fact grounds his account of our emotional engagement with music (the empathetic account). Although in this section I have argued that there is a serious worry about the account, I am sympathetic to Walton's phenomenological observation about the introspective element of our engagement with music, and I will return to it more carefully in the next chapter.

3.2 Arousal of Feelings

Another main theory of musical expression based on the dispositional idea is Derek Matravers's arousal theory (1998, 149):

> [A] piece of music expresses an emotion *e* if it causes a listener to experience a feeling α, where α is the feeling component of the emotion it would be

> appropriate to feel (in the central case) when faced with a person expressing *e*. Hence, music expressive of sadness will cause a listener to feel sad (or maybe the feeling component of pity if that is a different state), and this feeling will cause him to believe that the work expresses sadness.

According to this theory, expressive music arouses a *feeling* in the listener. This is a main difference of the account from Walton's: Matravers thinks what is aroused by expressive music is always a feeling and not the listener's imagining having a feeling. So, unlike Walton's account, in Matravers's theory imagination does not play any role. It is claimed that music expressive of, say, sadness causes a listener to experience a feeling, the one that she would feel "when faced with a person expressing [sadness]". The listener becomes aware of what music is expressive of through what she herself feels, because in real life when she has that feeling it is in response to a particular emotional expression. (My feeling sadness or pity is a typical response to someone's expression of sadness; so when I feel that/those feeling(s) while listening to music it is because of the music's expression of sadness.)

Two questions can be raised here: (1) Do we always or normally respond to the expression of an emotion (by a person) by experiencing one, or a few, specific feelings? (2) What is the nature of what we experience on such occasions?[47] Matravers's theory is based on a positive answer to (1). Notice that his account of musical expression is parasitic on our typical affective responses to people's emotional expression. If it cannot be assumed that there is regularity (at least to a good extent) in the ways in which we affectively respond to others' emotional expression, Matravers's account of musical expression collapses. The assumption in his account is that feeling, say, sadness or pity is

[47] My discussion of these two questions as regards Matravers's theory is indebted to Kivy's (2001) critical examination of that account.

normal in response to someone's (expression of) sadness. But there is serious worry in assuming such regularities.

How exactly we respond to someone's emotional expression seems to depend on the relevant situation. For instance, if I saw that someone is sad because she has failed to achieve something unethical, I would not become sad or feel sadness—I might even become happy. Also, my feeling pity is not always (or even most of the time) for someone's sadness. I might feel pity for someone's happiness—for example, when the person lives in a very bad social-economic condition and is happy for attaining the very basic necessities of life. Put generally, our response to someone's emotional expression is usually dependent on what we think of what the other person's emotion is *about*. And this causes irregularity between what one expresses and how we affectively respond to it. This consideration is also related to question (2) above, on the nature of the affective response to emotional expression.

In real life, when one experiences sadness in response to a sad person, most of the time what is experienced is the *emotion* of sadness. And such an emotion, according to a mostly accepted view on emotions, is constituted by both a feeling sensation and a cognitive component (belief or judgment).[48] For example, imagine Sina has become sad in response to her friend's sadness about losing her job. Here, besides what Sina feels (a sensation connected with sadness), there is a judgment in his sadness, a judgment connected with his friend's situation: he thinks/believes/judges that his friend's losing her job is, let us say, painful or unjust. Similarly, on many occasions, we experience an *emotion* in response to someone else's emotion.

[48] See Pitcher (1965)

Now, in Matravers's account the affective response to expression does not contain any cognitive component: it consists only of the "feeling component" of an emotion. This feeling component seems to be read as the sensation connected with an emotion (the sensations we feel when we are sad, for example). So, to him something like the following happens when one hears, say, sadness in music: sad music arouses in one a feeling, a sensation, which is the sensation she typically feels when confronted with another who is sad; and this then causes her to believe that the music is sad because she has felt that sensation as a result of listening to the music and she feels the sensation when she typically responds to people expressing sadness. We should not, though, read the last part as an inference made by the listener in hearing the musical emotion. Matravers's idea, at least in my understanding, is that a listener's attributing an emotion to music, based on the sensation she has felt, is automatic: the listener is accustomed to do the same in real life when she emotionally reacts to people's emotional expression.

But now, if we are to understand the account in this way, it seems we need to qualify it as follows: hearing an emotion in music is parasitic on the cases in real life in which we affectively respond to a person's emotional expression in an *undeliberate* and *automatic* way. For if we did not confine the account in this way, it would be undermined by the regularity problem discussed earlier. We showed that in many cases our affective reactions to other people's emotional expression are based on the beliefs/thoughts/judgments we form about the objects of the emotions expressed. And such deliberations can cause irregularity in the ways in which we affectively react to others' expression: sometimes it is appropriate to react to sadness with sadness, or pity, and sometimes with other emotions depending on what the (expressed) sadness is about.

(For example, in the case imagined earlier happiness can sometimes be an appropriate reaction to sadness).

I think, however, that the qualified reading of Matravers's account is more promising for facing the regularity problem. If we only consider our undeliberate affective responses to others' emotional expression, it seems more regularities can be found. Maybe our undeliberate reaction to sadness is typically feeling sadness or pity as our undeliberate reaction to someone's happiness and laughter is an automatic smile. Such reactions seem to be acontextual and non-judgmental and more related to how we are biologically wired (perhaps, motor mimicry plays a role). So, perhaps removing contextual consideration and deliberation can bring about the regularity of emotional cases in real life that Matravers's account of musical expression needs in order to get off the ground.

Yet, I do not want here to suggest that the qualified reading of the account is closer to Matravers's core idea. Particularly, what makes having such a reading of the account difficult is what Matravers says in the formulation of his theory, quoted above: "[…] where α is the feeling component of the emotion it would be appropriate to feel […]". The issue is related to his use of the term "appropriate," which allows us to consider deliberation in emotional reaction. (We can, as discussed, imagine cases in which it is inappropriate to react to someone's sadness by feeling sadness). If we should feel the feeling component of an *appropriate* emotion in responding to expressive music, how should we know our emotional reaction is appropriate if our reaction is unreflective (which is suggested in the qualified reading)? I am not sure then whether we can have the qualified reading of Matravers's account.

It is, however, obvious that to Matravers deliberation does not play any role in a listener's affective response to expressive music. After all, his attempt to revitalize the old arousal theory is in part motivated by the problems he sees in the cognitivist theories of musical expression. (The cognitivist view, generally speaking, is based on the idea that hearing musical expression is the result of recognizing properties *in* music. For example, the resemblance theory examined earlier is such a theory in claiming that hearing expression is the result of recognizing in music a similarity between music and the appearance of human expression.) One of his criticisms is worth considering here as it draws our attention to a phenomenological worry in cognitivist views like Davies's and Kivy's.

One of the interesting criticisms Matravers levels against the cognitivist view is that the expressiveness of a piece is not *a* property of it which is recognized by the listener at a particular moment of her listening experience. He challenges the view in this way:

> The problem is that in contrast to asking when, during the experience of the performance, you heard the emotion in the music, it does seem sensible to ask when, during the experience of the performance, you recognized some property of the music. That will be the time at which the belief is acquired. The acquisition of a belief at a time is not, however, an enduring unitary experience. Hence, it is difficult to see how the cognitivist could develop an account of the *experience* of expression over time in terms of the *recognition* that the music had some property. (1998, 120-121)

Experiencing musical expression is, holds Matravers, an enduring experience and not something that happens at a particular moment of listening. If the experience was, as the cognitivist claims, the recognition of a property in music, the listener would have that experience the moment at which she recognizes the relevant property. But this is a

momentary experience (recognizing the relevant property and forming the belief that the music is expressive) and not an enduring experience, which is, thinks Matravers, a characteristic of the experience of musical expression. To Matravers, the cognitivist then faces a problem in explaining how construing the experience of musical expression as the recognition of a property in music can capture the endurance of the experience.

I think Matravers is generally right but with some qualification. First, it is not phenomenologically obvious that our experience of expression is always enduring. Sometimes, just by hearing one or two musical qualities we get the expressiveness of music. Perhaps, in such cases the cognitivist is right: recognizing the relevant property leads one to think or believe that a part of a work is expressive. Second, we should be careful in what we ascribe to cognitivism. The cognitivist does not think that what is responsible for musical expression is a property of music that can be momentarily recognized. An expressive property for the cognitivist is a holistic or emergent property of music, a dynamic, temporal property that supervenes on some more basic musical properties.[49] By recognition, the cognitivist then means appreciating such temporally emergent properties in music.

What I believe Matravers is right about, though, is that the cognitivist owes us an explanation for the temporally dynamic nature of the experience of musical expression. It is not clear in the cognitivist theory how exactly a listener experiences musical expression by attending to the relevant musical properties that, in many cases, unfold in the process of listening. If expressiveness, as the cognitivist says, is an emergent property of music, we need to know what the experience of such properties is like and how they

[49] See the passage quoted from Davies in this regard.

can endure in the listening experience. Matravers's worry then plausibly applies here: it must be specified by the cognitivist how the recognition of musical expression is not in contrast with the experience of musical expression which, in many cases, is enduring and not momentary. I do not here intend to consider how a cognitivist could respond to this worry—I will take that up in the next chapter—but what I have said hopefully suffices to show that Matravers raises a warranted phenomenological worry about the cognitivist view of musical expression.

I want now to turn to Matravers's own thesis. I specifically consider how by his theory he responds to one of the main objections to the arousal theory. The objection goes as follows. A work of music is an object of aesthetic appreciation. A listener needs to attend carefully enough to a work in order to appreciate its different properties. Expression seems to require such aesthetic attention. If expression is music's capacity to arouse emotions in us, it would be disconnected from aesthetic appreciation. Anyone could just be aroused by expressive music no matter whether she listens to it with or without enough attention. In that case, music would become like a "feeling drug"—say, "happy pills"—that can arouse some feelings or sensations in its consumer.

Matravers tries to meet this objection through two arguments: (1) that there is a phenomenological difference between feeling an emotion by, say, consuming a feeling drug and feeling an emotion by listening to music; and (2) that Matravers's suggested dispositional account captures the observation that experiencing musical expression requires enough attention to music. I think Matravers makes a plausible point in (1) but is unsuccessful in establishing (2).

Regarding (1), he says:

> There is an enormous phenomenological difference between the experience of expressive music and the experience of having one's mood altered by a drug; a difference which is more than sufficient to account for the fact that the first causes the belief that the music is expressive and the second does not cause the belief that the drug is expressive. (1998, 118)

Matravers is right. In the feeling drug case, the consumer is aware that her feeling because of the drug is due to her mere consumption of it and allowing that it chemically affects her body/brain. She is also aware that the effect is independent of her conscious consumption of the drug (the drug would affect her body even if she were asleep or hypnotized). But in the musical case, the listener seems to be aware that her hearing musical emotions is in one way or another dependent on her *conscious* act of listening—or her conscious engagement with music. (She seems to be aware that her emotional engagement with music would not be the same if she were asleep or hypnotized, for example.)

I think thus, with Matravers, that if we consider the above phenomenological difference, an arousal view of musical expression would not necessarily turn into the idea that being emotionally aroused by music is like being aroused by, say, a feeling drug. An arousal view can stand if it could incorporate this phenomenological difference into its explanatory mechanism. And this is what Matravers believes his account can do, which gets us to (2).

He says:

> I have concluded that the properties which distinguish being caused to be sad (by the 'feeling drug', for example) from the experience of the musical expression of emotion are that for music to arouse the feelings, the listener needs to pay sustained attention to its properties, in particular its relational properties. (1998, 185)

And by attention to music's "relational properties" he means: "the listener must have some grasp of something analogous to narrative in fiction: he must have some idea of where the music has 'come from' and where it is 'going'. This requires listening for and recognizing themes, being aware of the structure of the piece, and so on." (ibid. 178)

The problem I think this account has is that it cannot explain how exactly a "sustained attention" to musical properties is intertwined with experiencing a feeling as a result of that attention. To my view, as I mentioned above, an arousal theory can be saved from this objection if it can explain how being aroused by music is the result of the way in which a listener attends to music. Without such an explanation, we would be left with the mystery we were initially concerned with: what is experiencing musical expression? If the answer, like Matravers's, is that the experience is the experience of a feeling due to paying attention to music in some particular ways, we still need an explanation what is specific about that kind of attention to some sounds that results in, or is connected with, feeling any emotion.

Notice that by saying that an arousal theory like Matravers's needs to explain the matter, I do not mean it has to give a *causal* explanation. I agree with Matravers (1998, 122) that explaining the *experience* of expression is not a concern with what in music, or outside music, *causes* a listener to have that experience. (This is what I tried to show with my discussion of the causal-phenomenological distinction as regards musical expression in the previous chapter.) My concern is specifically about the phenomenology of that experience, or the phenomenological explanation of it. I think Matravers is right, as discussed, that there is a phenomenological difference between being aroused by music and being aroused by a feeling drug; and that phenomenological difference is the

existence of an attention component in the musical case. What I am concerned with, though, is that Matravers's account not only does not sufficiently explore the attention component of the experience, but, more importantly, does not explain how that component fits with the feeling component of the experience. The account leaves us with a dichotomous understanding of the experience of musical expression. On the one hand, we are told that that experience is the experience of a feeling, a feeling that we have in different emotional experiences in the real life. And on the other hand, we are directed to see that there is a "sustained attention" playing an essential role in our hearing musical emotions and having the intended feeling. All of these, without any explanation of *how* the feeling and attention components are intertwined.

The problem I have raised is in the spirit of what Budd (1992, 125) has called the "heresy of the separable experience":

> Hence what is needed from a theory of musical expression is a less external, a more intimate, connection between the experience of what music expresses and the experience of the music itself: it is necessary in some way to fuse the experience of the mood, feeling or emotion expressed by a musical work with the experience of the music which gives it expression, or to integrate the experience of what music expresses with the experience of the music. It is necessary to avoid the heresy of the separable experience.

In one sense, what I have tried to show is that Matravers's account is not developed enough to "avoid the heresy." That is to say, although the problem is highlighted in his arousal theory, the theory cannot explain what is specific in the way we attend to expressive music that causes us to experience feelings. This is why I think that this interesting version of the arousal theory is an unsuccessful explanation of what the *experience* of musical expression *is*.

In this chapter, I examined three ideas central to some theories of musical expression that try to shed light on what it is like to hear a piece or passage of music as expressive. Appealing to imagination, metaphor, and feeling are three ways in which these theories have tried to explain the experience of musical expression. A merit of these theories is that they try to characterize the intended experience independently of the listener's awareness of what in music *causes* expression. (I argued why such an approach is implausible in the previous chapter.) However, all these theories suffer from a common problem: they dismiss the fact that as a conscious experience, the experience of musical expression depends on how the listener *attends to* music. My analyses in this chapter show that the discussed theories cannot clarify what sort of *aesthetic involvement* with music causes the listener to imaginatively (in the case of imaginative theories), metaphorically (in the case of metaphorical theories), or affectively (in the case of dispositional theories) respond to musical expression. It is not that the listener should be aware of what in music causes expression, but *how one listens to music* is a part of one's experience of musical expression. The theories examined in the chapter are not able to factor this matter into their considerations. In the next chapter, I aim to argue for the importance of this matter about experiencing musical expression and show how a form of dispositional theory can explain that. Specifically, I will argue how we can connect the experience of musical expression to aesthetic attention to music—to show how the two are intertwined.

Bibliography

Budd, M. (1992). *Music and the Emotions: The Philosophical Theories*. Routledge.

Boghossian, P. (2002). On Hearing the Music in the Sound: Scruton on Musical Expression. *The Journal of Aesthetics and Art Criticism*, *60*(1), 49–55.

Davies, S. (1994). *Musical Meaning and Expression*. Cornell University Press.

Davies, S. (2011). Artistic Expression and the Hard Case of Pure Music. In *Musical Understandings and Other Essays on the Philosophy of Music* (pp. 7–20). Oxford University Press.

Larson, S. (2012). *Musical Forces: Motion, Metaphor, and Meaning in Music*. Indiana University Press.

Levinson, J. (2000). [Review of *The Aesthetics of Music*, by R. Scruton]. *The Philosophical Review*, *109*(4), 608–614. https://doi.org/10.2307/2693633

Levinson, J. (2006). Musical Expressiveness as Hearability-as-Expression. In *Contemplating Art: Essays in Aesthetics* (pp. 91–108). Oxford University Press.

Levinson, J. (2007). *Music in the Moment*. Cornell University Press.

Levinson, J. (2011). Hope in The Hebrides. In *Music, Art, and Metaphysics: Essays in Philosophical Aesthetics* (pp. 336–375). Oxford University Press.

Karl, G., & Robinson, J. (2015). Yet Again, "Between Absolute and Programme Music." *British Journal of Aesthetics*, *55*(1), 19–37.

Kivy, P. (2001). The Arousal Theory of Musical Expression: Rethinking the Unthinkable. In *New Essays on Musical Understanding* (pp. 119–151). Clarendon Press, Oxford.

Walton, K. (1999). Projectivism, Empathy, and Musical Tension. *Philosophical Topics*, *26*(1 & 2), 407–440.

Walton, K. L. (1988). What Is Abstract about the Art of Music? *The Journal of Aesthetics and Art Criticism*, *46*(3), 351–364.

Walton, K. L. (1990). *Memesis as Make-Believe*. Harvard University Press.

Matravers, D. (1998). *Art and Emotion*. Oxford University Press.

Peltola, H.-R. (2017). Sharing experienced sadness: Negotiating meanings of self-defined sad music within a group interview session. *Psychology of Music*, *45*(1), 82–98.

Pitcher, G. (1965). Emotion. *Mind*, *74*(295), 326–346.

Scruton, R. (1999). *The Aesthetics of Music*. Oxford University Press.

Scruton, R. (1974). *Art and Imagination: A Study in the Philosophy of Mind*. Methuen & Co Ltd.

Tanner, M., & Budd, M. (1985). Understanding Music. *Proceedings of the Aristotelian Society*, *59*, 215–248.

Chapter 3: Hearing Musical Expression by "Moving With" Music

1

A longstanding opposition regarding musical expression is between cognitive and affective or arousal theories.[50] On one side, the cognitivist believes that expression is a *property* of music, and appreciation of musical expression is the *recognition* of such a property in a musical work. On the other side, the affectivist thinks musical expression results from the listener's affective response to music; that is, the listener hears music to be expressive when she *feels* or imaginatively feels an emotion.

A central challenge cognitivism should grapple with is to spell out the kind of property it appeals to in explaining musical expression and, relatedly, how the listener perceptually recognizes that property. Two main versions of the view by Stephen Davies (1994) and Peter Kivy (1981) suggest that the intended property is *resemblance*; that is, some formal properties of music can perceptually *resemble* the appearance of human expressive behavior (like the resemblance of a fast rhythmic pattern with a jaunty, cheerful behavior). So, the property appealed to is a relational, complex property: the resemblance of some formal, musical properties to human expressive behavior. But it is not clear how we should understand this idea that in hearing expressive music, we *hear* such resemblances. As Jerrold Levinson has pointed out, music can resemble many things and phenomena in the world; why should we auditorily

[50] In this paper, I will use "affective view" or "affectivism" as a broad phrase to include all the affective theories, whether they consider the listener's affective response as an actual feeling (Matravers, 1998) or include imaginative feelings in such affective responses as well (Walton, 1999 and 1988). Moreover, I will use the term "expression" specifically in the sense of "expressiveness" following how the terms have been distinguished, for example, in Kivy (1981)—the distinction being that "expression" is for psychological beings and "expressiveness" is for artistic work.

compare it to human expressive behavior?[51] Moreover, as Derek Matravers has argued, there does not seem to be a specific moment at which the listener perceptually recognizes musical expression as a property; hearing expression temporally unfolds during the process of listening.[52] A main worry behind such criticisms is that the cognitivist owes us an explanation of how the listener auditorily recognizes expression as a resemblance property in music.

On the other side of the debate, the main challenge the affectivist faces is to explain how *music* is responsible for expression. After all, some affective responses to music could be irrelevant to musical expression. A piece could make me sad without being expressive of sadness because the reason for my sadness could be some personal memories involving the piece and not any of its aesthetic qualities.[53] If the affectivist's reply is that only those affective responses to music are expressively relevant that are *aesthetically appropriate*, she needs to spell out what *kind* of affective responses to music are aesthetically appropriate and can thus lead to experiencing musical expression.

In this paper, I propose how the affective view can deal with the challenge described above. Setting aside the resemblance notion, which can, as mentioned, be problematic for the cognitivist view, a strength of the view is that its strategy is to pinpoint what *musical mechanism* causes expression. In section 2, I argue that to give a full explanation for musical expression, a theory cannot avoid the strategy of identifying a mechanism, especially if we agree that experiencing musical expression is the result of the listener's *aesthetic attention to music*. In section 3, I will suggest how the affective view can take such a strategy by appealing to musical

[51] Levinson (1996), p. 97
[52] Matravers (1998), pp. 120-1
[53] This is one of the main criticisms Kivy (1981) levels against the view.

movement. I will explicate and develop my idea in section 4 and will then consider and respond to two objections in the last section.

2

I want to start with a passage by Malcolm Budd (1992) in which he sets a plausible requirement for an affective theory:

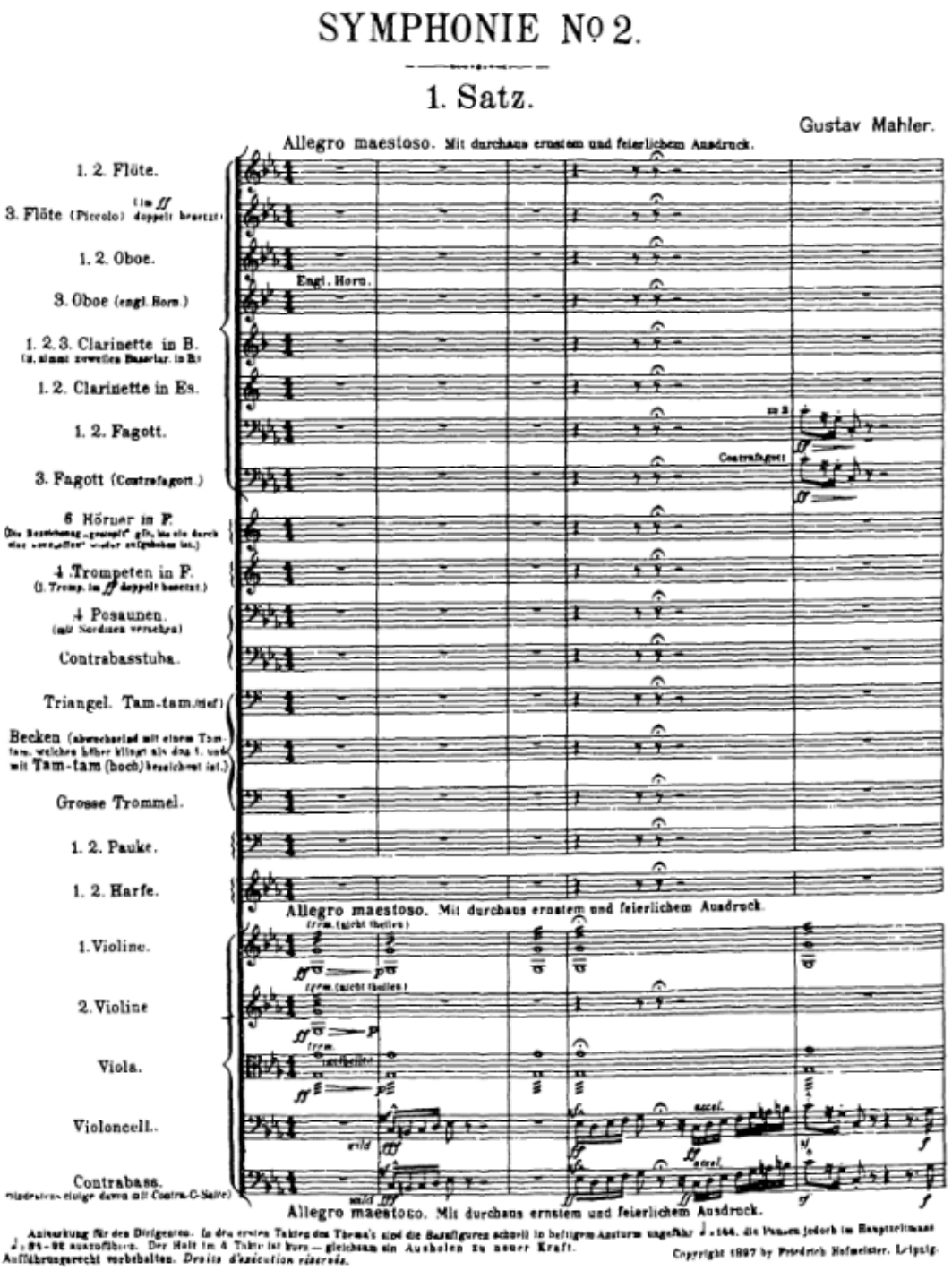
> [W]hat is needed from a theory of musical expression is a less external, a more intimate, connection between the experience of what music expresses and the experience of the music itself: it is necessary in some way to fuse the experience of the mood, feeling or emotion expressed by a musical work with the experience of the music which gives it expression, or to integrate the experience of what music expresses with the experience of the music. It is necessary to avoid the heresy of the separable experience. (125)

According to Budd, a theory claiming that the listener feels what music expresses ("mood, feeling, or emotion") is required not to separate that affective experience from the experience of the music itself. If, for example, a piece or passage is expressive of sadness, the listener's feeling of sadness should be intimately connected with her experience of the music itself. Setting aside the ontological matters regarding the notion of "music itself," I believe it is safe to translate "the experience of the music itself" to the experience of musical properties—after all, whatever ontological view we take about a musical work, we cannot deny that there are different musical properties that we hear while listening to music. Thus, Budd's requirement, in my understanding, is that an affective theory must be able to show how the affective experience of the listener is fused with the experience of musical properties, the properties that are expressively relevant.

I believe Budd's requirement is plausible because it is hard to deny that appreciating musical expression is a function of *how* one listens to music. If one could not appreciate those

properties of music that are expressively relevant, she would be unable to appreciate the related expressive qualities of the music. For example, for a person mostly accustomed to monophonic music, it is hard to appreciate the expressive qualities in polyphonic music (like in a symphony). Or, as another example, the Western ear is mostly accustomed to two main modes of major and minor. But in some cultures, there are several modes, and more than one may be used in one musical piece or performance. Some of the expressive qualities of such kinds of music come from various modal changes. Now, for a person whose ear is accustomed mostly to Western music, it would be hard to hear various modulations in, say, Iranian Dastgahi music, which has seven modes, each called a *Dastgah*. As a result, she would be unable to appreciate some of the expressive qualities of Dastgahi music. Or consider the beginning moments of Gustav Mahler's Resurrection Symphony (Fig. 1). The tense quality the passage creates in the beginning measures comes in part from the violin and viola sections playing, in the background, tone G shakily (using a technique called tremolo) and softly while the bass and cello sections introduce and develop the beginning theme. A full appreciation of the expressive quality of this passage is also dependent on the listener's attention to these background sounds while hearing the beginning theme. Such a dependence is what I mean by saying that the appreciation of musical expression is a function of how one listens to music.

Fig. 1. Gustav Mahler, Symphony No. 2, "Resurrection," mm. 1-5



The idea is not that appreciating expression is necessarily a function of one's musical and technical knowledge, although it can sometimes be. I mostly used the term "accustomed to" above to convey the idea that one's manner of listening to music basically forms culturally in the process of her listening to different pieces of music in her musical culture from childhood. I am concerned with this aspect of musical listening. I do not want to make the strong claim that to appreciate musical expression one must necessarily have musical knowledge. In most cases of expressive appreciation, the listener's cultural familiarity with the music in question, in the sense of her knowing how to listen to it, would be enough.[54] Whether the listener is *aware* of her manner of listening or not is a different matter and irrelevant to my argument. Regardless of such an awareness, the point is that the manner of listening is intertwined with the listener's appreciation of the expressive qualities of a piece.

Now, if this is the case, an affective theory claiming that expressive music arouses feelings (or imaginative feelings) in the listener would be incomplete if it does not meet the requirement I delineated above based on Budd's view. Such a theory should be able to spell out the aroused feeling in terms of how the listener attends to music (musical properties) or in terms of, using Budd's terms, "the experience of the music itself."[55] Without giving such an explanation, not only does an affective theory remain incomplete, but it also becomes musically uninteresting because it leaves the musical mystery of expression unexplained, the mystery being how *experiencing music* causes one to entertain (think of, feel, imaginatively feel, etc.) emotions.

[54] Using the epistemological distinction between knowledge-that and knowledge-how, what I mean here is very close to knowledge-how and not necessarily knowledge-that.
[55] I do no mean that a plausible arousal account has to necessarily give a causal explanation of musical expression. (Regarding the casual explanation see the related discussion in chapter 1). However, I think that an arousal account should be able to explain how the arousal of emotions by music depends on an aesthetically appropriate manner of listening to music.

The only affective theory that, to my knowledge, has directly addressed the above matter is Matravers's. The core idea in Matravers's theory is that expressive music causes a listener to feel an emotion and that feeling causes her to believe that music is expressive. I want here to set aside the detail of his theory because my focus is on whether and how his theory can meet the requirement I have explicated in this section. That is, I am concerned with how Matravers thinks experiencing a feeling (in response to expressive music) is intertwined with the experience of *music* (the music itself). Regarding this, Matravers says:

> I have concluded that the properties which distinguish being caused to be sad (by the 'feeling drug,' for example) from the experience of the musical expression of emotion are that for music to arouse the feelings, the listener needs to pay sustained attention to its properties, in particular its relational properties.[56]

And by attention to music's "relational properties" he means:

> [T]he listener must have some grasp of something analogous to narrative in fiction: he must have some idea of where the music has 'come from' and where it is 'going'. This requires listening for and recognizing themes, being aware of the structure of the piece, and so on.[57]

I think Matravers clearly acknowledges the issue and tries to address it partially, but what he says is not explanatory enough. He plausibly points out that affective response to expressive music is aesthetic in the sense that it depends on the listener's attention to the aesthetic properties of music, and such aesthetic, affective responses are different from non-aesthetic affective responses (to 'feeling drugs,' using his example). But what he is unable to do is to *explain why* paying "sustained attention to" aesthetic properties of music causes the listener to experience a feeling. Without such an explanation, the theory cannot help us understand how the realm of music (relational musical properties) is connected with the realm of human emotions. That is, as

[56] Matravers (1998), p. 185
[57] Ibid., p. 178

an affective theory, it cannot show how the listener's experience of a feeling is fused with her experience of music—although it acknowledges that the two experiences are fused. Now, I want to suggest how we can explain this in the context of an affective theory.

3

I will argue that the listener hears musical expression when she *moves with* music. I will explain the central thesis in this section and delineate in more detail how "moving with music" is at the core of the appreciation of musical movement in the next section.

It is a well-recognized phenomenon that, on many occasions, we hear movement when hearing musical sounds in relation to each other. We hear notes going upward or downward in a musical scale. We hear "ups and downs" in a melody. We hear harmonic resolutions as a kind of movement tendency—a dominant chord tending to go toward the tonic. It may be better to say that we hear *movement relations* or *movement tendencies* between ingredients of music (notes, chords, etc.). As Jackendoff and Lerdahl have articulated:

> [P]itch space fosters intuitions of tonal attraction […]. An unstable pitch tends to anchor on a proximate, more stable, and immediately subsequent pitch […]. Tonal attractions in turn generate expectations. The listener expects a pitch or chord to move to its greatest attractor.[58]

Delving into this psychoacoustic phenomenon itself is beyond the scope of the current paper. Nevertheless, I want to show how it grounds the perception of musical movement, a perception that has phenomenological characteristics relevant to hearing musical expression. But

[58] Jackendoff and Lerdahl (2006), pp. 51-2. The idea of movement relations stated in this passage is what I am primarily concerned with here. The idea of *generating expectations* mentioned by Jackendoff and Lerdahl is the basis of Stephen Mayer (1956)'s theory of musical expression.

before getting to the core of the problem, some clarification on the notion of intended musical movement.

The musical movement that will be our primary concern should be distinguished from a few other kinds of movement related to musical sounds and performance. Arnie Cox has described three forms of music-related movement: "acoustic motion, psychoacoustic motion, and the apparent motion of music from one physical sound source to another."[59] Acoustic motion is the movement of sound waves in the air. "Psychoacoustic motion," in Arnie's terms, "includes the apparent motion of sounds within the three-dimensional sonic space of stereophonic recordings (as in panning between left and right channels)." There is also a perception of movement when we hear different parts of a melody moving from one section in an ensemble or orchestra (for example, string instruments) to another (say, woodwind instruments). Moreover, a fourth form of movement is involved in performing music: a performer's bodily movements in a musical performance. However, the kind of musical movement we will delve into is distinct from the four kinds described above: we will be concerned with our perception of a *flow* in music, for example, how a melody *goes* from one point to another and rests at the end. I will consider what philosophical challenges this kind of movement perception in music presents and how they are related to our hearing of musical expression. I start with Scruton's theory of musical movement.

Scruton (1999) thinks we metaphorically attribute movement to music. He believes that although musical time is a real property (based on the duration of sounds), musical space is metaphorical; that is, we hear musical sounds as if there is a musical space in which they move toward each other. And regarding metaphor, he thinks experiencing something metaphorically is based on someone's imaginative engagement with the object of her experience. The core of his

[59] Cox (2016), P. 110

theory of musical movement is that in listening to music, we imagine a space in which musical sounds move toward each other.

Scruton's main reason for appealing to imagination to explain musical movement is that in music nothing literally moves.[60] When we hear the leading note tending to move toward the tonic in a major scale (for example, B to C in C major), nothing literally moves; the leading note and the tonic are two distinct events or objects. So, it seems that in hearing such a movement tendency, our mind "adds" something to what is literally heard. To Scruton, we *imagine* a (musical) space in which one note is heard as moving toward the other.[61] I believe that regardless of whether Scruton's specific claim is true or not, his analysis shows something important concerning how we perceive musical movement; that is, in hearing musical movement, the mind needs to *form* the related experience in a particular way. That is to say, the core idea with which I agree is that perceiving musical movement requires the (mental) *participation* of the listener in one way or another. I will argue that imagination cannot capture the nature of such participation, but before that, let me put forward two phenomenological observations about the nature of that participation itself.

The first phenomenological observation is as follows. Imagine a case in which you see an object move; for example, you see a ball thrown by a person, and you see the ball moving down to the ground. In such a case, you see both the *object* (the ball) moving and its *movement* (the ball's going down). In all typical cases of the perception of movement, these two components are present: *what* moves (the thing-component) and the *movement* of that thing (the movement-

[60] By "movement in music," it is meant movement in a musical *work*—whatever a work ontologically is—and not movements involved in performing music. There are indeed *bodily* movements in any performance, the bodily movements of the performer, which are not meant in the context of our discussions.
[61] Another form of the imaginative theory of musical movement has been proposed by Kania (2015).

component). This is also the case when we perceive movements in mathematical graphs, for instance, a sine wave represented in a sine graph. Even here, we see *something* whose *movement* is represented through the graph. For example, in a graph representing a sine wave, *something* has a movement that has a sine pattern. However, in the case of hearing musical movement, one component of movement perception is missing: the thing-component. Although we hear movement tendencies among musical elements (say, two notes or chords), we do not hear a particular thing in music that moves. Using the above example, we hear a movement tendency between the leading note and the tonic as if there is a motion, a flow, between them, but, at the same time, we hear them as two distinct things without hearing that either of them moves and changes its position; we hear them as two different entities the first of which ends at a point of time and the other starts immediately or shortly later. (We do not hear the tonic as the leading note whose position has changed!)

Another phenomenological observation with regard to musical movement—secondary to the one described above—is this. If we want to go beyond hearing movement between two notes (like in our above example) and consider hearing movement in, say, a melody, we need to consider that in hearing musical movement, the mind must reconstruct the movement in the process of listening. In an instant of listening to music, what was heard moments ago and the expectations for what will come shape the experience.[62] For the listener to perceive the movement of a quite long melody, she needs active listening; that is, she needs to attend to the parts of the melody (sounds, phrases, …) and how they are related. Appreciating movement in more complicated structures of music requires higher levels of active listening.

[62] See Montague (2011).

The active role of the mind in perceiving musical movement described in my second observation is not something peculiar to music. Even in visually perceiving complicated movements in, for example, a sports game, our mind actively processes its objects of perception and their relations to each other that unfold in longer spans of time. I mentioned the second observation only as a complement to the first one, namely that in hearing musical movement the thing-component of movement perception seems to be missing; we do not hear something that moves. In addition (and this was the second observation), in hearing musical movement, the mind must actively process different bits of the experience (relating the current moment of listening to what was listened to before and what is expected to come). My following argument centers on the first observation, which is, as argued, peculiar to hearing *musical* movement.

I suggest that the missing component of the perception of musical movement (the thing-component) is filled with the listener herself. Our perception of moving objects in the world is, in fact, the perception of a process of change. When I see a leaf fall from a tree, I perceive a continuous change of its position from the branch to the ground. A process of change, in other words, is at the core of any movement, and the perception of movement is, at its core, the perception of a process of change. What I am going to suggest and defend about the perception of musical movement is that it forms when, in the absence of anything that moves in music, the listener feels (or imaginatively feels) a process of change in her own body. If the discrete elements of music cannot themselves move or present a continuous process of change happening to one thing (in music), the listener herself feels or imaginatively feels a continuous process of visceral change in herself. Outside the context of music listening, when we make different bodily movements we feel the movement of our own body, we are proprioceptively aware of a process of change in our body. When our heart rate increases, we feel that visceral change: we are (or can

be) aware of a continuous process of an inward change. This is the kind of feeling that I suggest is involved in the perception of musical movement. And having this feeling is how our mind provides what is missing in the music regarding musical movement—the thing-component, or what moves. When this happens, we "move with" music.

However, why should we not say, as suggested by Scruton, that we *imagine* the thing-component missing in perceiving musical movement? The reason I think imagination is not a plausible solution to the problem is related to the intentionality of imagination (or the intentionality of the sort of imagination appealed to by Scruton).[63] When I imagine, I imagine *something*, or in imagining my mind is directed toward the object of my imagination. Even if I might not be aware, on an occasion, that I am imagining, it appears I am aware of my object(s) of imagination. That is, it might be that sometimes I am unaware of the *act* of imagination while imagining, but I am always aware, in one way or another, of the *object* of my imagination as something my mind is directed toward.[64] For example, if we wanted to consider dreaming as a form of imagination (as Walton (1990) has proposed), it might be that in a dream, I am unaware that I am dreaming (imagining), but it seems even in such a dream my mind is active or conscious to the extent that it engages with the objects being dreamt of. (I think this is one reason why we can, on many occasions, remember the contents of a dream, at least to some extent, after we wake up.)[65]

---

[63] I am not concerned here with Kantian imagination which has also been called *productive imagination* (Matherne, 2016). This form does not seem to be relevant to the imaginative theory, by Scruton, I am discussing here—or at least it is not explicitly stated by him that he is suggesting the specific form of imagination primarily used by Kant in the context of his transcendental philosophy.

[64] I reckon that even in the kind of Kantian imagination called "productive imagination," (Matherne, 2016) the mind is directed toward the object of imagination although the imaginer is unaware of the act of imagination.

[65] What John Locke says about dreaming seems to me to be related to the intentionality of that mental state: "And *Dreaming* it self, is the having of *Ideas*, (whilst the outward Senses are stopp'd, so that they receive not outward Objects with their usual quickness,) in the mind, not suggested by any external Objects, or known occasion; nor under any Choice or Conduct of the Understanding at all […]." (*An Essay Concerning Human Understanding*, Book

Now, if the listener imagines the thing-component of musical movement, what does she imagine exactly? In other words, if in perceiving musical movement, the listener imagines *something* moving, what is the thing imagined? Scruton believes the listener imagines a musical space, but there are two issues with this suggestion. First, the intended imagination seems to have two components: a musical space *and* a thing that moves in that space. So, saying that we merely imagine a musical space is not enough to explain musical movement: we also have to add that we imagine a moving thing in that imaginary space. But even if we suppose that the notion of musical space is unproblematic and set aside the question of what such a space is, should we say that what moves in musical space is something *over and above* any musical ingredients (like notes and chords) or those ingredients *themselves*?[66] The latter seems to be against the phenomenology of music listening. We hear musical notes, chords, etc., as discrete, temporal entities—things that begin and end—and not as things that are auditorily present at all points during listening to a piece. So, it does not seem that we imagine any of the musical ingredients in a piece, like a note, to move during that piece. But if the former alternative is the case—that is, we imagine something *over and above* musical ingredients that moves in music—then the main question is about the nature of that imagination: do we *auditorily* imagine the moving thing, or is there some form of *visual* imagination involved too? It seems to me that saying that we imagine *hearing* something that moves in music (or in musical space) just shifts the problem from hearing to imaginative hearing, and we still need an explanation of how we can *imaginatively hear* something that moves. If the idea is that visual imagery is involved here, it is unpersuasive, as it does not seem that whenever I perceive a movement, say, in a line of melody, I need to

---

II, Chap. XIX, Section 1) If in dreams we have such Lockean ideas, this means in having dreams our mind is directed toward, engages with, and, on many occasions, is able to recollect them afterwards.

[66] In my understanding, Kania (2015) suggests that we imagine something over and above musical ingredients, "'something I know not what'" (ibid., p. 168).

engage visual imagination. In fact, it seems that our perception of musical movement is an auditory and not a visual phenomenon.

So, appealing to imagination to explain the perception of musical movement raises several issues that need to be examined and dealt with. What I presented above is not indeed a decisive rebuttal of an imaginative theory of musical movement, but I think it shows that unless we are given informative suggestions about the nature and content of the intended imagination, our solution remains unexplanatory. Appealing to imagination cannot solve the problem automatically, considering that imagination is a very broad mental state with various elusive characteristics, and its forms in sensory modalities other than vision are much less explored than visual imagination. In what follows, I want to show how appealing to *feeling*, as adumbrated above, can more convincingly help us in explaining the perception of musical movement, particularly with regard to musical expression, in such a way that it will also meet Budd's requirement.

4

The core idea I described in the previous section was that the thing-component of musical movement—*what* moves when we hear musical movement—is provided by the mind itself through the listener's feeling of an inner movement. The idea is that the listener hears relations between discrete musical ingredients, say, between different notes in a melody, and those relations suggest movement tendencies in the music. Based on this sensory input (manifold) the listener then feels the process of change and the flow suggested by those relational qualities of music. It is in this sense that the mind subconsciously fills the gap—between the *discreteness* of musical events and the *continuity* suggested by them—by forming an inward process of change

and continuity felt by the listener. To put it roughly, the discreteness is in music, and the continuity in a form of a process of change is felt by (or in) the listener. However, these two realms are tightly intermingled because the relational qualities among the discrete events of music (musical ingredients) suggest a continuity, flow, and unity that calls for the active participation of the listener. And that active participation has a feeling component, as I have described above. I aim to flesh out this active participation and its feeling component in what follows.[67]

What we feel in experiencing musical movement is the same as feeling our own bodily movements. For example, walking slowly has a different feeling than jumping energetically. Such feelings are particularly more salient in different forms of our expressive behavior. For example, when we laugh and our face muscles move, or when we are sad and express it with a slow gait and by dropping our shoulders, each is felt differently. Jenefer Robinson (2013) has insightfully pointed out such feelings when we watch and aesthetically appreciate a dance performance and imaginatively feel the expressive movements of a dancer:

> The audience at a work of dance mimics in imagination the bodily movements of the dancers on the stage. Because of this, they can actually feel what the dancers are expressing. Of course the audience is not actually dancing about in the aisles; their own bodily activity is largely suppressed. But their muscles may tense, their blood may race, and they may feel as if they are extending their arms, flexing their legs, and so on even as they are sitting quietly in the theater. Moreover, because dance evokes bodily responses in the audience, it is peculiarly capable of communicating what it feels like to be proud and resentful or in the throes of young love or stricken with grief or rage.

[67] My focus will be on *tonal* music, any piece of music whose constituent notes are organized around a central tone. Modal music can also have this basic characteristic. For example, in Persian classical music, each improvisation or piece based a mode (*dastgah*) is centered on a specific tone (see Khaleghi (1937/2023) and Farhat (1990)). My upcoming arguments will not, however, be focused on aleatory music. Although I do not think we listen to aleatory music without hearing different tendencies between tones, I do not intend to consider the implications of my theory for aleatory music in the scope of this paper.

There are two key ideas in this passage: (1) feeling bodily movements that relate to expressing emotions can cause one to feel emotions, and (2) we have the capacity to *imagine* or mimic such movements, particularly when we respond to the expressive behavior of others, for example when appreciating an expressive dance performance. The former is the core observation in the James-Lang theory of emotions according to which when we are in emotional states, we feel physiological and behavioral changes (such as retreating and an increase in our heart rate when we are in danger and fear). And based on that theory, such feelings are the essential components of emotional experiences. Jesse Prinz (2005), a defender of this view, has argued that emotions are feelings, specifically when they are experienced consciously. There is also empirical evidence suggesting that different emotions are associated with distinct feelings—fear has a different feeling than sadness, sadness than frustration, and so on.[68] Regarding the second idea stated above, Robinson believes observing the expressive bodily movements of a dancer can cause the observer to imagine feeling those movements from inside. Such imagining is of the same sort as what Dominic Gregory (2010) calls *sensory* imagination, particularly "those imaginings where one imagines a sensation 'from the inside,' as for example when imagining what it feels like to have a sore throat […]."[69] Similar to imagining the feeling of a sore throat, we are able to sensorily imagine, or *mimic*, in Robinson's terms, what it feels like to have the bodily movements of a dancer. So, Robinson's main idea, in my understanding of her, is that aesthetically paying attention to the movements of a dance "evokes bodily responses" in the

[68] Damasio et al. (2000)
[69] p. 737

audience and the feeling associated with this bodily response is the core of the audience's appreciating the expressive quality of the performance.[70]

What I suggest is that appreciating expressive music involves such bodily responses. When I say the listener feels an inner process of change in response to the relational properties among musical ingredients that suggest continuity and flow, I mean the kind of bodily responses Robinson describes in appreciating dance. However, I believe the mechanism through which listening to expressive music evokes such bodily responses is different from the mechanism through which observing expressive dance does so. We do not, figuratively speaking, "observe" or literally hear an expressive movement in music and mimic it in ourselves. But before delineating the musical mechanism involved, I want to point out the pervasiveness and the appreciative characteristic of our bodily responses to music.

A tendency to make different bodily movements while listening to music, particularly in response to the music we enjoy and are familiar with, is a phenomenon known to many of us. Sometimes that tendency presents itself in dance movements involving different parts of our body, and sometimes only in more limited forms such as tapping or hand movements. One interesting thing about such reactions is that they do not seem to be merely automatic bodily reactions to music; they seem to be importantly related to music appreciation and aesthetic enjoyment.[71] On many occasions, a person who tends to make different and nuanced movements

[70] It seems that for Robinson such bodily responses are an amalgamation of actual and imaginative feelings in the audience: "[…] their muscles may tense, their blood may race, and they may feel as if they are extending their arms, flexing their legs, and so on even as they are sitting quietly in the theater."

[71] In discussing Scruton's idea about the connections between dance and music, Levinson says: "[A]part from the impulse to actual dance which music heard with understanding may inspire, musical understanding is itself essentially a kind of latent dancing, whereby one moves with musical movement, but on a psychic or imaginary plane." (2009, p. 424) Richard Shusterman (1991) has insightfully argued that somatic response to some forms of popular art is a part of their appreciation and getting enjoyment from them. Particularly about rock music he says: "Rock songs are typically enjoyed through moving, dancing, and singing along with the music." (p. 205) Although he mainly points out this feature of appreciating rock as a characteristic distinct from appreciating "highbrow"

while listening to a piece seems to be *getting* (and enjoying) what is going on in the piece; she seems to be appreciating the music in one way or another. When one does not understand some music at all, it would be hard for them to make bodily movements in response.[72] Imagine when you listen to an unfamiliar piece from a culture from which you have not heard any type of music before. Probably, figuring out how to listen to such a piece would be hard for you, let alone making movements to it while it is progressing.

The above observation can reveal an important characteristic of hearing and appreciating musical movement: the flow of music seems to evoke a movement response in the listener. The idea is not that perceiving musical movement necessarily represents itself in actual bodily responses. Even when appreciating musical movement is not bodily represented, a form of movement response to music seems to be latent in the music experience. This is the core of Charles Nussbaum's (2010) theory of musical movement and expression, and since my view is very akin to the general framework of his theory, discussing that theory will help me develop my view.

Nussbaum believes that musical tonal structures—particularly in Western, tonal music from the seventeenth century—prompt the attentive listener to imaginatively move in an imaginative musical terrain:

> The internal representations employed in recovering the musical structure from the musical surface specify motor hierarchies and action plans, which, in turn, *put the listener's body into off-line motor states that specify virtual movements through a virtual terrain or a scenario possessing certain features*.[73]

---

classical music, I think there is not a fundamental difference here. Even though somatic response is not a part of the typical *ritual* of listening to classical music in concerts, this does not rule out that giving aesthetically relevant somatic response to classical music can represent the listener's enjoyment and appreciation of that music.

[72] See Anderson (2011) as an example of the importance of bodily movements to music from the perspective of music education.

[73] p. 47

A central idea in Nussbaum's theory is that an action plan is musically represented to a listener while listening to music (what he calls the *musical plan*). His theory's notion of the musical plan is grounded on two other theories, one related to music cognition and the other to representational contents. Jackendoff and Lerdahl (2006) have argued that there are hierarchical tonal and rhythmic/metrical organizations in works of tonal music. For example, the rhythmic structure of a phrase can be divided into subgroups having similar metrical structures to each other (for example, see a passage from the song *Norwegian Wood* by the English rock band the Beatles in Figure 2). A phrase can also be divided into subgroups based on its tonal structure, how it starts and ends, and how different notes from the relevant musical scale constitute a piece's melody and harmonic structure. The main point, in Jackendoff and Lerdahl's terms, is that "[t]he understanding of [a melody like Norwegian Wood] goes beyond just hearing the sequence of notes […]"[74]; "the stability of a note cannot be determined by its pitch alone. The same pitch may play a different role at different points in the melody, and its relation to adjacent notes—at each level of abstraction from small to large—is crucial."[75] That is, according to Jackendoff and Lerdahl, understanding a musical work of this kind depends on getting how the music's tonal and metrical ingredients are *related* to each other. (Obviously, the claim is not that to have such an understanding, the listener must have a knowledge of the related music theory; an acculturated listener is well equipped to appreciate such relational musical properties.) A musical plan is in part related to such hierarchical organizations in music. These organizations present the kind of musical architecture each piece is based on and shape the flow of music—two

[74] Jackendoff and Lerdahl (2006), p. 53
[75] Ibid., p. 54

different melodies with two different tonal organizations present distinct progressions and distinct “ups and downs.”

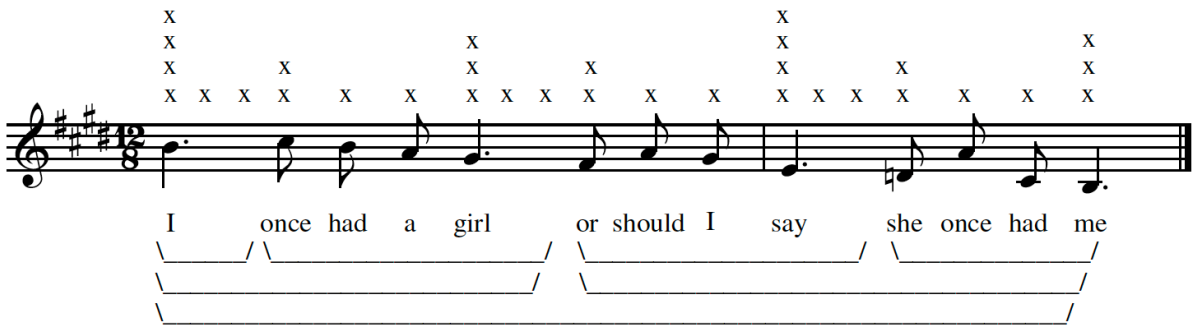


Figure 2. The metrical structure and grouping of a phrase in *Norwegian Wood* (from Jackendoff and Lerdahl (2006)

Nussbaum’s notion of musical plan is also grounded on a kind of representational content introduced by Ruth Garrett Millikan (1995) called “pushmi-pullyu representations.” According to Millikan, a pushmi-pullyu representation is simultaneously both descriptive and directive. Descriptive representations in language are like “Chicago is a city in the United States,” which represents how things are in a part of the world. On the other hand, a statement like “Run!” is a directive representation stating what needs to be done by a hearer of the statement. Millikan argues that besides these two representational forms, there are, using her examples, statements such as “No Johnny, we don’t eat peas with our fingers” or “The meeting is adjourned.” Such statements represent how things are *and*, at the same time, what kind of response the hearer is expected to give. Millikan believes the application of such representational content is not limited to the forms used in human language. She thinks some communicative forms among animals, like bee dances or a hen’s call to her chicks for food, are also examples of pushmi-pullyu representations. Some states of mind in humans, like intentions, are also such representations.

For example, to Millikan, an intention is a mental representation that is both a descriptive and a directive: it is descriptive as regards what is intended and directive as regards what needs to be done. An intention is then a pushmi-pullyu representation. It is important to note that pushmi-pullyu representations are not, holds Millikan, "the mere conjunction of a pure descriptive representation and a pure directive one but are more primitive than either."[76] By primitive, she means that entertaining a mental representation of this sort need not include any higher-order cognitive processes, such as inference. We do not necessarily analyze, say, an intention and decompose it into a belief (descriptive content) and a desire to act in this or that way (directive content): we entertain an intention's being simultaneously descriptive and directive at a more primitive level independently of such inferences and analyses.

Let us now see how the ideas of the two theories discussed above direct Nussbaum's theory of musical movement and expression. To Nussbaum, a musical plan is constituted by an organization of the sort suggested in Jackendoff and Lerdahl's theory and is also a pushmi-pullyu representation. The idea is that when the listener cognizes the hierarchical structures of tonal music (its metrical and tonal structures), that is, when she cognizes the musical plan, it causes internal pushmi-pullyu representations in the listener's mind. Because they are pushmi-pullyu representations, the listener is directed or evoked to act upon them (because like a hen's call to her chicks, they have a directive force and are not merely descriptive representations). And the listener's acting upon such representations is her moving (imaginatively) according to the music plan. However, like the process of grasping pushmi-pullyu representations, this process is not inferential: the process of music appreciation does not include the listener's first cognizing tonal and metrical structures of music and then reflectively imagining some sort of movement. The

[76] p. 186

cognitive process and imaginative response happen simultaneously, as far as I understand Nussbaum's theory.

This is a sophisticated and interesting theory for delineating the backbone of a mechanism through which music cognition may lead to a movement response to music, to move (imaginatively) while listening to music. The suggestion is that appreciating the relational qualities of music—presented to the listener via different hierarchal structures among musical pitches and rhythmic patterns—causes a representational mental state that is both about what is in music and how the listener should act upon what is musically suggested. Yet, there is an explanatory gap in Nussbaum's theory: why is the content of music a pushmi-pullyu representation? What is special about tonal and metrical structures of music whose cognition causes pushmi-pullyu representations rather than merely descriptive representations, representations of merely what is in music? Why can we not merely cognize the relational qualities of music and appreciate the different ways musical ingredients are related to each other without acting upon them? The problem with the theory is not its two main components isolated from each other, that musical relational qualities constitute the flow of music, and that appreciating such relational qualities has a participatory nature. I think both components are plausible observations about music appreciation on their own. What is missing here is how these two components are *related*: why does cognizing the relational qualities of music lead to a participatory response (a pushmi-pullyu state of mind)? I think Nussbaum's main response to this matter is speculations referring to the early stages of human evolution. The reason music presents pushmi-pullyu representations is, Nussbaum believes, the use of musical vocalizations by our ancestors for communicative purposes:

> "An intentional vocal imitation of the growl of a predator," [Merlin] Donald suggests […], would "communicate the specific nature of a threat and allow a specific response on the part of a tribal group." Not only, we might add, would such an imitation *allow* the response; it would *prescribe* it. *The mimetic vocal gesture would function, that is, as a pushmi-pullyu representation*.
>
> Conformity and coordination of group behavior are major accomplishments of mimetic culture. It is easy to imagine a tribal group of *Homo erectus* sitting around the communal fire, revisiting events of the day, or planning those for the next, while communicating by way of a mixture of gesture, pantomime, and songlike vocalizations. The quasi-spatial organization of musical pitches would enhance the gestural aspect of these vocalizations. […] Music displays a much higher degree of syntactic organization and is more complexly generative than any such mimetic symbolism. Still, there is good reason to regard the musical surface produced by a performer as a descendant of rudimentary song, and therefore as a mimetic utterance, an external pushmi-pullyu representation.[77]

Even if we accept the evolutionary speculation here concerning musical content being a pushmi-pullyu representation, there is a problem. It seems that our aesthetic responses to music now have a function weakly associated with the functions of music in those primitive stages of human evolution. Moreover, it seems a plausible assumption that our cognitive abilities for appreciating sound structures for aesthetic purposes have also evolved. If so, our music cognition seems to be evolved enough to separate, with a little reflection, what is communicated by music (its hierarchical structures) from its directive force in such a way that we seem to be able to appreciate the content of music (the musical plan) without acting upon it. As an analogy, consider how easily we can separate what is communicated in a pushmi-pullyu linguistic representation such as "No, Johnny, we don't eat peas with our fingers" from its directive force. That is, if Johnny is old enough and is a moderately able speaker of English, he can think of the content of this pushmi-pullyu representation without acting upon it. In other words, he can analyze it in a way that the chicks are unable to do in response to their mother's call for food

[77] pp. 98-9

(which is also a pushmi-pullyu representation, according to Millikan). What I am trying to point out is that with regard to linguistic pushmi-pullyu representations, our cognitive abilities seem to be evolved enough to easily separate the descriptive contents from the directive ones. Why should it not be the case with music? Why not entertain the musical plan, what is communicated in listening to music, without acting upon it considering our evolved music cognition compared with our ancient ancestors?[78] And if we can appreciate a musical plan without acting upon it—that is, we can appreciate the hierarchical organizations among musical ingredients themselves—then moving imaginatively with music (or acting upon that musical plan) becomes a "freebie" in music appreciation. And this is against what Nussbaum tries to establish, since he thinks acting upon a musical plan by moving imaginatively with music is part and parcel of music appreciation. Thus, the main concern is the lack of a convincing explanation of why the kinds of hierarchical structures existing in music based on Jackendoff and Lerdahl's theory motivate an attentive listener to move or imaginatively move with the music or why appreciating musical structures depends on moving with the music.

I should emphasize that I agree with the general framework provided in Nussbaum's theory as regards how attending to musical structures requires a moving response on the listener's part. I just do not think Nussbaum is able to give us a convincing explanation for why such a response is required. Moreover, it appears to me that if we should construe musical content as representational, a matter which is beyond the scope of this paper, it seems plausible to consider it a pushmi-pullyu representation, in the sense that it requires a participatory response

[78] For example, Karl and Robinson (1995) have argued that acculturated listeners can sometimes appreciate more sophisticated "musical plans" or "scenarios" in some works of Western classical music (for example, "dialogues" between different sections in a work), and that this can contribute to the appreciation of more sophisticated meanings in such works.

on the part of the appreciator. But let me set aside what kind of representation musical content is and focus on what I think is missing in the general framework Nussbaum's theory suggests.

The reason appreciating musical organizations of the Jackendoff and Lerdahlian sort requires a listener to move (imaginatively) with music and feel that movement in herself can be provided by my analysis of musical movement in the previous section. As you will recall, the thing-component (what moves) is missing in the music. We hear musical events (for example, tones and chords) as discrete. However, the relational qualities between these events suggest that different changes in music are purposeful and coherent in one way or another—or, they have a unity. For example, some notes in a particular musical scale (say, the ones constituting the tonic chord) might be heard to have more affiliations or proximity to each other than other notes of the same scale, or the changes of pitch in a melody may be heard as somehow purposefully related to each other, as suggesting a process of purposefully related changes. Pressing several random keys on a piano has a different auditory phenomenology than playing the melody of the song *Norwegian Wood* or the folk melody "Row, Row, Row Your Boat" (Figure 3). The differences in pitches in the latter have distinct affordances from those in randomly played pitches. I suggest that to hear coherency, flow, and purposefulness among discrete musical events, we move with music (imaginatively or physically). We move with music in the sense that we feel all the changes among discrete musical events as a coherent, continuous process of change happening inside us. This is, in other words, the way in which our mind makes our auditory experience (of discrete musical events) an experience of continuity, unity and purposefulness: the experience of *music*. Through moving with the music, perceiving the changes among musical ingredients becomes perceiving a purposeful process of change into a moving entity, which is us ourselves

as the listener. This is how I think musical movement is felt by an attentive, acculturated listener.[79]

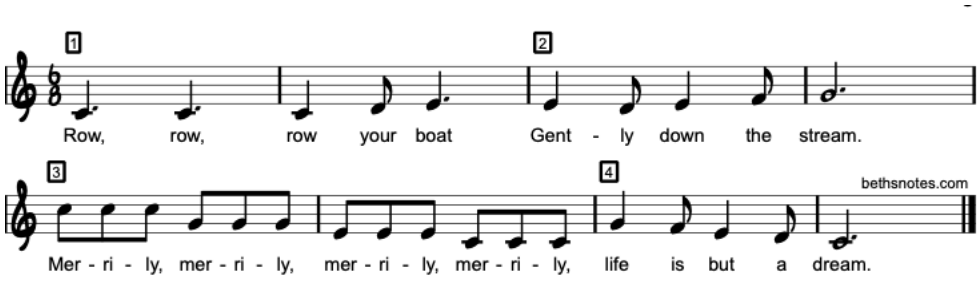


Figure 3. The folk melody "Row, Row, Row Your Boat"

From the perspective of music cognition, Arnie Cox (2016) has argued that musical movement results from the listener's *metaphoric* reasoning about or conceptualization of temporal events in music. He thinks that musical movement is not any real property that we perceive; it is only how we *describe*, *think of*, or *conceptualize* musical events. The backbone of his argument is related to how we usually think of time via special concepts. Two central observations he refers to in this relation are: (a) thinking of time as something that moves independently of us (for example, when we say "Summer is coming") and (b) thinking of time as a "landscape" in which we move (for example, when we say "We are getting closer to the holiday season"). He argues that we conceptualize time in this way because we primarily perceive time through perceiving actions. Actions occur in space; they are constituted by different stages, one coming after the other in space. The act of moving my hand is constituted by a process of change to the location of my hand occurring in a time span. (For example, when I

[79] Viewing "musical movement" from this perspective, the phrase needs to be clarified. By musical movement I do not mean a movement made by the music itself or inside the music itself; I mean a movement that is felt by the listener through attentively hearing the relevant relational qualities of music.

open a door by using a doorknob, my act will include different stages, one coming after another: I raise my hand, move it toward the doorknob, grab it, and so on.) In such an act, space and time are intermingled: the action takes place in space and goes through a time period. Now, holds Cox, when we think of time abstracted from space, we transfer our spatial terms used to perceive actions to our conceptualization of time: "When it comes to talking about temporal relations beyond actual spatial relations, we use the terms learned through actions. The concrete actions and space involved are the source domain, and the abstract temporality is the target domain."[80] For example, when we think of summer holidays (some temporal events), we might use spatial terms, such as "getting closer to (summer holidays)," borrowing the terms from our spatial description of an action in which we move and *get closer to* an object like our car. To Cox, the latter is the source domain from which we borrow a (spatial) term to describe summer holidays (the target domain). He thinks this is the main reason for thinking of music in spatial terms. Musical events are temporal entities, but to describe them, we use spatial descriptions (like how we think of summer holidays.) We heard and "passed" an event (say, a pitch) moments ago and are now hearing or "in" another event (another pitch) and anticipating or "going toward" an event "ahead" (the pitch coming). That is, we perceive musical ingredients (as temporal events) and metaphorically *conceptualize* or *describe* them spatially with movement concepts.

My main worry with this theory is that I am not convinced that musical movement is a matter of *describing* or *conceptualizing* our sensory input in music listening. It seems to me that musical movement itself has a sensory and non-conceptual characteristic. It does not depend on how I describe my experience; it is simply and non-conceptually built into that experience.[81] And

[80] Cox (2016), p. 120
[81] Levinson (2007) has argued that making sense of the unity of a musical piece in music experience is basically *non-conceptual,* specifically because of the transient nature of musical ingredients. An opposing viewpoint has been

interestingly, when it comes to articulating the gist of his theory, Cox himself shifts from the language of conceptualization to a psychological language related to feeling of musical movement, which is akin to the view I proposed earlier:

> In music, a feeling of a relatively unchanging state will motivate a sense of being in a place (as in being "in" a key), a feeling of change will motivate a sense of motion from one state-location to another (as in modulating "from" one key "to" another), and a sense of difference between states will motivate a sense of distance between location (as in interval and key "distances"). […]To the extent that music listening *necessarily involves changes in a listener's physical-psychological state*, perceptions of musical states, changes of state, and differences between states *necessarily* are perceptions not only of what one hears but also of what one *feels in response to* what one hears.[82] (Emphases mine)

In this passage, Cox insightfully suggests that hearing relations between musical events necessarily involves a psychological/feeling response on the listener's part. I am unsure how we can relate this psychological language to the metaphorical language in his theory, but I think what he suggests here can stand independently of his metaphorical theory of musical movement. As I have suggested, by moving with music, the listener fills the gap between the discreteness she hears among musical events and the continuity the relational qualities among those events suggest. A purposeful progress, flow, and continuity form in the music experience when the listener (imaginatively or physically) moves with music. In fact, I believe the listener's psychological response in the form suggested can explain musical movement; we no longer need to appeal to metaphorical thinking, reasoning, or conceptualization to explain the phenomenon. It might be the case that attributing movement to music is metaphorical in the sense that movement is not literally *in* music. Notwithstanding, this does not entail that to perceive musical movement,

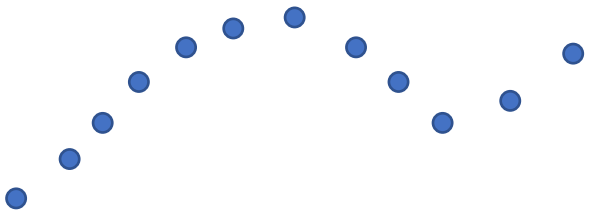
---

put forward by Kivy (2007), who construes the aesthetic experience of music, or at least Western classical music, as essentially conceptual.

[82] Cox (2016), p. 126

the listener has to perform any conceptualization or descriptive thought. Based on what I have proposed, musical movement is felt by the listener in a non-conceptual and visceral manner.

In the last part of my arguments in this section, I want to give a visual analogy to explicate my proposal from a different angle. Consider the following figure:



A part of what we see in this figure is a set of discrete objects (dots) scattered in a 2-dimensional space. However, we can also recognize that the dots are related to each other in space in such a way that they form a specific curve pattern. Particularly if we move our gaze from left to right on this pattern, we can recognize a movement suggested by the locational relations the dots have to each other. Now, suppose we do not see these dots statically but recognize a temporal pattern in the following way. Suppose each dot has a short life span; it appears and disappears after a short period, and suppose these appearances and disappearances are almost in order: initially, we do not see anything, then the first dot from the left appears, then after a short time the second one from the left appears while the previous one has disappeared, and so on. (Additionally, we can suppose that there is some level of overlap between appearances and disappearances, a supposition that will not affect the argument.)[83] The visual recognition of such a temporal pattern can be compared to the auditory recognition of the

[83] This example is similar to Kania's (2015) example of an array of lights that appears and disappears in a specific order.

temporal pattern, for example, in a simple melody (the appearances and disappearances of notes one after another in the context of their tonal relations to each other).

By the above modification to the case, the transient nature of what we see successively will not overshadow our visual recognition of the relevant movement pattern. We will still be able to recognize the curve movement suggested by the relative positions of the dots appearing and disappearing one after the other. However, this transience (the process of appearances and disappearances) will require us to be more involved in our perception of that movement (or the perceptual recognition thereof) compared with the static case, particularly if we want to keep track of the overall pattern the transient dots can make. Something similar happens in the case of music. In the same way, hearing the relational qualities among discrete musical elements requires more attentional involvement in the process of perception. Notwithstanding, there is this important difference: in the absence of a perceived space in which such (musical) movements occur, I think the listener "connects the dots" by herself moving, by experiencing an inner movement. (You will recall that I argued earlier that saying that the listener imagines a musical space in which music moves is not a promising suggestion.[84])

5

Now that I have argued that in response to the relational qualities among discrete musical events, the listener moves with the music, it is time to connect this idea to musical expression. I will also consider two possible objections to my view on musical expression.

[84] By this I do not mean that imagination does not play *any role* in hearing musical movement. As will be explained in the next section, my suggested theory will also appeal to imagination. What I consider as unpromising, based on my earlier discussion, is that the listener imagines *a musical space*. I am not convicted that this particular form of imagination occurs when we hear musical movement.

As I stated in the previous section, the main idea is that the listener feels an inner movement by moving with music (imaginatively or with some sort of physical movement). This feeling has similarities with the feeling component of our emotional experiences. Not that the listener must consciously think of such similarities, but when she feels the musical movement, her feeling experience can become phenomenologically similar to experiencing/feeling an emotion. I say "can become" because I do not think feeling musical movement is always similar to feeling an emotion. There is a movement, say, when the tones of the C major scale are played from C to C (an octave above), and I think to understand the flow of the tones so played the listener moves with the music and feels that movement. However, I do not think this flow of musical tones has any salient expressive quality and I do not think moving with such a succession of tones leads the listener to experience a feeling saliently akin to any emotional feeling. However, sometimes moving with music can become saliently related to emotional feelings. As an example, consider the first movement of Florence Price's Symphony No. 1. When you follow the succession of musical events by moving with and along the passage from measures 22 to 39, you can feel that the particular passage from measures 37 to 39 has a salient expressive quality—it is an emotional moment in music.[85] In fact, the music from measures 22 to 36 prepares you to anticipate, perhaps with some level of nervousness, a moment of bombastic release, a moment that is emotionally charged. (Notice the violin sections in measures 31 and 32 and the trills of the woodwinds in measures 34 to 36, for example.) The movement you feel as a result of attentive listening to this passage is emotionally salient enough to motivate you to describe it with emotive terms.

[85] See the score online in https://issuu.com/scoresondemand/docs/symphony_no_1_58893. In the following performance of the work, the intended passage is between 00:50'' and 01':28'': https://www.youtube.com/watch?v=V-bLUQ2Ec9E (Accessed on May 24, 2023)

It is the musical movement, that is, our moving with music, and the feeling of this inner movement that cause us to hear the emotional qualities of musical passages such as the above-mentioned. As discussed in the previous section, human expression is typically associated with bodily movements. When we are happy, we express it behaviorally differently than when we are anxious or melancholy. Also, when in emotional states, we experience inner bodily changes (an increase in heartbeat, for example). The feeling components of emotions are associated with feeling our body's external and internal movements. The feeling component of happiness can be associated with external and internal movements. For example, when I become happy and smile and jump out of joy, I feel external movements, such as movements of my face (smiling) or my body (jumping), and internal movements, such as my heart racing. As I argued, when I move (imaginatively or physically) in response to music, I also feel that movement and a process of change. This feeling can sometimes be reminiscent of emotional feelings. This is how in moving with music, sometimes I hear and feel musical emotions.

Two objections might be leveled against the view I have developed about musical movement and experiencing musical expression. First, it might be argued that emotions are not merely feelings; they necessarily include cognitive states such as judgments or beliefs about things in the world. When I am angry, I am angry about something, and I make a negative judgment about that thing. I am angry at Saba since I think what she has done is wrong and hurts me. Although my anger has a feeling component (I feel in a specific way when I am angry), what makes my mental state the emotional state of anger is necessarily the judgment (the cognitive state) that comes with anger. The idea is that emotions necessarily have cognitive components.[86]

[86] Pitcher (1965) has defended a version of the cognitive view. Peter Kivy (2001) has used the cognitive view of emotions in criticizing the arousal view on musical expression.

Now, if this theory were correct, our feeling this or that way related to musical movement could not explain why we attribute specific emotional qualities, such as sadness or happiness, to music.

My response to this objection is twofold. First, the cognitive view about emotions is itself a controversial viewpoint. It might not be the case that a cognitive state is the necessary component of an emotion. The totally opposite viewpoint that emotions necessarily are feelings has been a main rival of the cognitive view.[87] Or perhaps a more moderate viewpoint about emotions needs to be taken, that emotions are necessarily a combination of feelings, appraisals (judgments), and action tendencies. In Robinson's (2013) terms, "emotion is a process in which appraisals result in states of action readiness which are then (often) experienced as emotional feelings […]."[88] Or, as Solomon (2008) has stated: emotion is "a complex of many experiences; sensations; various ways of being aware of the world, our own bodies, and intentions; and also thoughts and reflections on our emotions, all melded together in what is typically encountered as a single more or less unified experience."[89] Now, if emotional states are such complex entities, including feelings, and considering the existence of empirical evidence suggesting that have different feelings associated with them,[90] then I think the feeling component I have proposed there is in hearing musical movement contributes importantly to attributions of specific emotions to music.[91] When I move with music, it puts me in a psychological/affective state similar to the state I experience when I feel emotions. This can lead me to describe my (musical) feelings with emotive terms.

[87] Prinz (2005) has defended a version of the feeling view.
[88] p.165
[89] p. 244
[90] Domasio et al. (2000)
[91] It *contributes to* expressive attributions but, as will be explained shortly, might not be sufficient for such attributions.

My second response to the above objection is as follows. I believe that a theory of musical expression should be tolerant about the *emotional ambiguity* or *ineffability* of musical expression.[92] It does not seem that our conceptualization of the emotional qualities of different works is as specific as sometimes assumed. For example, in the passage from Price's symphony talked about earlier, I do hear the emotional qualities of the passage, but I am to some extent unable to describe my experience with exact and highly specific emotive terms. Such conceptualizations and descriptions can be done with a higher level of specificity, for instance, in visual representational works such as Kramskoy's *Inconsolable Grief*, which represents a sad face, and the sadness of the face can be perceived without even paying attention to the title. The movement felt as a result of appreciating expressive music makes the listener aware of an emotional quality in music in a visceral and non-conceptual manner. To conceptualize with which emotion that feeling is associated may sometimes require other factors to come into play. And I believe a salient factor can be culture. Cultural associations can significantly influence how we *describe* our musical experience, particularly with emotional terms. A form of music mostly used in a particular cultural setting can bear some specific expressive association even when heard outside that setting.[93] It is beyond the scope of this paper to investigate how culture contributes to attributing emotional terms to music. But what I want to point out is that my

[92] See Newcomb (1984) and Cook (2001). My claim is not that musical expression is always ineffable or ambiguous. But a theory can be tolerant about many cases in which the expressive qualities of musical passages or works enjoy ineffability and ambiguity.

[93] Nettle (1996) believes in different cultures music is "thought" differently. Blacking (1974, p. 17) has an example of how the expressive quality of a passage can be interpreted in a way that is opposite to how it is emotionally heard in its cultural context if we just focus on the sonic pattern and ignore the relevant cultural interpretations. Movie soundtracks are also other examples of this matter. Musical passages from horror movies such as Jaws or Psycho could be associated with ominous events and carry ominous expressive qualities. Some songs are also played with specific events (such as Mendelssohn's "Wedding March") in a culture which might cause them to carry specific expressive associations and tones. (I am indebted to Sherri Irvin for these examples.) In Iranian musical culture, the timbre of the Ney (a woodwind instrument), carries specific expressive associations because of its being used in funerals and sad events.

theory is meant to provide an account of how the listener's feeling response to music is a function of her aesthetic attention to musical, relational properties; I did not address the problem of how such feelings should be *described* with specific emotive terms. As suggested, by their similarities to emotional feelings, feeling components of our music experience can contribute to such descriptive attributions, but they are not the only factor in this relation.

A second objection to my theory is as follows. It might be noted that in different places in my arguments, I used the terms "hear" and "perceive" in relation to musical movement. For example, I have articulated at several points that the listener hears musical movement. Also, as regards musical expression, I have said that the listener *hears* the expressive qualities of music. But this perception language is in tension with my theory suggesting that musical movement is *felt* by the listener. And if musical expression is based on feeling musical movement as suggested by my theory, then musical emotions are also felt and not heard in music. The objection is that I cannot have the liberty of switching from feeling language to perception language. If we hear musical movement, it means that movement is something *in music* that we hear. On the other hand, if we feel such a movement, then the movement is *in us* and not in music in such a way that it can be heard in it. Either we perceive musical movement or we feel it; it is either in music or in us.

My response is that affective states can be at the core of, part and parcel of, our perceptions of the world. Perception is not only receiving sensory input from the world but the *processing* of input by the perceiver. And it is not an unfamiliar or ungrounded idea that our perception of the world, particularly how the sensory input is processed by the mind, can be infiltrated at its core by different cognitive states (beliefs, attitudes, etc.). I see the tree near my apartment from our bedroom window. I later see that same tree from the dining room window. My angles and, as a

corollary, the sensory input that I receive are different in the two cases. But this change in the sensory input does not make me think that I am perceiving a different object: I understand how vision works and relates me to the objects in the world. There might be a belief shaping (or processing) my sensory input in the two cases, a belief containing the basic idea that *difference in point of view does not mean any difference in the object I see, although it changes my sensory inputs*.[94] This is an example of how cognitive states such as beliefs, thoughts, and attitudes can be at the core of our perceptual experiences.

If this is the case as regards cognitive states, why can it not be so about affective states? I do not think we can simply rule out the idea that affective states can also be central to some perceptual experiences in the sense that they can be an integral part of how our mind processes some sensory input. For example, when I am in danger, perhaps before my beliefs come into play, it is my feeling(s) that make me aware of a dangerous thing happening around me.[95] In other words, it appears that my feeling(s) in such a situation contributes to how my mind processes the relevant sensory input to perceive danger. From this perspective, affective states can play the same role cognitive states can play in perceiving the world. This is why I think they can be considered to be (sometimes) part and parcel of perception. And I tried to show how this is the case regarding musical movement and musical emotions. I have argued in this paper that the feeling of our own movement is the way in which the mind processes the sensory input coming from discrete musical events and their relations in order to perceive the unity and purposefulness of music. This, I believe, legitimizes the use of "hearing musical movement" and

[94] I do not obviously think that a perceiver consciously thinks or conceptualizes their experience in this way in order to perceive the tree. What I mean is that it seems most perceivers somehow have this experiential knowledge about their visual capacity and how it informs them about the world.
[95] Michelle Montague (2009) has argued that feelings, like cognitive states, are intentional states of mind: they are *about* things in the world.

"feeling musical movement" interchangeably, since that feeling is at the core of that hearing/perception.

And the same goes for musical expression, for hearing the emotional qualities of music, I have argued, is based on the feeling of musical movement. Thus, we both hear and feel musical emotions. Regarding whether such musically generated feelings are in the listener or in the music, it depends on how we formulate it. In the sense that these feelings are mental states that the listener is aware of, they are in the listener. But in the sense that they are mental states representing musical qualities (the music's purposefulness and unity), they are representations of qualities in music. But I do not want to make strong claims about this issue because I think the problem of what is *in* music depends on our ontological view about musical works, a territory beyond this paper's scope. To meet Budd's requirement, I tried to delineate in this paper how our feeling responses to music are a function of our aesthetic attention to the relational qualities among musical events/ingredients.

Bibliography

Anderson, W. T. (2011). The Dalcroze Approach to Music Education: Theory and Applications. *General Music Today*, *26*(1), 27–33.

Blacking, J. (1974). *How Musical Is Man?* The University of Washington Press.

Budd, M. (1992). *Music and the Emotions: The Philosophical Theories*. Routledge.

Cook, N. (2001). Theorizing Musical Meaning. *Music Theory Spectrum*, *23*(2), 170–195.

Cox, A. (2016). *Music and Embodied Cognition: Listening, Moving, Feeling, and Thinking*. Indiana University Press.

Damasio, A. R., Grabowski, T. J., Bechara, A., Damasio, H., Ponto, L. L. B., Parvizi, J., & Hichwa, R. D. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature Neuroscience*, *3*(10), Article 10. https://doi.org/10.1038/79871

Cowen, A. S., Fang, X., Sauter, D., & Keltner, D. (2020). What music makes us feel: At least 13 dimensions organize subjective experiences associated with music across different cultures. *Psychological and Cognitive Sciences*, *117*(4), 1924–1934.

Davies, S. (1994). *Musical Meaning and Expression*. Cornell University Press.

Farhat, H. (1990). *The Dastgah Concept in Persian Music*. Cambridge University Press.

Gregory, D. (2010). Imagery, the Imagination and Experience. *The Philosophical Quarterly*, *60*(241), 735–753.

Jackendoff, R., & Lerdahl, F. (2006). The capacity for music: What is it, and what's special about it? *Cognition*, *100*(1), 33–72.

Kania, A. (2015). An Imaginative Theory of Musical Space and Movement. *British Journal of Aesthetics*, *55*(2), 157–172.

Karl, G., & Robinson, J. (1995). Shostakovich's tenth symphony and the musical expression of cognitively complex emotions. *The Journal of Aesthetics and Art Criticism*, 53(4), 401-415.

Kivy, P. (1981). *The Corded Shell: Reflections on Musical Expression*. Princeton University Press.

Kivy, P. (2001). The Arousal Theory of Musical Expression: Rethinking the Unthinkable. In *New Essays on Musical Understanding* (pp. 119–151). Clarendon Press, Oxford.

Kivy, P. (2007). Music, Language, and Cognition: Which Doesn't Belong? In *Music, Language, and Cognition* (pp. 214–232). Oxford University Press.

Levinson, J. (1996). Musical Expressiveness. In *The Pleasures of Aesthetics: Philosophical Essays* (pp. 90–125). Cornell University Press.

Levinson, J. (2007). *Music in the Moment*. Cornell University Press.

Levinson, J. (2009). The Aesthetic Appreciation of Music. *British Journal of Aesthetics*, *49*(4), 415–425.

Khaleghi, R. (1937/2023). *A View on Iranian Music (Sections 1 and 2)*. Safi Ali-Shaah. (First published in 1937) Farsi title:

Locke, J. (1979). *An Essay Concerning Human Understanding* (P. H. Nidditch, Ed.; 1st ed.). Clarendon Press.

Matherne, S. (2016). Kant's theory of imagination. In *The Routledge Handbook of Philosophy of Imagination* (pp. 55–68). Routledge.

Matravers, D. (1998). *Art and Emotion*. Oxford University Press.

Meyer, L. B. (1961). *Emotion and Meaning in Music*. University of Chicago Press.

Millikan, R. G. (1995). Pushmi-Pullyu Representations. *Philosophical Perspectives*, *9*, 185–200.

Montague, E. (2011). Phenomenology and the 'hard problem' of consciousness and music. In *Music and Consciousness: Philosophical, Psychological, and Cultural Perspectives* (pp. 29–43). Oxford University Press.

Montague, M. (2009). *The Logic, Intentionality, and Phenomenology of Emotion*. 171–192.

Nettl, B. (1996). Ideas about Music and Musical Thought: Ethnomusicological Perspectives. *The Journal of Aesthetic Education*, *30*(2), 173–187.

Newcomb, A. (1984). Sound and Feeling. *Critical Inquiry*, *10*(4), 614–643.

Peacocke, C. (2009). The Perception of Music: Sources of Significance. *British Journal of Aesthetics*, *49*(3), 257–275.

Pitcher, G. (1965). Emotion. *Mind*, *74*(295), 326–346.

Prinz, J. (2005). Are Emotions Feelings? *Journal of Consciousness Studies*, *12*(8–10), 9–25.

Robinson, J. (2013). Three theories of emotion—Three routes for musical arousal. In T. Cochrane, B. Fantini, & K. R. Scherer (Eds.), *The Emotional Power of Music: Multidisciplinary perspectives on musical arousal, expression, and social control* (pp. 155–168). Oxford University Press.

Shusterman, R. (1991). Form and Funk: The Aesthetic Challenge of Popular Art. *The British Journal of Aesthetics*, *31*(3), 213.

Scruton, R. (1999). *The Aesthetics of Music*. Oxford University Press.

Solomon, R. C. (2008). *True to Our Feelings: What Our Emotions Are Really Telling Us*. Oxford University Press.

Walton, K. L. (1988). What Is Abstract about the Art of Music? *The Journal of Aesthetics and Art Criticism*, *46*(3), 351–364.

Walton, K. L. (1999). Projectivism, Empathy, and Musical Tension. *Philosophical Topics*, *26*(1 & 2), 407–440.

Chapter 4: Sounds and Sound Sources

1

When we hear a dog barking outside in the yard, what do we hear or perceive exactly? One thing seems obvious: we hear a sound, the barking. But the way we talk about the experience—"I am hearing a dog outside"—seems to suggest that our object of perception is the dog, the source of the sound. So, maybe we hear the sound *and* the dog. But how should we distinguish hearing the dog from hearing the sound it makes? It might be that my hearing the dog is *through* my hearing the barking. But now the problem is how we should understand this "through" relation (that is, how a sound causes us to perceive its source, if it does).

An idea common to some philosophical views on sounds and auditory perception is that sounds are perceived as distinct from their sources. For example, related to his arguments on musical expression and following Stanley Cavell's and Italo Calvino's reflections on hearing, Kendall Walton (2015) says:

> Sounds have lives of their own. They are distinct from and independent of their sources, of the things that emit, cause them. They have their own temporal dimensions as well: a sound arrives at our ears after the event that give rise to it; in some cases perceptibly after the event. Sounds fill rooms, and travel across streets. Sights don't do that.[96]

To Walton, the way we perceive sounds independently of their sources is related to our perceiving and treating them as independent objects: "[W]e reify or objectify sounds. We speak of clatters, bangings, whinnys, murmurs, echoes, creaks, clangs, rustles, grumbles, gurgles."[97]

[96] p. 43.
[97] Walton (2015), p. 169.

The same idea has also appeared in Roger Scruton's non-physicalist theory of sounds. Scruton argues that we hear sounds as independent objects, but objects of a special kind: objects that are perceived in the world but whose existence is totally mind-dependent. He calls such an object a "secondary object," "an object all of whose properties are ways in which it appears" (58). According to this view a sound is constituted merely by a set of secondary qualities audible by a perceiver. To Scruton, this bundle of audible qualities is perceived by the hearer as an object distinct from a sound source.

To support his view Scruton refers us to our auditory experience. For example, in explicating why he thinks sounds are "listened to in isolation [from the sources or physical events causing them]" he says: "The striking thing is that sounds, thus emancipated from their causes, are *experienced* as independent but related objects, which form coherent complexes with boundaries and simultaneities, parts and wholes."[98]

A similar phenomenological idea has been used in O'Callaghan's physicalist theory of sound.[99] According to O'Callaghan, "sounds are events located at a distance from their perceivers. They occur at or near their sources, and travel only if their sources travel."[100] This is a different view from Walton's and Scruton's; but in O'Callaghan's theory a sound is treated as a "*particular individual*," an object, as well. This itself is founded on a phenomenological reflection on auditory experience. To O'Callaghan, we hear a sound as a distinct object, as an entity that "survives changes to [its] properties across time."[101] In my understanding, what O'Callaghan means is that similar to the way in which a particular can change with respect to

[98] Scruton (2009), p. 58 (emphasis mine).
[99] O'Callaghan (2017), particularly Ch. 1.
[100] Ibid., p. 22.
[101] Matthew Nudds (2001) also finds a similar idea compelling, "that our experience of sounds is of things which are distinct from the world of material objects […]" (p. 210).

one of its properties without turning into a different particular, a change in one property of a sound does not necessarily make it a different sound; for example, the same sound can be quiet or loud. For O'Callaghan this is another reason for considering sounds as objects or particulars.

As we can see, a common theme appearing in these philosophical views on sound is that introspecting audition shows that sounds are perceived as objects distinct from their sources. I will refer to this phenomenological observation on auditory perception as the Distinction Observation (DO). However, (DO) seems to be in tension with another phenomenological observation about hearing which can also be supported with empirical evidence.

Think of the many occasions on which we talk about hearing sound sources directly, for example when one says, "I hear a dog barking outside" or "I hear a violin." It seems that in such situations we can directly hear objects or physical events in our surroundings through the sounds they make. There is also empirical evidence showing how our auditory system enables us to perceive sound sources—to perceive the source's location[102], to perceive its size[103], and to perceive some mechanical features related to how an object is stimulated to vibrate and make sound.[104] (We will return to this matter in more detail later.) As Stephen Handel says about the auditory world: "[A]coustic properties adhere to *objects*; the properties belong to and at the same time characterize the source."[105] If Handel is right and the audible properties we hear in the course of auditory perception *belong to* sound sources, one upshot is that in hearing sounds we directly perceive sound sources. And it supports the way in which we sometimes talk about hearing material things and happenings in the world. Let us call this observation about our

[102] See Darwin (2008).
[103] See Patterson, R. D., Smith, D. R. R., van Dinther, R., & Walters, T. C. (2008) and van Dinther, R., & Patterson, R. D. (2006).
[104] Giordano, B. L., & McAdams, S. (2010).
[105] Handel, S. (1995), p. 425 (emphasis added).

auditory experience the Immediacy Observation (IO).[106] According to (IO) in hearing we directly auditorily perceive sound sources since our auditory system seems to be able to perceptually/auditorily present sound sources to us. And by sound source it is meant an *object* (a flute, a metal bar, a human vocal tract, etc.) that is stimulated or vibrated in an event.

Now, one problem before us with regard to auditory perception is whether and how (DO) and (IO) can be consistent with each other, that is, whether and how it could be the case both that we hear sounds as objects perceptually distinct from their sources (DO) and that hearing is the perceptual presentation of sound sources (IO). For example, when one hears the sound of a violin being played (without seeing the instrument itself), could it be the case both that the sound one is hearing is an object of her perception distinct from the violin *and* that through the auditory properties one is hearing, one perceives the violin being played?

Let us be clear. The intended tension between (DO) and (IO) is not that the content of auditory perception cannot consist of two objects. If it was shown that a sound is an ontologically distinct object, it might also be the case that both a sound and a sound source can be objects of auditory perception. In this way there is no tension between (IO) and (DO). Generally speaking, it is obvious that there is no inherent inconsistency in saying that two distinct objects can feature in the content of perception in any sense modality. For example, when one sees a butterfly sitting on a flower, both the butterfly and the flower, among other things, are in the content of that visual experience.

However, the tension between (IO) and (DO) is this. If (IO) is the case, then sound and its source seem to share some of their audible properties; that is, there seem to be properties we hear

[106] (IO) is basically what Leddington (2014, p. 324-5) calls *the Heideggerian view* on sound and hearing.

in sounds that belong to their sources. And this is against what is claimed by (DO), that the properties we hear in sounds only belong to those sounds and not their sources. What we need to remove this tension is an understanding of whether and how both (IO) and (DO) can be accepted.

In what follows I want to delve into this problem and consider what it entails about the ontological nature of sound (what sound *is*). As will be clear, the reason that I will consider the ontological implications of my discussion is that our approach toward the problem I have raised regarding how sound and its source are perceived to be related is entangled with how sound is construed to be ontologically. I will first critically examine two views in the literature on the nature of sound and how it is related to its source: the *causal view* and the *mereological view*. Both of the views appeal to (DO), and I will argue that neither of them can cogently explain (IO). I then propose my solution according to which hearing a sound is auditorily directly perceiving properties of a material object and the event happening to that object (like bouncing or colliding). Sound is, I will suggest, a *complex property*. I will argue that this view can explain both (IO) and (DO).

2

The first view I want to examine is that the relation between a sound and its source is causal (the *causal view*).[107] According to the causal view, sounds are perceived as distinct particulars and effects having causes in our surrounding environment.[108] From this perspective, a source can be thought of in a mediated way by a perceiver, but it is not perceived directly. For example, when I see a ball moving, I might think of the thing that caused the ball's movement (a

[107] Scruton (2009); and Nudds (2001). "Experiencing the Production of Sounds," *European Journal of Philosophy* 9:2, 210-229.
[108] For a critical discussion of the causal view, see also O'Callaghan, C. (2011), pp. 375–405.

person, wind, etc.). When I do not directly see the cause, my thinking of it is mediated by seeing the ball moving. Similarly, according to the causal view, what is directly perceived in an auditory experience is the sound as an effect. And then this perception might lead the hearer to think of the source (the cause). But how?

What the advocates of the causal view appeal to in order to explain how a hearer can perceive a sound as being *of* or, specifically speaking, as causally related to something in the world is the hearer's past experience or knowledge about the world. To elucidate the idea, let us consider a famous passage by Berkeley on hearing: "when I hear a coach drive along the streets, immediately I perceive only the sound, but from experience I have had that such a sound is connected with a coach, I am said to hear the coach."[109] According to Berkeley whether or not the person can be said to hear the coach (source) is dependent on her past experience or, more particularly, dependent on whether or not she *knows,* based on her past experience, that the sound she is hearing is the sound of a coach. I think "past experience" in this context can be interpreted in two ways. It can be a hearer's visual-auditory experience, an experience in which the hearer sees a source and, at the same time, hears what kind of sound it makes—for example, seeing a coach moving and hearing the sound it makes or seeing a violin and hearing its sound. Based on such a visual-auditory experience, a person can probably perceptually recognize a familiar sound even when the source is visually hidden.[110] But "past experience" does not need to be limited to such visual-auditory experiences; it can also be the general knowledge of a hearer about the world. For example, suppose one hears the sound of the Chinese instrument the guqin without seeing the instrument. Suppose also that the person has not ever seen the instrument making any

[109] Quoted in Frank Jackson (1977), p. 7.
[110] Nudds (2001) suggests that such visual-auditory experiences can explain how a hearer hears a sound as causally related to its source.

sound before. So, she does not have the related past visual-auditory experience. But suppose the hearer has heard this sound before and was told by someone whom she trusted that that was the sound of a guqin. Now when the person is hearing the sound (of a guqin) she can recognize it and say, "I am hearing a guqin."

The main point is that according to the causal view, in cases that someone only hears a sound without seeing the related source, if the person does not have a related past experience, either in the form of a visual-auditory experience or some other non-perceptual form of knowledge, she would not be able to connect the sound she is hearing to any source (a thing in her surroundings). In other words, she cannot be said to *hear* the source. She merely hears the sound. As Scruton says: "It is in some sense an accident if we can attribute a sound to a particular—to say that it is the sound *of* this thing, caused by changes *in* that thing, and so on."[111] To my understanding, he thinks it is an accident because it is an accident that we know (based either on visual-auditory experiences or on non-perceptual knowledge) that a sound is caused by this or that thing in the world.[112]

There are three main worries, however, with the causal view. First, by appealing to past experience to explain how a hearer can hear a sound as being of a thing, the advocates of the view seem to conflate perception and knowledge. As Jackson, in response to Berkeley's passage quoted above, has said:

> This is a puzzling claim. Berkeley purports to be talking about hearing the coach, hearing the sound, and about the relation between the two; but what he says is plausible only if construed as being about *believing* (or *knowing*) that one is hearing a sound and hearing a coach. For example, 'the experience that I have had that such a sound is connected with a coach' is irrelevant to whether I hear a coach. There is such a thing as hearing a coach for

[111] Scruton (2009), p. 62.
[112] This is why Scruton thinks that "It would be quite possible for us to be surrounded by sounds […] which we individuate, order, and interpret without assigning to any of them a physical process as origin or cause." (Ibid.)

> the *first* time, and so, one may hear a coach in the absence of the past experience Berkeley refers to. Past experience is only relevant to the quite separate question of whether I believe or know that the sound is that of a coach, and hence to whether I hear *that* there is a coach outside.[113]

The distinction Jackson has plausibly drawn is between what one perceives and what one knows about her object of perception. This is applicable to visual experience as well. I might see an object that I do not know anything about (its name, whether it is animate or inanimate, etc.). However, this does not mean that I do not *see* the thing. I think Jackson is right that this can be said about our auditory experience as well. Whether one can form any belief about the source of the sound she is hearing is distinct from whether one hears the source. In an auditory case, I may not know that I am hearing the sound of a guqin; so on such an occasion I cannot *say* or form the *belief* that "I am hearing a guqin." But this does not necessarily rule out that I *hear* the guqin. We can then say that past experience is primarily related to *forming beliefs* about the object of auditory perception and not necessarily about the very *perception* of the object. So, the fact that in auditory cases in which a hearer does not see the source she should somehow refer to her past experience in order to *say* that she hears the sound of a particular thing does not show whether or not the hearer directly perceives that thing in the absence of such a past experience.

The second worry about the causal view is related to the phenomenology of auditory experience. Besides the fact that in many cases of hearing we describe our auditory experiences in such a way that we seem to be able to directly perceive things in the world (hearing a *dog*, *a door being slammed*, etc.), our auditory experience of sound sources seems phenomenologically to be different from experiencing causal relationships between two entities. For example, imagine you see a ball having hit the window close to which you are sitting, and you infer that it

[113] Jackson (1977), p. 7.

must have been thrown by someone from outside.[114] In such a case the effect directs you to make an inference about the cause. That is, by visually perceiving the effect you form a *thought* or *belief* about its cause. The cause is not visually *perceived* here. The link you make between the effect, which you perceive, and its cause is *purely* cognitive (inferential).[115] However, when you hear a sound coming from a skillet banged by a spoon in the kitchen, it seems that in the core of your auditory experience you auditorily *perceive* a physical happening in your surrounding (in the kitchen) to which your attention turns. In other words, in hearing the sound in such a situation your attention is directed to somewhere and something in the world, the source of the sound. However, this is not what happens in the case of the hitting ball; it would, I think, be implausible to say that in seeing the hitting ball your attention is automatically directed to the cause, to somewhere or something beyond what is happening at the window. Here, seeing the hitting ball can be considered as a perceptually distinct scenario from seeing its cause (for example, seeing the person who is throwing the ball)—that is, the two are not perceptually intermingled. But in the auditory case imagined, *hearing* the sound and *hearing* the related physical happening (banging of the skillet) are perceptually intermingled. The point is not that the hearer, in the auditory case, might not draw any inference about the source of the sound (she

[114] The idea that we seem to hear sources directly does not deny that in auditory perception we might *also* infer some things about the source. It suggests that in hearing we seem to have some direct or immediate perception of a source. If that is the case, it undermines the causal view claiming that our immediate and primary object of hearing is a sound (effect) and not its source (cause).

[115] It might be the case, as Susanna Siegel (2012) has argued, that causality can be visually *perceived*. What I am arguing here is not, however, necessarily against such a viewpoint. Siegel's argument is, in my understanding, concerned with those cases in which someone sees both the cause and the effect, and the view is that in such cases the person also visually perceives the causal relationship between the two, or, in other words, the causal relationship is part of the content of the person's visual perception. However, in my imagined case the perceiver does *not* see the cause (that is, the cause of the ball hitting the window is visually hidden from her). What I am saying is that in such cases, the perceiver has to make inference(s) about the cause based on background knowledge. Or, let's put it in this way: suppose there is a causal relationship (CR) between X (effect) and Y (cause) and let's formulate it as X-CR-Y. Siegel's argument is that when one can see both X and Y, CR is also part of her visual perception (she *sees* CR). In my case Y is visually hidden from the perceiver. Now, it might be either that only X is in the content of the related visual experience or that both X and CR are. Either way, what I am saying is that the perceiver has to make inference(s) about Y in such a case.

might infer that her friend or husband is in the kitchen making dinner, she might infer what is being cooked, and so on and so forth). My point is that at the core of this auditory experience hearing the sound gives the hearer a perceptual/auditory access to something located somewhere in her surroundings.

Here, a causal theorist might argue that we recognize the causal relationship between a sound and its source in a cross-modal experience in the following way. We see a sounding object (say a violin being played); we, at the same time, hear the sound (the violin's sound); and we conclude that what we hear (the sound) is something caused by what we see. Now, it might be argued that the view that sound is a distinct entity caused by its source can phenomenologically be corroborated by considering such cross-modal experiences in which we hear sounds and see their sources at the same time.

However, I think the above argument is question-begging and cannot phenomenologically support the causal view. It is question-begging because it assumes what is at issue: it assumes that what we hear in the depicted cross-modal scenario is a distinct *entity* caused by the object or event we see. Unless one is theoretically inclined to think that sound is a distinct entity caused by its source, the scenario cannot clearly support it. It of course shows that our *hearing* of the sound is caused by what we see. But this does not mean that in such a case we phenomenologically experience a causal relationship between *two entities*: a sound and a sounding object. The scenario can in fact be phenomenologically described in a way opposite to the claim of the causal view. It can also be said that in such a scenario the perceiver experiences the same thing (an object participating in an event) in two perceptual modes at the same time: auditory and visual. That is, instead of saying that the perceiver in such cases hears *something* and thinks it is causally related to what (the event) she sees, we can plausibly say that she hears

the same thing (the event) that she also sees.[116] And I do not see why we should prefer the causal theorist's interpretation of such cross-modal scenarios to the alternative one just given.

So, the second problem with the causal view, which I examined through the comparison of the hitting ball case and the banged skillet case, and through my above discussion of the cross-modal scenarios—is that our phenomenology of hearing goes against the causal view.

A third worry about the causal view is that it is inconsistent with some empirical evidence about our perception of acoustic properties. If we accepted the view, we would have to say that in auditory experience the acoustic properties we hear (timbre, pitch, etc.) are exclusively those of a sound and not of its source. For according to this view, a sound is a perceptually independent entity which is causally related to a source but has its own properties that are perceptually distinct from the source. But there is different empirical evidence suggesting that a sound and its source are deeply perceptually intermingled. In fact, a body of empirical research is focused on how through audition we *perceive* sound sources.[117] There is a central theme in these studies: "We rely critically on our ability to identify everyday objects and events from a sound to function normally in a world that at any given moment is largely out of view."[118] From this perspective, our auditory perception is considered as one of our evolved abilities to deal with *objects* in the world. As William A. Yost has said: "Auditory sensation may be viewed as processing the physical attributes of sound: frequency, level, and time. Auditory perception may be the additional processing of those attributes that allow an organism to deal

[116] And we can also say that the perceiver is aware that both her auditory and visual perceptions, which are concurrent in such cross-modal cases, are *caused by* the object participating in an event (for example, by a violin being bowed). However, we should notice that such a causal description of these cross-modal cases does not go in the direction of the causal view under examination because that view does not merely say that an auditory *perception* is caused by a source but that the source causes *another entity* (sound) to come into existence.
[117] See, for example, Yost, et al. (2008).
[118] Lutfi, R. A. (2008), p. 13.

with the sources that produced the sound."[119] This broader understanding of auditory perception shows how what we perceive in hearing seems to go beyond auditory sensations. Hearing timbre is an example of this.

In perceiving the sound of an object or happening, one property that our auditory system recognizes is timbre. The sound of the flute has a different timbre than the sound of the piano, and the sound of a skillet banged by a metal spoon has a different timbral quality than when it is banged by a wooden spoon. There is empirical evidence showing that what timbre represents to an ear is in part the physical characteristics and mechanical processes involved in producing a sound, also known as source mechanics. For example, based on their examination of empirical studies on the timbre of different musical instruments, Giordano and McAdams (2010) conclude: "Overall, the analyses reported here are consistent with current theories of timbre, according to which this multidimensional attribute of auditory sensation is indeed the perceptual correlate of the mechanics of the sound source."[120] Their study specifically supports the idea that there are factors "in the mechanics of the sound source [that are] actually differentiated in timbre perception."[121] The two main factors they identify are: the physical qualities that place a musical instrument in a certain family (string instruments like cello or guitar, membranophones like tympani, idiophones like xylophone, and wind instruments like flute), and the excitation type (how the instrument is struck to sound). This result more specifically characterizes Stephen Handel's (1995) earlier finding about timbre, that "the [acoustic] cues that determine timbre quality are interdependent because all are determined by the method of sound production and the physical construction of the instrument"[122]; it also justifies why Handel "use[s] the term timbre

[119] Yost, W. A. (2008), p. 1.
[120] P. 165.
[121] Ibid. p. 164.
[122] p. 441.

to refer to the perceptual qualities of objects and events; that is, 'what it sounds like.'"[123] These empirical studies suggest that in hearing a timbral quality our auditory system enables us to directly auditorily perceive physical features related to the source such as how a sounding object is struck to sound, that is, what kind of event it participates in to sound, and what kind of physical construction or material the sounding object has.[124] Let me clarify what I mean by saying that one can *directly auditorily perceive* such features.

Suppose you hear the sound of a violin being bowed but do not see the instrument and do not even know which object or event in your surroundings you hear. After this experience, if later you hear another sound made by the violin, again by being bowed and again without seeing the instrument, you would be able to auditorily recognize that you hear the same (type) of sounding object in your surroundings. However much more you hear the violin being bowed in different auditory experiences, your auditory recognitional capacity with regard to the instrument and its making of sound (via bowing) would increase. Specifically, regarding timbre, we can say that you can auditorily recognize, in these experiences, the instantiation of basically the same timbral property. And if in the perception of timbre some features of source mechanics are represented (as discussed earlier), we can say that in the imagined hearing cases your ear is able to auditorily recognize the same type of physical process a particular (type of) object (the violin) undergoes.[125] This recognition, however, is independent of whether you know the *name* of the instrument or the *name* of that form of stimulation; this is peripheral to the main idea being argued for: the point is not about how we *conceptualize* our perception but about the *perception*

[123] Ibid. p. 426.
[124] See also Balzano (1986) and Gaver (1993).
[125] A body of empirical research on timbre centers on this recognitional capacity of audition, and specifically on the hypothesis that "the sensory dimensions that compose timbre serve as indicators used in the categorization, recognition, and identification of sound events and sound sources." (McAdams (2013), p. 46)

itself. If you had previously been told that you were hearing *a violin being bowed*, you would be able to name or conceptualize your subsequent auditory experiences in that way; but this is irrelevant to the core auditory recognitional capacity delineated above, the fact that you can *auditorily perceive* the same property no matter whether you can conceptualize it or not. Comparing this with a visual analogy can clarify the point more.

Imagine you see an object having a particular color unfamiliar to you—say, seeing the color violet for the first time in a piece of cloth. Based on that experience, your visual system basically enables you to recognize other instances of that color in your future experiences of it regardless of whether you know the name of that color or how you particularly conceptualize it.[126] As Susanna Schellenberg has said: "[…] perception is constituted by employing perceptual capacities–for example, the capacity to discriminate and single out instances of red from instances of blue." Singling out instances of color is a perceptual capacity of vision. This is applicable to properties like shape, size, or texture as well. I think that this discriminatory capacity of perception is distinct from the conceptualization of perception.

What I am saying about timbre is related to the perceptual capacities of audition. For sounds having the same source, for example the sounds of the violin being bowed, the auditory system enables us to recognize the same timbre (or similar timbral quality), no matter whether we can conceptualize that perceptual recognition or not.[127] Considering this together with the

[126] It might be said that such perceptual recognitions necessarily have a cognitive core; for example, in the recognition of a color in an object the perceiver must have a general belief about her experience in the form of, say, "There is an object which is in some particular ways in front of me (having colors, shape, etc.)." However, even if it was the case, this does not undermine what I am saying here, because my point is that in the perceptual recognition of different instances of a specific color the perceiver does not seem to need to form beliefs *specifically* about that *particular* color (that is, naming or conceptualizing that specific color) although she might need to form beliefs *generally* about her related (perceptual) experience.

[127] By conceptualization I mean using a correct linguistic label to describe our experience, for example, experiencing a red color, or the timbre of violin, and describing it with the relevant correct term "red" (or "the sound of violin").

empirically supported fact that timbre is constituted by the features of source mechanics supports my claim that in audition we can hear things such as *a violin being bowed*, *a guitar string being plucked*, *a wooden piece being knocked* and so on and so forth. Audition enables us to perceptually recognize such features and conceptualize them, although this conceptualization is, as argued, distinct from and additional to the intended recognitional capacity.

Now, the discussion of timbre and our auditory recognitional capacity with regard to it has hopefully clarified the third worry I raised against the causal view. The view claims that sound is a distinct object of perception and timbre is a property of this object. The timbre discussion showed that through perceiving timbre the hearer perceives (aconceptually or conceptually) features of an object (for example, that it is metal) and source mechanics (for example, that it is being scratched or struck). This goes against the causal view because it shows that what the hearer perceives (aconceptually or conceptually) in sound through timbre are features of *the source*. That is, some of the features heard in sounds *belong to* their *sources* and do not belong to the sound postulated by the causal view as an object caused by and distinct from its source. We also argued that audition enables us to perceptually recognize physical objects and processes in our surroundings based on how they sound. (I hear *a violin being bowed*, without seeing the instrument itself. Through the perception of timbre my auditory recognitional capacity enables me to recognize the same (type of) object undergoing the same physical process—a violin being bowed—when I am having other similar auditory experiences. And if someone said to me, during the first such experience, that I was hearing a violin being bowed, I would also be able to conceptualize my later similar experiences in that way. Compare: I see an object with a pentagonal shape. The visual system enables me to recognize that similar shape when I see

pentagonal objects or drawings later. And if someone told me that I saw a pentagonal shape the first time, I would be able to conceptualize the other similar visual experiences in that way.)[128]

The three worries discussed undermine the plausibility of the causal view. The view cannot explain why we sometimes seem to be able to directly hear sound sources, as the way in which we sometimes talk about our auditory experience suggests—what we called the Immediacy Observation (IO). In fact, the causal view eliminates (IO) in favor of the Distinction Observation (DO) by claiming that sound is a distinct and independent object of perception. As we tried to show, appealing to past experience does not help the causal view to explain (IO) away. On the contrary, both phenomenological considerations and empirical evidence, particularly on timbre, suggest that we need to take (IO) seriously. If we can find a way to explain both (DO) and (IO), that will be a plus for our theory. I want now to turn to a more promising proposal and its problems.

3

Another suggestion for how sounds and their sources can be related has been put forward by O'Callaghan (2007 and 2011). Before discussing how O'Callaghan views that relation, we need first to consider what sound *is* according to him. O'Callaghan argues that the most plausible candidate for the nature of sound is *event* rather than *acoustic wave* or *property*.

[128] What I have said about the recognitional capacity of perception is not necessarily tied to the realist view of properties, that in each case of property perception the perceiver recognizes the instantiation of a universal or repeatable. Even in the context of a trope theory we can also talk about such perceptual recognitional capacities. For example, in the case of, say, perceiving red in a red object, the visual system enables the perceiver to recognize that the perceived red-trope is *similar to* the other red-tropes she experienced in the past. This is applicable to the perception of timbre as well.

He rejects considering sounds as waves existing in a medium (like the air) between a source and the hearer based on two phenomenological considerations. The first is that the wave view makes the locational feature of sound illusory. The phenomenology of hearing suggests that we hear sounds as located at or near their sources; we hear them at roughly specific places around us, where their sources are. (While walking you hear a bird, and you hear it somewhere above you on your left. You turn and see the tree from which the sound is coming, and by some more effort you will probably be able to locate the bird on a branch with the help of the sound you hear.) Now, if we accepted the wave view, we would have to consider the locational feature of hearing as illusory because sound waves are located from the source to the hearer—or we can say they travel from the source to the hearer—and are not located at or near the source as the phenomenology of hearing suggests.

O'Callaghan's second phenomenological consideration to weaken the wave view is the following. When we hear sounds, we typically hear them as things lasting in a specific (or roughly specific) time span, as things that start and end at specific (or roughly specific) points of time. I hear, from my apartment, a honking, and I hear the honking as lasting a time span, say, about two seconds. However, if the sound of honking were a wave in the air, it would exist even after I hear it to end—more than, say, two seconds. This means that if sound were a wave in the medium the time period in which I hear the sound to last would be illusory, because typically sound waves last much longer than sounds seem to us to last.

Both of these phenomenological considerations about the wave view show that the view, as Casati, Dokic, and Di Bona (2020) have put it, "[accounts] for auditory perception in terms of a massive error theory." Considering this, it would be an explanatory advantage for a metaphysical theory of sound if it could incorporate two of the main characteristics of the

phenomenology of hearing—that is, locationality and temporality. The wave view is not able to do that, and I agree with O'Callaghan that this makes it an unconvincing theory of sound.

O'Callaghan also thinks that the property view—another main candidate for understanding the nature of sound—has serious flaws. According to the property view a sound should be understood as a property of an object similar to other sensible properties an object can have (colors, shapes, etc.). He discusses two main problems with this view. First, he argues that the property view "fails to account for the essential temporal characteristics of sounds"[129] because it dismisses the fact that a sound temporally behaves like the *bearer* of properties and not a property itself. His reason is that sound, like any other particular possessing its own properties, can survive changes to its properties. For example, to O'Callaghan the pitch of a sound can change without the sound becoming a different one. A musical example he gives is "when an instrument seamlessly shifts from playing C sharp to playing B."[130] He thinks that here we have *one* sound one of whose properties (its pitch) changes, but the sound itself survives that change. Second, he argues that sounds, unlike properties, do not seem to qualify other objects or events. In his view, sounds themselves are particulars being qualified by the properties of loudness, pitch, and timbre. An example of this can be when an instrument plays a middle C; in this case the pitch does not qualify the instrument but the sound it makes.

I do not intend here to examine O'Callaghan's arguments against the property view. As I will argue in the next section, both the arguments can be responded in my suggested version of the property view. But I am going here to discuss O'Callaghan's own theory according to which a sound is neither waves in a medium nor a property of the sounding object but an *event*.

[129] Ibid., p. 22.
[130] O'Callaghan (2017), p. 24. As a specific example consider when the violin plays a few tones in a single bowing.

To see what kind of event O'Callaghan thinks a sound is let us consider an example. Suppose a hollow metal bar is hit with a wooden stick. As a result of this happening the bar vibrates and causes sound waves in the air which itself causes a person close enough to hear a sound. One main event in this scenario is the vibration of the bar. The other event is the disturbance or molecular motions of a column of the air between the bar and the hearer. (This can be controversial depending on whether we want to regard sound waves as events in a medium; but here it is assumed that sound waves are events.) Besides these two events in our imagined scenario, O'Callaghan believes that there is also another event: the vibrating bar's *disturbing* the air. That is, O'Callaghan thinks the vibrating bar's disturbing the air in an oscillational fashion is itself a distinct event in the scenario. A compelling reason for considering it as a distinct event is that there would not be any disturbance of a medium if the bar vibrated in a vacuum. This shows that the vibration of the bar is not identical with its disturbing the air. Now, O'Callaghan identifies sounds with such disturbance events. He says:

> According to this account, sounds are particular events of a certain kind. They are events in which a moving object disturbs a surrounding medium and sets it moving. The strikings and crashings are not the sounds, but are the causes of sounds. The waves in the medium are not the sounds themselves, but are the effects of sounds.[131]

What I want to highlight in this passage is the distinction O'Callaghan draws between sound, as a disturbance event, and what *brings about* that disturbance, the source of that disturbance (events such as strikings or crashings). Here, O'Callaghan construes the relation between these two—source and sound—as *causal*. This points in the direction of the causal view we examined in the previous section; that is, it suggests that sound is a distinct entity (a disturbance event) caused by its source. However, in his later work, O'Callaghan (2011) argues

[131] Ibid., p. 61.

that the best way to account for the relation between sound and its source is *mereological* and not causal. What he suggests there is that sound is a *part* of its source:

> [T]he audible relations between sounds and sources are mereological. Sounds audibly are constituent *parts* of everyday audible events, such as collisions and vibrations, which involve material bodies.[132]

O'Callaghan's main reason for appealing to such a mereological relation is to explain how everyday happenings including material bodies are audible when we hear sounds in our surroundings, to explain what I called the Immediacy Observation (IO). The suggested part-whole relationship can explain (IO) because there is a strong sense in which by perceiving a part of something we directly perceive the thing itself—for example, when I only see the front cover of a book, there is a strong sense in which I see the book itself, the object including that front surface.[133]

So, considering this mereological idea of how a sound and its source are related, let me formulate O'Callaghan's event view thus. Everyday happenings such as the collision of two material objects are complex events comprising different parts. A part of such complex events is the very event causing a material body or bodies to vibrate, for example the event of collision. O'Callaghan does not think this event is sound. The other part is another event type which is the result of, say, a collision, that is, the *vibration* of a material body or bodies. Again, O'Callaghan does not think sounds are such events. However, a further part of everyday complex happenings is vibrating material bodies' *disturbing* a medium like air or water. As discussed earlier, this disturbing (of the medium) is what O'Callaghan identifies with sound. What is important here is

[132] O'Callaghan (2011), p. 395.
[133] As Jackson (1977) has argued, we perceive a whole in virtue of perceiving its parts. The idea that wholes are (indirectly) perceived in virtue of their parts has been criticized by Mohan Matthen (2010). According to Matthen, we perceive wholes directly, and sometimes we perceive parts of a whole in virtue of perceiving the whole.

that based on his mereological proposal we should understand such a disturbance event (sound) as a *part* of the source of the sound. That is, in my understanding, the collision of two material objects (as an example) *plus* the vibration of those objects constitutes the source of a sound, and sound is a part of this source.

I want to begin my discussion of this theory with a phenomenological concern. Let's suppose that a sound's relation to its source is a part-whole relation. The issue is when O'Callaghan suggests that this mereological relation between a sound and its source is *audible*. For example, in the passage quoted above he says: "Sounds *audibly* are constituent *parts* of everyday audible events […]" (the first emphasis added). I should emphasize that the concern is not that a sound cannot metaphysically be a part of its source. This might be the case *metaphysically*. What concerns me is that O'Callaghan seems to try to back up this metaphysical claim by the phenomenology of hearing, by saying that we *hear* a sound to be a part of its source. I think this specific claim is not phenomenologically convincing for the following reason.

How can we make sense of the idea that we *hear* a sound to be a *part* of its source? Perhaps vision can give us some clues here. Suppose you see a book and its different parts, for example its front cover, back cover, and the spine. You also see, for example, the spine to be a *part* of the book. That is, we can plausibly say that you see a part-whole relation between the spine (a part) and the book (the whole). To me, the reason that one can see such a part-whole relation is because one can in principle see more than one of the book's parts in relation to each other comprising the book. If one could not see some other (non-spine) parts of the book and their relations to each other, she would be unable to *see* that the spine is a part of the book. To clarify more, imagine a case in which you can *only* see the front cover of a book—you cannot see any other parts. Perhaps the book is behind a big wooden panel, and through a rectangular

hole at the center of the panel you can see only the book's front cover. Can we say that in such a situation you can *see* the front cover as being a part of the book? I think not. If you do not have any visual access to any other parts of the book, it appears you can *only* see the front cover. It can of course plausibly be said that by some background knowledge—for example, by knowing how book covers typically look and recognizing that what you see is a book cover—you can *infer* that what you see is a part of a book. But this is pure inference and not visual perception. My point is that if in such a case you do not have any visual access to any other parts of the book and how they are related to each other and the front cover, it would be unreasonable to say that you could *visually perceive* that the front cover is a part of the book.

Let me generalize this observation about the phenomenology of visual mereology. Suppose P is (metaphysically) a part of W. Now, in order for one to be able to see P's being a part of W, that is, in order for one to be able to see a mereological relation between P and W, one must be able to see some other (non-P) parts of W and see their relations to each other and to P.[134] Otherwise, we cannot say that one can *see* any mereological relation between P and W. It is worth noticing that I am not saying that if the above condition did not hold then by seeing P one would not see *W*. To me it is a plausible idea that in virtue of seeing a part of something one sees the thing itself. However, the point I tried to show is that if the above condition did not hold, one would not be able to see *a part-whole relationship* between P and W.

Does hearing meet the requirement that we have gotten at by reflecting on visual mereology? I think it does not, because in my opinion in hearing we do not hear anything but

[134] I am using the term "see" in its perceptual sense here and not in the sense used in a sentence like "I see how hard this problem is." Also, by seeing a relation between P1 and P2, which are parts of a whole W, I mean seeing P1 and P2 being somehow attached to each other and making a whole (for example, the way in which we see two legs of a table being somehow connected in the whole table).

sounds or, to put it differently, in hearing anything we hear is heard *in* sound. If by changing our perspective we are able to see different parts of, say, a book or a car, as a result of which we can visually perceive that, for example, a wheel is a part of a car or a spine is a part of a book, we cannot do such things in the case of hearing. Suppose you hear a sound the source of which is far from and unrecognizable to you. But when you get closer to the source, it becomes clear to you that what you hear is a violin being bowed. Here, it makes perfect sense to say that by getting closer you clearly hear *the source*, or, put it more precisely, by getting closer you can auditorily recognize that the sound you hear is the sound of a violin being bowed. However, this does not show that when you get closer you hear the source as something distinct or different from the sound: you hear the source *in* the sound. In other words, by getting closer you hear the sound more clearly, as a result of which you can auditorily recognize its source better; but this does not mean that by changing your place you can hear any non-sound thing (whatever that means!). This is unlike a visual case in which you can, for instance, only see the front of a car but by changing your perspective or place come to see the car's other parts. This is why I think the phenomenology of hearing does not support the idea that we hear sounds as parts of their sources.

I should, though, clarify that my argument is not that we cannot hear a sound as part of another sound. That is, I am not denying that we can hear complex sounds and that sometimes changing our (auditory) perspective enables us to recognize that a sound we hear is in fact a complex sound. For example, suppose you hear the sound of a guitar from a distance, and you think it is one sound. You get closer and you hear the same sound again and recognize that what you hear (and heard moments ago) is in fact composed of two sounds generated by two of the guitar strings plucked together. In such a case you can hear relations between two or more

sounds. But this is not what I argued against. My argument above was for undermining the idea that we hear mereological relations between sounds and their *sources*. This in no way rules out the plausible idea that there can be mereological relations between sounds, that we can hear complex sounds or sonic structures.

So, I tried to show that O'Callaghan's idea that there is an *audible* part-whole relationship between sound and its source is not phenomenologically convincing. My main consideration against it was based on the phenomenological observation (and the related requirement) I gave for visual mereology (*seeing* part-whole relationships between objects and their parts).

But it might be objected that the phenomenology of visual and auditory perceptions can be different in various respects. Accordingly, the objection goes, it is implausible to let the phenomenology of visual perception guide us in exploring whether we can hear sounds as parts of their sources.

I concede to this objection to some extent. I agree that we should not assume that the phenomenology of hearing and that of seeing are necessarily similar in any respect, and mereology could be one of the respects in which they differ. Yet, I simply do not know how we can phenomenologically explore whether a sound is a part of its source without appealing to, or at least getting some clues from, visual experiences of parts and wholes. That is, I do not know how to make sense of the idea that we *hear* sounds as parts of their sources only in auditory terms. It does not pretheoretically appear to me that when I hear a sound, I hear it as a part of its source. If a theory claims that such a part-whole relation is *audible*, it should provide some supporting phenomenological considerations. I considered how similar visual experiences—when we see something as a part of something else—could be clarificatory. But my analysis

went in the opposite direction; it showed that we do not seem to hear sounds as parts of their sources.

However, I think it is time now to set aside such phenomenological analyses. For even if we had convincing phenomenological considerations to doubt that we *hear* sounds as *parts* of their sources, it might be the case that they *are* parts of their sources. That is, it might be the case that our phenomenology of hearing should not be trusted here: although we do not seem to hear sounds as parts of their sources, they might, *metaphysically*, be parts of the objects or events to which they are related. In other words, perhaps O'Callaghan is right in a metaphysical, though not a phenomenological, sense. Let's examine this.

An issue with O'Callaghan's theory is that in explicating his ontological view on sound, specifically in O'Callaghan (2007), he mostly uses causal language to describe a sound's relation to its source; that is, in the context of what a sound *is*, he construes sounds as being generated or caused by their sources. However, in the context of examining how sounds and their sources are related, specifically in O'Callaghan (2011), he suggests that the relation is mereological. The concern is whether using these two languages affects the metaphysical coherency of O'Callaghan's view. The problem is about the nature of the relation between a sound and its source in the context of that theory. Is the intended relation causal, or mereological, or both? In one place he uses mereological language to spell out that relation—particularly as a response to and as a rejection of the view that sound is an entity caused by sounding objects—and in another place he uses causal language and construes sound as a disturbance event (the event of disturbing the medium) *caused* by other events such as striking or collision of two objects. Using these two languages seems to cause a tension in O'Callaghan's general view on sound. I am going to

suggest two ways in which this concern can be responded to in the context of O'Callaghan's theory—although I will argue that neither is convincing.

It can be argued that the tension described above is only apparent because the notion of causality is meant in two different ways in how O'Callaghan construes sound in relation to its source. What O'Callaghan objects to in rebutting the causal view on sound is the idea that sound is an *object* caused by its source. That is, he denies that the relation between a sound and its source is a causal relation between a cause (a sound source) and a distinct object—a sound with its own qualities. However, the event view he suggests is that sound is an *event* (a disturbance event) caused by another event (a happening like the collision of two objects), and these two events are collocated and intimately connected. To clarify this reading of O'Callaghan's view, let's consider two examples for each kind of causal relationship mentioned.

Suppose we have a machine that produces wooden cubes with specific colors. Imagine you see a red cube made by this machine. Whether you know that the cube you see was produced by the machine or not, what you see is a distinct object with its own qualities such as size, color, texture, and so on. Although the cube was produced by the machine, or the latter has caused the former to exist, the cube is an independent object of perception that after coming to existence does not have any intimate connection with the machine that produced it. They do not need to be collocated, for example, or share any of their qualities.

But now, let us consider a more intimate causal connection, say, between a tooth and the body in which it resides. The relation is causal because teeth form as a result of processes in the bodies in which they reside. In this respect, this relation is like the causal relation between the machine and the cubes it produces. But what to some extent distinguishes the tooth-body causal relation from the other relationship is that the former is an intimate relationship which seems to

be both causal and mereological. That is, a tooth both is caused by some biological processes in the body in which it resides and is a *part* of that body. However, we do not have this intimate, dual relationship in the case of the machine and the cubes it produces. This latter relationship is primarily causal and not mereological—the cubes are not parts of the machine that caused them.

It should be noticed that the above distinction between a primarily causal relation (like the case of the machine and the cubes it makes) and a causal-mereological relation (like the tooth example) might have implications for perception. Perceiving an effect is not necessarily perceiving the cause when the relation between the two is merely causal. For example, in the machine-cube example, seeing a cube made by the machine is not seeing the machine itself; the two do not necessarily share their location and their qualities (when, for example, you see one such cube in your house while the machine is in a factory miles away). However, because a tooth is a part of the mouth and of the body in which it resides, perceiving the tooth of a person gives us stronger perceptual access to the person: through perceiving the tooth we can perceive *a part* of the person's body, and, at least, their location.

To recap our discussion above, it seems that O'Callaghan construes the relation between a sound and its source as not merely causal but both causal and mereological: a sound source as an event, like the collision of two objects, causes another event to happen, that is, it causes the disturbing of the medium (air, water, etc.). And this disturbance event, to O'Callaghan, is also a part of that former event. In other words, the former event is a complex event comprising (*and* causing) the event of disturbing.[135] I tried to spell out this causal-mereological relation by the example of a tooth and body of which it is a part. Some objects/parts seem to have causal-

[135] It should be noted that what includes the disturbance event is not merely, for example, the collision of two objects but a complex entity including the collision and two objects. In my understanding of his view, it is this complex entity that O'Callaghan thinks also includes the event of disturbing the medium.

mereological relations with their wholes. And if this is the case, the relation between a sound and its source might be of this type—at least this is what O'Callaghan claims, to my understanding of his view.

However, I think there is still a concern with this proposal. This construal of course tells us in which sense hearing a sound is hearing its source: that is, it gives an explanation for the Immediacy Observation (IO) via the proposed mereological relation between sound and its source. (As discussed earlier, perceiving a part of a whole is in a sense perceiving the whole itself.) However, the explanation provided for IO is to some extent unconvincing.

It is unconvincing because we still need to know how exactly a sound is a part of its source in order to understand better how perceiving a sound can be perceiving its source. I see two possibilities. First, a sound could be a distinct object with its own distinct properties that is merely collocated with its source and is a part of it. Such a (mereological) relation is like the relationship between a mirror and the car it is, say, an exterior mirror of. Second, a sound could be a portion of its whole, like the relationship between a portion of a wall and the whole wall. The former alternative is, as I will argue below, a weaker sense and the latter a stronger sense for how perceiving a sound can be the perception of its source.

According to the first alternative, sound is itself a distinct entity which happens also to be a part of an aggregate whole, a complex event. The relationship between different parts of a car to the car itself can be an example of such a relationship between a part and its whole. Wheels, seats, or interior and exterior mirrors of a car are stand-alone objects; we can perceive them independently as distinct, independent objects. However, on some or perhaps most occasions we see them as being connected to each other and forming an aggregate object, a car. I do not intend to explore the metaphysical nature of such mereological relations. What is important for my

purposes here is a specific characteristic of them: namely, when something is a part of something else according to this alternative, the part is itself a distinct entity having *its own distinct properties* that can be perceived independently of its whole (such as the aggregate parts of a car as mentioned). Now, this characteristic has the following implication. If X is a stand-alone object (like a wheel) and on some occasions it also belongs to an aggregate whole Y (like a car), the very perception of X does not necessarily imply the perception of Y. Suppose in a film you see a close shot of a car wheel without seeing any other things: you *only* see a wheel (or a major portion of it). Can we say that in this case you were seeing the car? I think if we wanted to answer carefully, we should say it depends on the context in which the wheel is located and being perceived. If the camera zooms back and you see the bigger scene, you can then decide whether the wheel you were seeing was one of a car. In other words, you would need to see if the wheel is *connected to* any other parts of the car in order to say if your perception of the single wheel is also the perception of a car; otherwise, it might be the case that you are seeing only a wheel standing alone on a street (it can happen!). What I am trying to say is that nothing about a car's wheel necessitates that the mere perception of it, when you do not have the bigger picture, is ipso facto the perception of the car.

A second alternative for the part-whole relationship, which has a different characteristic than the one considered above, is as follows. This time imagine a big wall. When I can only see a portion of the wall, I see in fact a part of the wall (the whole), and like the earlier case of a car and its exterior mirror, the perception of a part/portion of the wall is also the perception of the wall/whole. But there is a difference in the latter case: a part of a particular wall typically instantiates many properties that are specific to that wall as an object (for example, the material the wall is made of, its texture, color and so on and so forth). Put differently, a part of a wall is a

mostly reliable *sample* of its whole because it instantiates several properties that are specific to its whole. However, this is not the case for an exterior mirror and a car: arguably, a mirror, which can be used as an exterior mirror in a car, is not a reliable sample of the car, a sample that instantiates the main properties of its whole.

It should be emphasized that contrasting the two above cases is not meant to oversimply the mereological relationships between parts and wholes and all the relevant metaphysical nuances. For example, it can reasonably be argued that the comparison oversimplifies the situation because it ignores the fact that a car is a complex object including lots of parts while a wall is, in comparison, a mereologically simpler object. Such considerations are important in examining the mereological aspects of the example cases I have given; however, they are beyond the main point I want to convey via the comparison. What I wanted to show is that there are two extremes when we regard the relation between a part and its whole: sometimes a part (like a part/portion of a wall) represents its whole by instantiating some or several of the whole wall's main properties; and sometimes a part is a stand-alone object (like a mirror) that can, on some occasions, be a part of a complex whole (a car, for instance). An (intended) distinction between the two is that in the former case a part can be seen as a *sample* of some of the main properties of its whole; however, in the latter there is no such a sampling role taken by a part.

The above distinction is particularly important from the perspective of *perceiving* parts and wholes. When a part plays a sampling role (the former case), by perceiving its properties one also perceives some of the properties of the whole—by perceiving a part/portion of a wall, one perceives some of the properties of the wall itself. However, when there is no such a sampling role taken by a part, perceiving the part only delivers to the perceiver the distinct properties of

that part itself (or, in other words, it might or might not deliver any main properties of the related whole).

Let's return now to the case of sounds and O'Callaghan's theory. To him, a sound is a part of a complex happening (sound source) and it has its own properties (pitch, loudness, and timbre). That is, timbre is not, according to him, a property of the source the sound is a part of but a property of the sound itself. And the same for loudness and pitch. Construing a sound as a distinct entity (a disturbance event) that possesses its own distinct properties and is a part of its source is very close to the second (non-sampling) mereological case discussed above, that is, sound, to O'Callaghan, does not possess any properties of its source (the whole). But now a problem for this theory is to explain how hearing a sound is *hearing* its source, given that hearing a sound delivers to us only the properties of the sound itself.

Suppose we hear a sound and do not have any prior (visual) experience or background knowledge of what object(s) and event(s) constitute its source. Based on O'Callaghan's theory, in such a case what is perceptually/auditorily presented to us are the distinct properties of the sound itself. In other words, the properties presented to us basically carry information about the sound itself and not about the source (especially given that in this case there is no relevant background knowledge or prior visual experience on the side of the perceiver). The only information about the source that is claimed to be carried and directly presented to the perceiver by the sound (in our case) is the source's *location*, because the sound, as claimed, is a part of its source and collocated with it. (Recall that the view is that sound is a part of a complex happening consisting of different parts—objects and events—and one event-part in this complex happening is the disturbance of a medium, like air, caused by another event-part such as the collision of two objects.)

And now it is time to see a main worry about this theory. The mereology this theory provides with regard to sound and its source is akin to the non-sampling case discussed earlier (the car case). This suggests that when one hears only a sound without having any kind of background knowledge about the relevant source, they would not have any *auditory* access to the source except for its location. The reason is that according to this theory the properties being auditorily perceived by a subject while hearing a sound are specific to that sound as a stand-alone entity and, as a corollary, by hearing *only* the sound the perceiver would not be presented with any property of other parts of the complex happening the sound is a part of. So, when one hears only a sound (without knowing anything about the source) one cannot be said to be able to *hear* the source—they can only hear the sound.

But I believe an unconvincing implication of the above conclusion is that human audition per se (without the help of some form of background knowledge) is mostly disconnected from the world of objects that seem to be an integral part of sonic happenings. For according to this view by hearing (only) sounds we have auditory access only to their specific properties and the location of the sonic happenings of which sounds are part, and nothing more. That is, sound is considered not to auditorily reveal any property of its source to hearer (except for its location). Hearing (approximately) how big the source of a sound is, what type of mechanical procedure (say, collision, rubbing, scratching) constitutes the relevant sonic happening, or what material constitutes the relevant object (wood, metal, etc.) does not seem to be able to be explained by this theory. It might be argued that all such information can be multimodally revealed to a hearer, and O'Callaghan in fact briefly suggests this in the context of his theory.[136] The idea is that when a hearer has some form of background knowledge about the source of a sound, they are then able

[136] O'Callaghan (2011), p. 400-401.

to also hear, say, the relative size of the source by hearing only the sound it makes. For example, if I have seen several times how the double bass and violin are played and make sounds, when I hear only their sounds, I can probably recognize (auditorily) the instruments and hear their relative sizes. So far, so good; we have an explanation for how hearing can reveal some information about a sound source to the hearer. However, it should be noticed that in the core of the proposal is that audition has the mentioned capacity *when* it is infiltrated by a form of background knowledge or relevant prior visual experience. The idea is not that audition itself has that capacity to reveal information about some properties of sound sources, or, the idea is not that *in* sounds we receive such information. It is suggested that such information is given basically multimodally (or inferentially) to a hearer and not purely auditorily.

The issue is not about the coherency of this suggestion: it might coherently be possible that the intended properties of sound sources are delivered to a hearer multimodally or inferentially and not directly *in* sounds. However, we are not given any justification why it should not be the contrary: that the properties we auditorily perceive in sounds do belong to sound sources, to events *and* objects involved in sonic happenings. This prospect is corroborated by empirical evidence showing that hearing reveals to the perceiver properties of the involved objects and sonic events (source-object and source-events), properties like the material source-objects are made of (whether they are, say, strings or wooden bodies) or the excitation mechanism, how an object is excited to make sounds. If, as Giordano and McAdams (2010) have discussed, acoustic cues in sounds are perceptual correlates of such properties—properties that belong to sound sources—we have stronger evidence for thinking of acoustic properties heard in sounds as properties of their sources rather than as distinct properties of sounds, which are

construed, in O'Callaghan's theory, as independent entities mereologically related to their sources.[137]

But if the empirical evidence suggests this and we seem to be justified in considering sounds in terms of *properties* of their sources in one way or another, we still need to spell out in which way exactly because: (1) we need to explain how timbre, loudness, and pitch are properties of sound sources, and (2) it cannot be dismissed that sound sources seem to have both object and event components, raising the question: sounds are properties of what exactly? And this will be the business of the following section.

4

In this section, I am going to lay the foundation for the idea that in hearing sound we perceive properties of an *object partaking in an event*. The intended event is the vibration of an object (a piano, an animal's vocal tract, a glass, etc.) when it is stimulated somehow (for example, by being hit). That is, the vibration happens when an object is stimulated by an external force, whether that force is another material object (like when a china cup is banged by a metal spoon) or an external force like air pressure or heat (like when a piano's string breaks, because of, say, a sudden change in temperature, and the instrument makes a noise). As we saw in O'Callaghan's theory, sonic events seem to have a complex nature being constituted, at least in part, by the two mentioned events: vibration and excitation. I will argue that in hearing a sound

[137] It seems O'Callaghan's main justification for considering sounds as independent entities is because we (can) treat them as such in musical experiences: in musical experience we attend to sounds themselves as independent objects of perception. However, as will be discussed in the last section, treating sounds thus in musical experience, which itself is controversial, seems to be the result of *aesthetic* attention and not related to the nature of sounds or ecological hearing. (I use "ecological hearing" in the sense of hearing sounds in our surroundings in contrast with hearing sounds in an aesthetic experience such as listening to music.)]

we perceive properties of an object participating in a sonic event. But before getting to defending this idea, let us reflect a little on the nature of events themselves.

Two questions about events will be relevant to our subsequent arguments: (1) What is an event? (2) What is the perception of an event? Approaching either of these questions in depth will be beyond the scope of this paper. I will only consider the implications of these two questions for my view on sound, starting with the first question.

Two major philosophical views on the ontological nature of events are as follows. (1) Events are *individuals* like any other objects. (2) Events are, in Moltmann's terms, "derived objects, objects entirely dependent on individuals, properties and times."[138] A main defender of the first view was Donald Davidson, who believed events are individuals that are referents of event terms. For example, in the sentence "Nina ran yesterday morning," according to Davidson, the event of running is itself an individual entity, on par with Nina, that is being referred to by the term "ran." In other words, setting aside the specific time involved, in this sentence, two entities are being referred to: a person and an event. On the other hand, according to (2), events are not individuals but complex *structures involving* other individuals, their properties, and specific times. For example, Kim has suggested that an event is the exemplification of a property (P) by an entity/individual (x) at a specific time (t). For instance, in the above case, that specific event of running has these constituents: Nina (x), the property of running (P), which is a generic event, and a specific time t (a point of time one day before the utterance of the sentence). Moltmann, in another version of this view, has suggested that an event has this structure: a

[138] Moltmann (2002), p. 189

change of properties possessed by an individual *a* during a specific time period (*a*'s having property P at t to *a*'s having property Q at t').

I intend neither to delve into the metaphysical details of these two viewpoints on the nature of events nor to adjudicate between them. What I want to assume in my following arguments is that an event, whether it is an individual or derived object, has its own properties. I believe this assumption is in harmony with both the views adumbrated above. The intended properties are those that can be explained only by referring to an event either as an individual or as a derived object. For instance, as regards our running example, the speed associated with Nina's running is such a property. It is not hard to see that if the event of running is an individual, the speed is that individual's property. That is, if we wanted to construe Nina and her running as two distinct individuals, the relevant speed seems to be a property of the latter and not the former individual (Nina *ran at a speed of*, say, 6 miles per hour). On the other hand, if the event of running is a derived object, for example, the exemplification of a property by an individual at a time—as Kim suggests—or a change of properties possessed by an object during a time period—as Moltmann says—the relevant speed seems to be a property of this derived object—the exemplification or the process of change. Nina herself, standing alone, does not have any property of speed, but it is when she runs that we can talk about that property; in other words, speed is a property of *Nina's-running-in-a-specific-time-period* considered somehow as a structure/derived object (in Kim's, Moltmann's, or another way) and not Nina herself. The point is that whether events ontologically are individuals or derived objects, it is a plausible assumption that they have properties of their own distinct from those possessed by the entities or individuals they happen to.

This assumption can also be corroborated by the considerations relevant to the perception of events, particularly how events are perceptually recognized. In perceiving events, we are able to recognize different instances of the same event (or the same type of event). For example, we can recognize the event of running while seeing an adult, a child, or even an animal running. Some have applied James J. Gibson's idea of *transformational invariants* to explain this phenomenon of event recognition.[139] The idea relates to the fact that an object can perceptually be recognized in different situations in which it might undergo some changes. For example, we can recognize the same human face, say, the face of our friend, although the face takes on different expressions or changes due to aging. Gibson has argued that the reason for this is that we can recognize some constant features in objects, what he calls transformational invariants, besides the variant variable features they might present in different situations. Some people have used this idea about object recognition to explain event recognition. As Shipley (2008) has articulated:

> Robert Shaw […] had the critical insight that the same logic used for object perception could be applied to event perception; there is some property, a *transformational invariant*, that is present for certain transformations and allows us to identify the transformation even when it is applied to different, novel objects […].[140]

In line with this account, I will employ the following assumption in my argument:

> [PoE] An event has properties of its own, distinct from the properties of the entity or entities it happens to.

Based on [PoE], I want now to argue that a sound is *a complex property* that includes properties of the object and properties of the event involved in making it. An upshot of this view

[139] See Shipley (2008).
[140] p. 17

is that when we hear a sound, we hear an instantiation of a complex property possessed by the relevant object and event. Here, we can distinguish between a sound-type, which is a complex property, and a sound-token, which is the instantiation of a complex property.[141] So, by "hearing a sound [on a particular occasion]," I mean hearing a sound-token that is, according to my proposal, an instantiation of a (complex) property. For example, in hearing the sound of a metal bar when it is struck and vibrates, we hear a sound-token or the instantiation of a complex property that includes some properties of that metal bar and some properties of the event causing it to vibrate. Now, let me elucidate how I will use the term 'complex property.'

The notion of a complex property has notably been used by David Armstrong in his theory of universals.[142] He defines a complex property based on the conjunction of some other properties: "Given that F and G are distinct universals, then F&G can be a universal, provided always that a particular exists at some time which is both F and G."[143] Armstrong calls the condition used in this definition the *co-instantiation condition*. An example of a complex property can be a ball's being both round and red. In this case, *redness-and-roundness* counts as a complex property in such a way that the ball is red-and-round. Complex properties such as red-and-round are, in fact, mere conjunctions of other simpler properties; we can call them *conjunctive* complex properties.[144] However, there is also another type of complex property called *structural property*.[145] A structural property is one that requires that parts of the object

[141] Or, if properties are tropes, the sound heard on a particular occasion can be considered a trope with a complex structure possessed by the relevant object and event.
[142] The term "universal" means property. As a realist about properties, Armstrong believes that the exact same property/universal such as redness can be possessed by different particulars (for example, a ball and an apple can both possess the same property of redness).
[143] Armstrong (1997), p. 31.
[144] See Fisher (2017).
[145] See Armstrong (1986), and Fisher (2017). Lewis (1986) critically examined construing such properties as properties over and above their constituents.

instantiate various properties that are not co-instantiated by the object as a whole. For example, a molecule of methane is constituted of one hydrogen and four carbons. That is, one hydrogen and four carbons in specific structural relations to each other constitute a methane molecule. However, although a methane molecule instantiates the property of *being methane*, it is neither hydrogen nor carbon (or instantiates neither the property of being hydrogen nor that of being carbon). Complex properties such as being methane—which has the described characteristic—have been called *structural* properties.

Now, by saying that sound is a complex property, I mean it includes some other properties, such as pitch, timbre, and loudness. Let's show the property of having a particular pitch as P, that of having a particular timbre as T, and that of having a particular loudness as L. Sound, in my view, is then a conjunction like P&T&L. However, although I am more inclined to consider the complex property of sound as of the conjunctive type, I do not want to make any strong claim on this matter. To ground the core thesis, let me focus only on the idea that sound is a complex property in the form of P&T&L; whether the conjuncts are in structural relations to each other is something that can be investigated after basing the core idea, which is beyond the scope of my arguments here.[146] But an important initial important initial matter we need to clarify about this complex property is related to co-instantiation. Based on Armstrong's definition given earlier, the conjuncts in (or constituents of) a complex property should be co-instantiated by a particular. What particular instantiates sound?

I suggest that what instantiates sound is *an object and an event in which it is participating*. Depending on our metaphysical view about the nature of events, my suggestion

[146] By bringing up these two alternatives for a complex property I meant to suggest that there could be two versions of the view I am developing about the nature of sounds. However, I do not intend to adjudicate between these two versions in this paper.

can be formulated in different ways. You will recall that we considered two main metaphysical viewpoints on the nature of an event: an event as a derived object or as an individual. If an event is a derived object including other individuals, properties, and times, then what instantiates sound is, based on my suggestion, such a derived object, specifically the individual constituting it and the derived object as a whole itself. That is, if a sonic event is a derived object, then the sound as a complex property is partly instantiated by the individual (such as a metal bar) that constitutes that derived object and partly instantiated by the derived object as a whole (*the-metal-bar-being-struck-at-time-t*). On the other hand, if a sonic event is an individual, then the sound as a complex property is partly instantiated by that event and partly by the individual (entity) participating in it. Regardless of how exactly my proposal should be formulated based on what metaphysical view we take on the nature of events, the core of my proposal is that sound as a complex property is instantiated in part by an object, which becomes vibrated in a sonic event, and in part by the event that happens to that object and causes its vibration. (The assumption [PoE] is the basis of my latter claim—an event's having its own properties.)

It should also be noticed that there is a peculiar characteristic about the two instantiating bases of a sound. Even if we endorse the metaphysical view that an event is an individual ontologically on par with other individuals, such as ordinary objects, events are *not detached* from objects/substances in the world. As Davidson has put:

> [M]ost events are understood as changes in a more or less permanent object or substance. It even seems likely to me that the concept of an event depends in every case on the idea of a change in a substance […].[147]

He also thinks that this dependence is symmetrical:

[147] Davidson (2001) pp. 173-4

> Substances owe their special importance in the enterprise of identification to the fact that they survive through time. But the idea of survival is inseparable from the idea of surviving certain sorts of change—of position, size, shape, colour, and so forth. As we might expect, events often play an essential role in identifying a substance.[148]

In Casati and Varzi's terms:

> Objects are prime actors in events; objectless events are uncommon. But so are eventless objects; events make up the lives of objects.

My point is that by saying that a sound is a complex property whose constituent properties are instantiated by objects and events, that what instantiates a sound is an object and an event that is occurring to that object, I am not attributing that complex property to an arbitrary combination of two distinct things, even if events are themselves individuals. The two instantiating bases of a sound—an object and an event it participates in—are intimately related. What bolsters this intimate relation in the case of sounds is that an ambient sound is always connected with both objects and events: an object makes a sound when it vibrates due to participating in an event.

The idea I have been arguing for so far is that sound is a complex property in the form of P&T&L whose conjuncts are instantiated in part by an object and in part by an event the object participates in. But now we need to answer two more urgent questions: (1) In which sense does an object possess one or more properties constituting the complex property of sound? (2) Which constituent properties of the complex property of sound are instantiated by the relevant object, and which by the relevant event?

Regarding question (1), consider the property of *agility* we attribute to some animals such as a cheetah. It seems clear that the property of agility can be possessed by an animal alongside

[148] Ibid., p. 175

other properties such as weight, body shape, height, etc. However, agility has a specific characteristic compared with properties like size or shape. Agility is mostly a potentiality in an animal, an ability that can reveal itself on some occasions, particularly when the animal *moves* or *participates in events* such as hunting or running. However, a cheetah does not have to move or participate in any event to have a particular weight, size, or body shape. It seems these properties are always "actualized" in the animal while agility is not. If agility is in the animal or possessed by it as a property, it is so as a potentiality. Specifically, what is important for my purposes is that the realization of this potentiality is tied to some events; for example, a cheetah's agility is actualized when the animal runs fast. Perhaps this is why the concept of agility is tied to an event concept, that of movement. According to the Cambridge Dictionary, the term "agile" means: "able to *move* your body quickly and easily" (emphasis mine). Buoyancy is another example of such event-dependent properties, defined as "the tendency of a body to *float* or to *rise* when *submerged in* a fluid" (emphasis mine).[149] Like agility, buoyancy is a potentiality in an object which gets actualized when the object participates in a relevant event, say, in the event of being submerged in water.

Thus, there seems to be a distinction between those properties that are essentially related to *how a substance behaves in some events* and those properties like size or density that are not. What density, size, or shape a body has is not primarily about how it behaves in this or that event. But the agility of an animal or whether a substance is buoyant is about how it behaves in the relevant events such as running, in the case of agility, and being submerged in a fluid like water, in the case of buoyancy.[150]

[149] Merriam-Webster Dictionary, https://www.merriam-webster.com/dictionary/buoyancy
[150] The distinction I have drawn is not the same as Lockean distinction between primary and secondary qualities. Locke considers secondary qualities as "powers" in the substances. But by saying that properties relating to the

Using this distinction, I suggest that among the properties being included in the complex property of sound, the properties that are instantiated by an object are those properties of that substance that determine how it distinctly behaves or participates in a sonic happening (striking, collision, etc.). A wooden object behaves differently when it gets hit and vibrated than a metal one. Knocking on a wooden door makes a different sound than knocking on a metal one. A metal bar makes a different sound both from a wooden bar of the same size and from a metal door even when all the objects are struck in the same way. I suggest that one or more properties being included in a sound (as a complex property) are instantiated by and represent the uniqueness of the substance making the sound.

I earlier considered a sound as a complex property in the form of P&T&L, as having pitch (P), timbre (T), and loudness (L) as its main constituents. I do not want to make a strong and exact claim about which one of these three constituents is instantiated by the substance and which one by the event it participates in. I think this matter should primarily be investigated by empirical science. However, the empirical studies on timbre I discussed earlier[151] suggest that timbre, or at least a part of it, is instantiated by the substance making a sound. These studies have shown how timbre is directly correlated to features of a substance such as its material, shape, and structure, besides being correlated to how the substance is excited to make sound.[152] Let us

behavior of a substance are potentialities in it I do not mean they are secondary qualities. How a substance behaves in different events could be a part of the substance's primary qualities. Moreover, reflectance physicalism about colors considers, in my understanding, colors as a property relating to how a substance's surface *reflects* light when it is illuminated. It seems to me that construing colors in this way can have similarities with what I have said about properties like agility: much as color is related to how a substance participates in the event of reflecting light when it is illuminated, agility is related to how a substance/animal participates, say, in the event of running or hunting.

[151] Giordano, B. L., & McAdams, S. (2010), Handel, S. (1995), McAdams, S. (2013).

[152] This is related to what has been called *timbre constancy*, the fact that timbre bears in one way or another the identity of its source (a clarinet or trumpet for example). (See Dowling (2010), p. 245). As McAdams and Giordano (2016) have put it: "[timbre] is one of the primary perceptual vehicles for the recognition, identification, and tracking over time of a sound source (singer's voice, clarinet, set of carillon bells) […]." (p. 114)

consider two examples. Striking a metal bar and striking a wooden one lead to different sounds. These sounds could have the same volume (and possibly the same pitch), but they would still differ in respect of their timbral qualities. Or, as a musical example, the string instruments oud and guitar (see Figure 1) can be struck in the same way (by plucking their strings) and make sounds of the same volume and pitch, but all the sounds produced by each instrument would still have a distinct quality in terms of timbre. Let me clarify this point a little more with some preliminary concepts related to the physics of sound.

When an object vibrates and makes a sound, its vibration causes waveform motions in the medium (in the air, for example). This waveform motion is typically called and depicted as a *sound spectrum.* Figure 2 shows the sound spectrum related to the G note played by the flute. Each sound spectrum has different wave components and a particular shape. A periodic wave is one that repeats itself periodically and has a particular frequency and amplitude (Figure 3). Objects such as musical instruments whose sounds are heard to have specific pitches cause periodic wave components in their sound spectra. Figure 4, the top diagram, shows the periodic components (waves) of a sound spectrum depicted in the bottom diagram. Generally speaking, the pitch and loudness of sounds are considered to be the perceptual correlates of, respectively, the fundamental frequency and amplitude (intensity) of a sound spectrum, respectively. (In Figure 2, the pitch and loudness of the sound are mainly correlated to the frequency and amplitude/height of the red wave.) The main idea is that sounds that differ in their spectra (particularly in terms of the frequency and amplitude features of their spectra) have different pitch and loudness attributes. However, the situation is different with timbre: two sounds with the same sound spectra can be heard presenting different timbral qualities. As explained by Wolf et al.:

> Introductory physics text books sometimes give the impression that the spectrum is the dominant contribution to the timbre of an instrument, and that certain spectra are characteristic of particular instruments. […] [T]his is very misleading. Some very general or vague comments can be made about the spectra of different instruments, but it is not possible to look at a harmonic spectrum and say what instrument it comes from. Further, it is quite possible for similar spectra to be produced by instruments that don't sound very similar. For instance, if one were to take a note played by a violin and filter it so that its spectrum were identical to a given spectrum for a trumpet playing the same note, the filtered violin note would still sound like a violin, not like a trumpet.[153]

To recap, the wave motions the vibration of a substance causes in a medium like air, which are captured theoretically as a sound spectrum, have, inter alia, frequency and amplitude features. The pitch and loudness of a sound (a musical sound, for example) are considered mainly to be the perceptual correlates of those two features. However, two identical sound spectra, which, in most cases, give the same pitch and loudness, can perceptually present two distinct timbral qualities. (The violin-trumpet example in the above quotation and our metal/wooden bars and oud/guitar cases earlier are examples of this.) This suggests that timbre cannot be wholly captured by the features or the shape of a sound spectrum; that is, it cannot be captured wholly by the features of the vibration patterns of an object.

Figure 1. The oud (left) and the guitar (right) with distinct sounds, particularly in terms of timbre

[153] Wolfe, J., Smith, J., Tann, J. and Fletcher, N.H. (2001) "Acoustics of classical and modern flutes: a sound spectra, sounds and fingerings". (The quotation is in this subpage: http://www.phys.unsw.edu.au/jw/sound.spectrum.html)

The above conclusion is bolstered by the empirical studies I mentioned earlier showing that timbre is, in fact, a perceptual correlate of features of an object *itself,* such as its shape, material, structure, etc. Among the conjuncts in P&T&L (the complex property of sound), timbre (T) is the property of the substance that is vibrated in a sonic happening. And using our discussion of properties like agility and buoyancy, I believe the instantiation of timbre by a substance is like the instantiation of agility by an animal—both the properties are event-dependent: just as a cheetah *runs with agility*, a violin *vibrates with a distinct timbre*. Both agility and timbre are properties of their relevant substances, the properties that cause them to behave distinctly in relevant events.

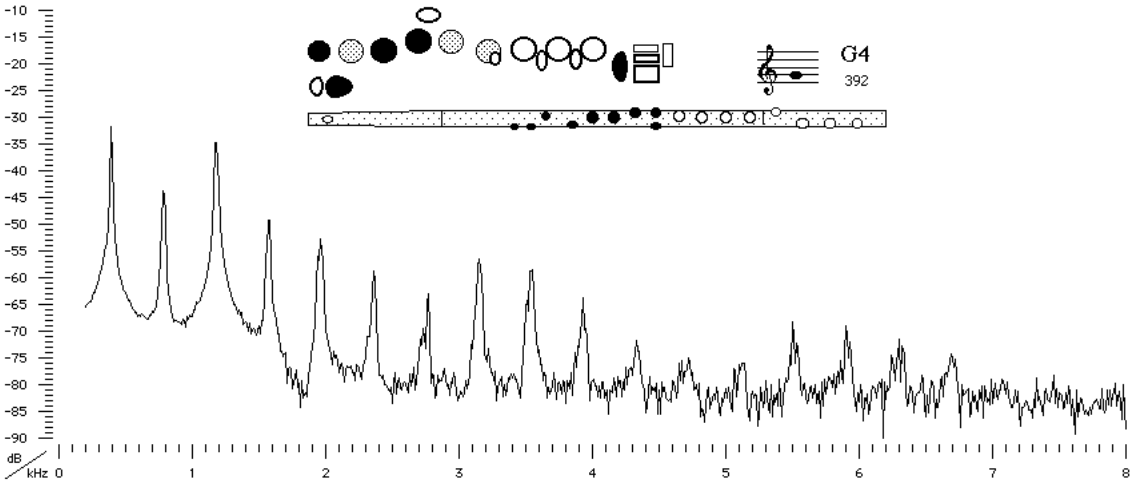


Figure 2.

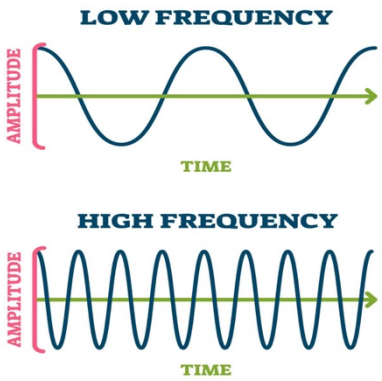


Figure 3.

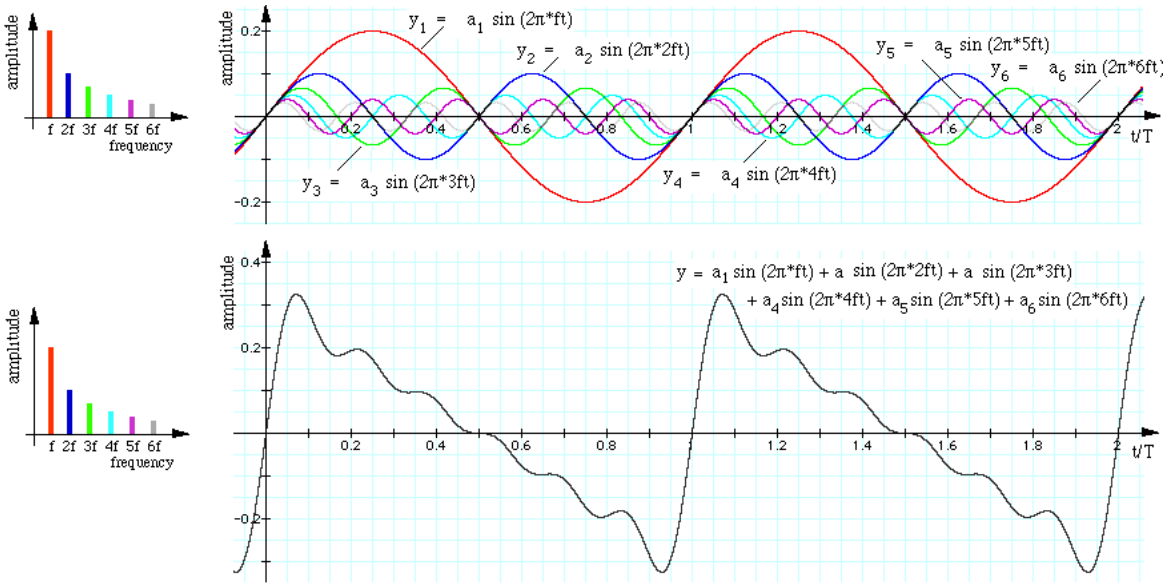


Figure 4

However, it should be emphasized that I do not intend to make the strong claim that timbre is wholly instantiated by a substance and pitch and loudness by the event it participates in to make a sound. I do not intend to put a strict border here. Regarding timbre, empirical studies

also suggest that *the manner of excitation* of an object and the features and shape of a sound spectrum are among the determinants of timbre. So, I think it is more plausible to leave it open how much of timbre is determined by the physical characteristics of a substance and how much by the characteristics of the event the substance participates in to make sounds. The weaker claim that I have tried to make is that timbre is at least in part instantiated by a substance partaking in a sonic happening. In fact, I do not think we have to necessarily consider timbre as a simple property; it may itself be a complex property constituted partly by the physical characteristics of an object and partly by features of the relevant sonic event.

So far, I have argued that sound is a complex property in the conjunctive form of P&T&L, and that the constituents/conjuncts of this complex property are instantiated by the substances and event(s) involved in a sonic happening. Ultimately in this section, I want to point out a perceptual or phenomenological characteristic of this complex property.

Borrowing the concept from Frank Jackson, I argue that hearing sounds is *indexically rich*. With regard to how an object's perceptual properties are represented in our experience, Jackson says:

> In representing something about shape, a visual experience *ipso facto* says something about color (in the wide sense that includes white, black and grey); and a similar point applies to extension, location and motion. Equally, you cannot prize the texture and temperature bits from the shape bit of a tactile representation. Something cannot feel to have some shape or other without feeling to have a texture or a temperature (in the wide sense that includes being neither hotter nor colder than one's limb).[154]

In the same sense, I believe hearing sound as a complex property is indexically rich. Pitch always comes with timbre and loudness. When we hear the sound of a musical instrument being played, we hear it with a specific volume, the distinct timbre of the instrument, and pitch, and all

[154] Jackson, F. (2003), p. 269.

these qualities come together. It is like seeing the size of an object which comes automatically with seeing its shape. If shape and size and probably color properties are related to each other in this way in our visual experience, I think timbre, loudness, and pitch (in the case of musical sounds) are so related and indexically rich. It is hard to separate (phenomenologically) the auditory perception of these sonic qualities in any case of auditory perception of sounds. When your friend speaks, their voice has a distinct timbre, but in recognizing their voice (via its timbre) you also hear the voice as loud, or quiet, or somewhere between. As said, I believe this is a phenomenological characteristic of hearing sounds as complex properties of their sources (objects and events).

5

We are now in a good position to provide a convincing explanation for the phenomenon that in hearing, we seem to directly auditorily perceive sound sources (objects and events in our surroundings), what I have called the Immediacy Observation (IO). We perceive things when we perceive their properties. I see an apple before me by perceiving properties such as its shape, size, and color. If sounds are complex properties of objects and events, hearing sounds is hearing their sources. Considering the issues the mereological view has, as I discussed in section 3, the property view I defended can explain (IO) more convincingly by appealing to the familiar idea of perceiving things through perceiving their properties.

A merit of the property view I have suggested is that it accounts both for the dynamicity of sounds (their temporal nature) and for their role in object recognition. Via timbre, hearing sound allows the perceiver to recognize a particular object involved in a sonic happening. By identifying some constituents of sounds with properties of events, the view provides an

explanation for the dynamicity of sound. We hear sounds as temporal entities because the relevant events are temporal. The specific time span in which a sonic happening occurs (for example, the period in which an object vibrates) is the main determinant of why we hear sounds as having specific durations. With no events occurring, there is no sound to hear. On the other hand, my theory can explain why by hearing the sounds of a particular object, which are made in different events, those sounds allow the perceiver to recognize that particular object. The reason is that they are constituted in part by some features of the object, because timbre is in part instantiated by the object.

The dismissal of the dynamicity characteristic of sounds is the main issue with the property views considering sounds as properties of substances on a par with properties like color.[155] According to John Kulvicki: "Sounds are stable properties of objects that seem to have them. More specifically, sounds are dispositions of objects to vibrate in response to being stimulated."[156] The main motivation for Kulvicki's view is *auditory/sound constancy*:

> Sound constancy of the right sort would suggest that the auditory system is structured so as to track features of objects that are stable across different circumstances. […] The [features] that matter are those that suggest objects sound the same across changes in the way that they are stimulated to vibrate, and thus across some changes in the pressure waves that they produce in response to those stimuli.[157]

I think Kulvicki is right, but he is not giving the whole picture. His focus is what I described as the recognitional characteristic of sound. Kulvicki plausibly points out one important aspect of auditory perception: that in various sounds of an object, we hear something constant: the object that vibrates and makes those sounds. Yet, this is partly true. I hear a sound

[155] Kulvicki (2008), and Pasnau (1999). Although Pasnau (1999) considered sound as a stable property of an object, he changed his view and suggested that sound is an event happening to an object (Pasnau 2009).
[156] Kulvicki, J. (2008), p. 2.
[157] Ibid., p. 9.

from the piano when one of its lower keys is pressed forcefully. Then I hear another sound from the instrument, but this time one of its upper keys is pressed gently. One thing remains constant in both cases: the particular timbre of the piano. But obviously, I also perceive salient changes with regard to loudness and pitch in the two cases. We would have to dismiss something important about sonic qualities and our auditory capacity if we ignored the auditory recognition of these *changes* (besides the auditory recognition of the aforementioned constancy). The constancy relates to the fact that the same object with the same physical, stable features vibrates on those two occasions; and the changes relate to the fact that that object participates in two different events, two different forms of vibration (with different frequencies) and pressing of the keys with different forces (which causes different amplitude attributes in the vibrations). As another example, consider how the loudness of a car crash is an important indicator of the force of the crash; that is, by perceiving the loudness, we hear features of the relevant event (the crash). Thus, it seems that our auditory system perceptually presents to us both the dynamic features of sonic happenings and the stable features of the objects participating in them. Although Kulvicki's account captures the latter (constancy) plausibly, it ignores the former (dynamicity).[158]

On the other hand, the view according to which sound is an event dismisses the auditory constancy related to timbre. A version of this view has been defended by Casati and Dokic.[159] The main difference between their event theory and O'Callaghan's discussed earlier is that

[158] In his most recent paper, Kulvicki (2017) argues that features of an event causing an object to vibrate can also be heard, but he still considers sound as the disposition of an object to vibrate. In my understanding, he does not justify why hearing sound should be considered as separate from hearing sonic events—why not claim that hearing sounds is also hearing features of sonic events?

[159] Their view has been formulated (in English) in Casati et al. (2013 and 2020).

Casati and Dokic do not consider sound as *a part* of a sonic happening but as a sonic happening itself. There is, in other words, no mereological relation in their theory in accounting for sound:

> The sound is not a proper part of a distinct event that is its source; it is identical with what we called the event source. The collision you hear is the sound you hear; the vibration of the stuff the tuning fork is made of is the sound. There is no difference between hearing the sound and hearing the event source, because there is no difference between the sound and the event source.[160]

Regarding what they mean by event sources:

> [I]n ordinary parlance two types of source are distinguished: thing sources (such as bells) and event sources (such as collisions or vibratory events at the object), events being here construed as spatiotemporally located individuals.[161]

So in their view, sounds are event sources in the sense given above. One thing that is not clear here is which event in a sonic happening a sound should be identified with: the event that causes an object to vibrate or the vibration itself. But we can set aside this matter and suppose a sound can be a complex event containing other events. The core idea in their theory is that sound is the event itself not a part of an event source.

One issue with such an event view of sound is connected with auditory constancy. In the theory under discussion, there is a distinction between an event source and a thing source. This seems to suggest that different thing sources can participate in the same (type) of event source. For example, two different objects can vibrate in the same way, with the same number of wave components, and with the same frequency and amplitude attributes in the associated sound spectra. In other words, as discussed in the previous section, the vibration of two distinct objects can have the same sound spectrum. This means that we have the same vibration (type) happening to two different objects. But as we discussed, the sameness of vibration (the event source) does

[160] Casati et al. (2013), p. 463.
[161] Ibid., p. 462

not guarantee the sameness of sound; that is, two distinct objects can vibrate in the same way, but their sound can still be distinguished. This suggests that sounds appear to perceptually represent something beyond their event sources; they appear to bear some characteristics of their object sources. But the event view under consideration cannot capture this because it identifies a sound with an event source. It cannot explain the auditory constancy we have discussed at different places in this paper so far.[162]

Another concern with this event theory is that when hearing sounds we seem to hear also objects participating in relevant events, what I have called (IO). Concerning this matter, Casati et al. say:

> The immediate object of auditory perception is just the event that is the sound, and then, in virtue of the perception of it we may recognize the thing source. 'In virtue of' means, roughly, that one has indirect awareness of x when one's awareness of x is caused by one's awareness of something else […]. Therefore, we say that our awareness of the thing source is caused by our awareness of its sound.

And then, they point out that understanding what this proposal exactly entails depends on the metaphysical nature of events and their relations to the objects they happen to. I personally do not know how this theory can cogently and successfully explain the idea that in hearing sounds in our surroundings, we hear objects and events. For example, if events include objects (according to one metaphysical view about the nature of events), then if sounds are identical with

[162] It might be argued that an event is always and necessarily connected to the thing it happens to. The idea is that an event always reveals in one way or another the features of the object participating in it. If that is the case, then considering a sound as an event can also explain how sounds reveal features of the objects participating in sonic happenings. Regarding such a viewpoint, it should first be pointed out that neither of the event views discussed in this paper explicitly endorses this construal. For example, according to Casati and Dokic, sound is not a thing source but an event source, and they do not provide any suggestion how event sources reveal anything about thing sources. Moreover, the viewpoint outlined above depends on a particular metaphysical view on events: that events ontologically are entities that somehow include or reveal features of the objects they happen to. This seems to be a version of or consistent with the view on events described earlier in this paper, that events are complex entities including objects. If we wanted to apply this viewpoint to our theory of sounds, we need first to show why such a metaphysical view of events is justified.

events—as this theory claims—then sounds *include* objects. But this seems a bizarre conclusion: sounds do not seem to be entities such that ordinary objects are wholly included in them!

An advantage of the view I defended in this paper is that besides explaining auditory constancy it also provides an answer for the (IO). According to my view, when an event happens to a substance/thing and the latter vibrates, a complex property (sound) is instantiated whose constituents are instantiated partly by the event source and partly by the thing source. What a person hears on such an occasion is this complex property, and the reason that they hear both the thing *and* the event is that they perceive properties of each in the complex property of sound. In this section, I showed how this view has advantages over the view according to which sound is a stable property of a thing and the view according to which sound is identical to the event happening to an object.

6

In this last section, I want to say how my theory can explain (DO), that we seem to sometimes hear sounds as distinct objects, or as objects distinct from their sources. In the first section, I explained how (DO) and (IO) seem to be in tension with each other. In the sections after it, I tried to give a convincing explanation for (IO) and argued that the complex property view of sound can provide that explanation better than the rival theories. I want now to explain (DO) based on the complex property view and show that, from this perspective, the two observations are not in fact in conflict.

Taking up (DO), I will argue that the distinction between a sound and its source is *not* ontological.[163] Sound is not something ontologically distinct from its source; that is, it is neither a distinct entity that is caused by its source nor a distinct part of its source. The (DO) comes from how we are capable of distinguishing between the perceptual appearance of a thing from the thing itself. When we tend to think of sound as something distinct from its source we attend to the *auditory appearance* of things (events and objects) in the world.

I think when we hear ambient sounds, our attention is on things and events in our surroundings. However, when listening to music, or at least in some forms of listening to music, we might tend to aesthetically attend primarily to sonic qualities themselves and not the sound sources. I suggest that considering sound as a distinct object is a product of *aesthetic auditory attention*, which typically happens in the case of music listening.[164] In an aesthetic experience, we might tend to pay our primary attention to how things *appear to us* rather than to those things themselves. From this perspective, I propose that a sound, considered as a distinct object, is *an auditory appearance* of its source.

An auditory appearance, in my intended use of the phrase, is basically similar to the visual appearance of things. Suppose you see an apple in a painting. What you see in this case, if the painting is somewhat realistic, is not a real apple but the way in which a real apple might *look* to you. You see the *visual appearance* of an apple in the painting, and this might cause you to *imagine* that you see a real apple. This visual appearance is a combination (conjunction) of a few properties characteristic of an apple: its color, shape, size, etc. Such properties that are

[163] In this section, I use "source" as a term for a combination of an object and the event happening to it: that is, an event source and a thing source together.
[164] To support the idea that sound is itself a distinct entity with its own qualities, both O'Callaghan (2017) and Scruton (2009) refer to experiencing sounds in music.

perceptually presented to us together in the painting constitute the visual appearance of the represented apple. This combination of properties is in fact a complex property according to the definition we gave in the previous section. For example, suppose a real apple has a particular shape (S), color (C), and size (Z) that are perceptually presented to us together when we see it. We see the related conjunctive property of S&C&Z in the drawn apple in a painting. That is, what we see in the apple painted is such a conjunctive/complex property.

I suggest that in the same way, an auditory appearance (in music listening) is constituted by the complex property of sound. When we hear a sound in music listening, we hear the auditory appearance of its source, and this auditory appearance is nothing but a complex property. Now, we have two basically similar aesthetic cases: as a representational painting contains the visual appearance of things and those visual appearances are complex properties, a musical piece contains the auditory appearance of (sound) sources, and these auditory appearances are complex properties too.[165]

What I want to point out now is that in our aesthetic experiences, sometimes how things *appear* to us in an artwork, and not those things themselves, is the primary object of our attention. For example, in a painting, it might not aesthetically matter that a specific *thing* (an apple, a house, a tree, the sky, etc.) has been painted but what might aesthetically matter is *how the thing looks* or has been represented in the painting. That is, our aesthetic attention is on the visual appearances themselves and not the things those visual appearances are appearances of. I think such an approach may also be typical in music listening: our musical attention might be primarily to how sound sources appear, or how the sources make sounds, in a piece rather than

[165] Visual appearances perceived in a painting might be considered as properties of the painting itself (a particular). In that case, I think musical sounds in a recorded performance can be considered as (complex) properties of that record itself.

what the sources are. That is, our aesthetic attention in music might be on sounds themselves instead of the sources those sounds are sounds of. Sounds, as complex properties, are perceptually rich enough to grab our aesthetic attention in a musical piece that typically contains lots of sounds that are interestingly related to each other in various ways.

The claim is not that in music we do not pay attention to sound sources. For example, sound sources most probably are the base of how our ear groups similar sounds together in a musical piece. But what I have tried to show is that in at least some forms of music listening, the *primary* aesthetic attention is on sonic qualities themselves and not on their sources. (Like the way in which in an appreciation of a painting how the color or shape of an object has been captured under the sun from a particular angle might be more aesthetically interesting than what that thing itself is.)

These arguments show how the Distinction Observation (DO) may be a result of aesthetic attention to sounds. I have argued that when we consider sounds as objects distinct from their sources, we tend to pay attention to them distinctly: we tend to make them distinct objects of our perceptual attention. And this mostly happens in music listening, when sounds become distinct objects of our music attention. This is how I think (DO) can be explained not as an ontologically relevant observation but as something related to how our mind can attend to the world and things and events in it in different ways. Sometimes the complex properties of objects and events can be distinct objects of our attention, and this is why sometimes sounds (as complex properties) can be distinct objects of attention. But this distinctness should not lead us to consider sounds as distinct objects ontologically on par with other substances and events in the world.

In this paper, I argued for a theory of sound according to which sound is a complex property possessed by an object *and* an event it participates in. I tried to show that this view can more convincingly explain the empirical evidence regarding timbre and our capacity to auditorily recognize objects and events in our surroundings in the sounds they make. I also showed how my theory can explain both the Distinction and Immediacy Observations by arguing that the former is a phenomenological aspect of audition more particularly occurring in auditory aesthetic experiences.

Bibliography

Armstrong, D. M. (1986). In defence of structural universals. *Australasian Journal of Philosophy*, *64*(1), 85–88.

Armstrong, D. M. (1997). *A World of States of Affairs*. Cambridge University Press.

Balzano, G. J. (1986). What Are Musical Pitch and Timbre? *Music Perception: An Interdisciplinary Journal*, *3*(3), 297–314.

Casati, R., Dokic, J., & Di Bona, E. (2020). Sounds. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*.

Casati, R., Di Bona, E., & Dokic, J. (2013). The Ockhamization of the event sources of sound. *Analysis*, *73*(3), 462–466.

Casati, R., & Varzi, A. (2020). Events. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*.

Darwin, C. J. (2008). Spatial Hearing and Perceiving Sources. In W. A. Yost, A. N. Popper, & R. R. Fay (Eds.), *Auditory Perception of Sound Sources* (pp. 215–232). Springer.

Davidson, D. (2001). The Individuation of Events. In *Essays on Actions and Events* (pp. 163–180). Oxford University Press.

Fisher, A. R. J. (2018). Structural universals. *Philosophy Compass*, *13*(10).

Gaver, W. W. (1993). What in the World Do We Hear? An Ecological Approach to Auditory Event Perception. *Ecological Psychology*, *5*(1), 1–29.

Giordano, B. L., & McAdams, S. (2010). Sound Source Mechanics and Musical Timbre Perception: Evidence from Previous Studies. *Music Perception: An Interdisciplinary Journal*, *28*(2), 155–168.

Handel, S. (1995). Timbre Perception and Auditory Object Identification. In *Hearing* (pp. 425–461). Academic Press.

Jackson, F. (1977). *Perception: A Representative Theory*. Cambridge University Press.

Jackson, F. (2003). Mind and Illusion. *Royal Institute of Philosophy Supplement*, *53*, 251–271.

Lewis, D. (1986). Against structural universals. *Australasian Journal of Philosophy*, *64*(1), 25–46.

Kulvicki, J. (2017). Auditory Perspectives. In B. Nanay (Ed.), *Current Controversies in Philosophy of Perception* (pp. 83–94). Routledge.

Leddington, J. (2014). What We Hear. In *Consciousness Inside and Out: Phenomenology, Neuroscience, and the Nature of Experience* (pp. 321–334). Springer.

Kulvicki, J. (2008). The Nature of Noise. *Philosopher's Imprint*, *8*(11), 1–16.

Kulvicki, J. (2017). Auditory Perspectives. In B. Nanay (Ed.), *Current Controversies in Philosophy of Perception* (pp. 83–94). Routledge.

Matthen, M. (2010). On the Diversity of Auditory Objects. *Review of Philosophy and Psychology*, *1*(1), 63–89. https://doi.org/10.1007/s13164-009-0018-z

McAdams, S. (2013). Musical Timbre Perception. In D. Deutsch (Ed.), *The Psychology of Music* (3rd ed., pp. 35–67). Academic Press.

McAdams, S., & Giordano, B. L. (2016). The perception of musical timbre. In S. Hallam, I. Cross, & M. H. Thaut (Eds.), *Oxford Handbook of Music Psychology* (2nd ed., pp. 113–124). Oxford University Press.

Moltmann, F. (2002). Events as Derived Objects. *Proceedings of the Colloque de Syntaxe et Semantique, Paris 2001 (CSSP01)*, 189–201.

Nudds, M. (2001). Experiencing the Production of Sounds. *European Journal of Philosophy*, *9*(2), 210–229.

O'Callaghan, C. (2011). Hearing Properties, Effects or Parts? *Proceedings of the Aristotelian Society*, *111*, 375–405.

O'Callaghan, C. (2017). *Beyond Vision: Philosophical Essays*. Oxford University Press.

Pasnau, R. (1999). What Is Sound? *The Philosophical Quarterly*, *49*(196), 309–324.

Pasnau, R. (2009). The Event of Color. *Philosophical Studies*, *142*, 353–369.

Patterson, R. D., Smith, D. R. R., van Dinther, R., & Walters, T. C. (2008). Size Information in the Production and Perception of Communication Sounds. In W. A. Yost, A. N. Popper, & R. R. Fay (Eds.), *Auditory Perception of Sound Sources* (pp. 43–76). Springer.

Schellenberg, S. (2019). The Unity of Perception: Content, Consciousness and Evidence (Summary). *Analysis*, *79*(4), 709–713.

Siegel, S. (2012). *The Contents of Visual Experience*. Oxford University Press.

Scruton, R. (2009). Sounds as Secondary Objects and Pure Events. In M. Nudds & C. O'Callaghan (Eds.), *Sounds and Perception: New Philosophical Essays* (pp. 50–68). Oxford University Press.

Shipley, T. F. (2008). An Invitation to an Event. In T. F. Shipley & J. M. Zacks (Eds.), *Understanding Events: From Perception to Action* (pp. 3–30). Oxford University Press.

van Dinther, R., & Patterson, R. D. (2006). Perception of acoustic scale and size in musical instrument sounds. *The Journal of the Acoustical Society of America*, *120*(4), 2158–2176.

Walton, K. L. (2015). Two Kinds of Physicality in Electronic and Traditional Music. In *In Other Shoes: Music, Metaphor, Empathy, Existence* (pp. 36–53). Oxford University Press.

Yost, W. A. (2008). Perceiving Sound Sources. In W. A. Yost, A. N. Popper, & R. R. Fay (Eds.), *Auditory Perception of Sound Sources* (pp. 1–12). Springer.

Yost, W. A., Popper, A. N., & Fay, R. R. (Eds.). (2008). *Auditory Perception of Sound Sources*. Springer.

Conclusion

This dissertation was based on the following two research questions:

(1) how can music (a set of organized sounds) be heard as expressive of emotions such as sadness or happiness?

(2) how are sounds (not limited to musical sounds) and their sources related to each other?

In what follows, I want to explicate how my findings contribute to addressing these questions. I will also sketch how what I accomplished in the dissertation will direct my research in the future.

The Role of Aesthetic Attention

The gist of my answer to question (1) is that hearing musical expression is grounded in how we perceive music's continuity, flow, and movement. To develop this idea, I argued in Chapter 1 that we need to distinguish the phenomenological and causal aspects of musical expression. I showed that the main reason for the importance of this distinction is that when we sometimes hear music as expressive, we are not necessarily aware of what specific musical properties are responsible for the expression. As a result, the phenomenological dimension of expression can be explored independently of what musical properties cause expression. My critical examination of the appearance theory in Chapter 1 was based on the relevance of this distinction. By suggesting that musical expression is a matter of the hearer's detecting similarities between musical properties and human expressive behavior, the appearance theory conflates these two dimensions in its explanations. The problem with that theory is that it is not true to musical experience: I can hear a piece or passage of music as being expressive of, say, sadness without identifying a similarity between music and anything else.

The analytical framework I built in Chapter 1 led me to focus, in Chapter 2, on how music experience, as regards hearing musical expression, can be explained independently of the causal dimension of expression. I examined three main ideas some theories have used to give such an explanation: the imaginative idea (that we imagine personae in music expressing emotions), the metaphorical idea (that we metaphorically attribute emotions to music), and the dispositional idea (that expressive music arouses emotions in us). My critiques of these theories showed a common issue in them: they don't account for the fact that our hearing of music as expressive depends on how we *attend to* different musical properties. How an acculturated listener pays attention to music determines whether they can hear music's musically *and* culturally relevant expressive qualities. The theories examined in Chapter 2 miss how this attention is part of the phenomenology of hearing musical expression. As I said in that chapter, it is important to keep in mind that paying attention to the musical properties responsible for musical expression is not the same as being aware that this or that specific musical property causes expression. (I can perceive the beauty of, say, a painting by attending to the relevant visual properties, but this does not mean that I am necessarily aware of what specific properties are responsible for the perceived beauty and how they make the painting beautiful.)

The analyses of Chapter 2 paved the way for my argument in Chapter 3 to show what role music attention plays in hearing musical expression. I first explained the role of music attention as a requirement, particularly for an arousal view of musical expression. I then analyzed the notion of musical movement and its dilemma—that we sense continuity, flow, and movement by hearing the discrete events of music without hearing that any of those events themselves move. I argued that to perceive that flow, motion, and continuity, the listener gives a feeling response to music. The idea is that the listener pays attention to the relational qualities of discrete

musical events, and to make sense of how those events make a continuous, coherent whole, she feels a movement, a continuous process of change, in herself. Now, such a feeling is a component of perceiving music, and this affective component can sometimes contribute to attributing specific emotions to music. The reason is that the feeling response we give to music can sometimes be similar to those feelings we have when experiencing emotions (the feeling of sadness, for example). However, I argued that the affective component of music perception is one factor in attributing emotions to music, and possible cultural influences should not be ignored in this relation.

Thus, in response to the research question (1), I developed a dispositional (or arousal) theory of musical expression. However, by explaining how music attention is an essential part of our feeling response to music, my theory can solve a longstanding problem the arousal theories have grappled with in connecting the feeling experienced by the listener in response to expressive music and her aesthetic attention to the music itself. My theory also provides a solution for the problem of musical movement by delineating the process in which music flow, continuity, and movement are *affectively perceived* by an attentive listener. Moreover, as sketched in the Future Directions section below, what I have proposed in response to (1) sets a framework for different cultural considerations as regards musical expression.

Constancy and Dynamicity in Sounds

My response to the research question (2) given in Chapter 4 is that by hearing sounds we hear sound sources, because sounds are properties of their sources. Much as by seeing the color of an object, we see that object, by hearing the sound of a metal bar being struck, we hear that metal bar and its being struck. Sound sources are considered objects *and* the events they

participate in to make sounds. And sounds are considered, according to my theory, as *complex properties* of their sources: properties that are constituted by properties of objects and properties of the events those objects participate in to make sounds.

This view of sound can better explain dynamicity and constancy heard in sounds. Sounds of a specific object (such as the flute) present a constant timbral quality. The violin has a different timbre than the flute, a wooden bar has a different timbre than a metal one, and so on. This constancy has led some to construe sounds, like colors, as properties of objects. However, sounds made by a specific object also present a dynamic character. Pitch and loudness seem to be the best candidates for the dynamic character of sounds. Although the sounds of the violin remain mostly constant in terms of their timbre, they can differ in terms of their pitch and loudness. This dynamic character is basically connected with how the object is excited in an event to vibrate: different ways of striking an object result in different vibratory patterns and, in turn, different pitches and loudness. Our auditory capacity can recognize both the constancy presented in the sounds made by a specific object and their dynamic characteristics. A merit of my proposed view on the nature of sounds is that it can give us a straightforward explanation of the mentioned perceptual capacity: in a sound, we hear properties of objects (constancy) *and* of the events they partake in (dynamicity), since sounds are *complex properties* instantiated by these objects and events.[166]

People who construe sounds as entities distinct from their sources usually argue that we *hear* sounds as distinct entities with their own properties (pitch, loudness, and timbre). Sounds, they argue, are not possessed, as properties, by other entities but possess their own properties.

[166] Sometimes through timbre we hear the properties of more than one object—for example, the timbre of cymbal stuck by a drumstick is different from when it is struck by a felt mallet. In developing my theory in future I intend to explore this character of timbre.

This is why we should consider them as entities ontologically distinct from their sources. I explained how this phenomenological consideration is in tension with another one according to which we can also hear things through their sounds (we can hear a car, a piano, someone walking, etc.). To remove this tension within my proposed view on sounds, I argued that hearing sounds as distinct entities is the result of our aesthetic attention to sounds, specifically in the music experience. Arguably, it might be the case that in paying aesthetic attention to sounds in music listening, what is important for listeners are pitches, loudness, and timbre from different aesthetic points of view and not the relevant sound sources. However, even if such a consideration were correct, the fact that we give distinct attention to sonic properties is not a good reason to propose an ontology according to which sounds are distinct entities. My view can explain the phenomenological consideration in question: complex properties can be distinct objects of attention, particularly due to their complex nature—and this might, arguably, happen in the case of music.

Future Directions

Two possible directions toward which my current research can go in the future are as follows.

(1) In Chapter 3, I argued that the feeling response the listener gives to music, with regard to musical movement, can contribute to attributing emotions to music. However, I also pointed out that this contribution is partial, considering that feelings do not seem to be the only distinguishing factor between different emotional states. I mentioned the important role of culture in this relationship. We do not listen to music in an aesthetic vacuum: our aesthetic appreciation of music is tied to different social and cultural contexts in which we listen to music. It would thus be implausible to assume that cultural factors are not aesthetically

relevant to musical expression. My theory of musical expression is able to incorporate different cultural considerations. When it comes to a listener's connecting what she feels during listening to music with a specific emotion and attributing that emotion to music, cultural associations play a key role. Conceptualizing my feeling response to music as, say, sadness can also depend on the cultural context in which the music is listened to, the texts or narratives in connection with which that music is thought of, and so on. I anticipate developing my theory of the role of cultural context in musical experience and expression in the future.

(2) Another subject that needs more exploration is the relationship between my theories of sounds and of musical expression. I described in the Introduction how a musical curiosity initiated my research on the nature of sounds and auditory perception. My theory of sound suggests that hearing sounds is hearing concrete objects and events in the world. A question that arises now is: What does this view entail about the content of musical experience? One thing it can entail is that when we record a performance having occurred, say, five decades ago, we do not merely capture the sounds but also properties of the objects and events existing and happening in concrete social and cultural contexts at that time, and we hear such historical entities by listening to that recorded audio. We can compare this with seeing people in photos taken years ago. When I see a photo, arguably, I do not merely perceive a bunch of colors and shapes but the things possessing (or having possessed) those visual properties.[167] My theory of sound might lead to such a conclusion with regard to recorded sounds. What that theory entails about the content of auditory or musical perception is another subject I am curious to develop my current research toward.

[167] Regarding this feature of photography, see Walton (1984). Frazier and Fried (2023) have also insightfully argued about his matter in their "Left Behind: Anti-Fatness, Photography, and Aesthetic Agency."

Bibliography

Frazier, C. and Fried, J. (2023). Left Behind: Anti-Fatness, Photography, and Aesthetic Agency. [Manuscript in progress]

Walton, K. L. (1984). Transparent Pictures: On the Nature of Photographic Realism. *Critical Inquiry*, *11*(2), 246–277.