The role of expectations in musical experience

Jenny Judge and Bence Nanay

Almost every facet of the experience of musical listening—from pitch, to rhythm, to the experience of emotion—is thought to be shaped by the meeting and thwarting of expectations. But it is unclear what kind of mental states these expectations are, what their format is, and whether they are conscious or unconscious. Here, we distinguish between different modes of musical listening, arguing that expectations play different roles in each, and we point to the need for increased collaboration between music psychologists and philosophers in order to arrive at a more detailed characterisation of conscious musical experience and the role of expectations therein.

## I. EXPECTATION in music: a survey

In *Emotion and Meaning in Music*, Leonard Meyer (1956) put expectation squarely on the agenda for psychologists of music[[1]](#footnote-1). Though Meyer’s particular interest was in the role of expectations in the emotional experience of music, almost all aspects of musical experience—pitch, rhythm, melody, harmony and the emotional experience of music, for instance—are now thought to be shaped by expectations, which arise both from general perceptual organisation processes and also learned stylistic patterns.

Eugene Narmour’s influential ‘Implication-Realization’ theory (Narmour 1990) proposes that melodic expectations are predicted on the basis of principles of melodic shape. For Narmour, the presence of a large interval gives rise to an expectation that it will be followed by a change in direction; a small interval elicits an expectation for another small interval in the same direction. A family of subsequent theories has attempted to streamline and quantify Narmour’s original theory (see, for example, Cuddy and Lunney 1995, Krumhansl 1995, Schellenberg 1997, Margulis 2005). Other influential models include Steve Larson’s (1997–1998, 2004) model of expectation by analogy with physical forces, as well as Fred Lerdahl’s (2001) theory of tonal attraction.

The experience of harmonic structure has also been discussed in terms of expectations. If two chords are heard as harmonically related, it is thought that hearing the first chord (or the ‘prime’) sets up an expectation of hearing the second (the ‘target’). Jamshed Bharucha and colleagues (Bharucha and Stoeckig 1986, 1987; Tekman and Bharucha 1992, see also Krumhansl 2000 for discussion) presented subjects with two successive chords, before asking them to judge whether or not the target (second) chord was in tune, measuring how long it took them to deliver this judgment. In this way, the experimenters thought, they could indirectly measure the extent to which the prime set up an expectation of the target chord. They observed that the reaction time on the tuning judgment was shorter when the prime and target chords were part of the same diatonic set of chords (the set of triads in a given key, in other words) than when they were not.

Expectation is also thought to be central to rhythmic experience. Entrainment is the process by which a psychological rhythm can become synchronised with some regularly occurring event in the environment. We can tap our foot to a beat with ease; that we can do so is thought to show that we accurately predict or expect the timing of events in a regular pattern, and then coordinate our behaviour with that pattern. Various authors have suggested that entrainment is not just related to rhythmic experience, but that it also lies at the heart of auditory attention as well. Mari Reiss Jones and colleagues (e.g. Jones 1981, Large and Jones 1999, Barnes and Jones 2000) argue that attention to periodic events, like musical rhythms, is underpinned by dynamic expectations: we do not pay attention equally at all moments. Rather, our attention is directed toward the most likely moments of stimulus onset. Attention waxes and wanes, in other words, in line with the expectations of onset that a pattern sets up.

All this work on expectations in music stems from Meyer’s seminal contribution, but David Huron’s work represents a more direct successor, insofar as he targets affective experience—the experience of emotion in response to music. In the widely-cited *Sweet Anticipation* (2006), Huron argues that expectations are the cornerstone of that experience . When melodic, harmonic or rhythmic expectations are met, a positively-valenced prediction response is elicited. As I follow a regular rhythm, for instance, I expect the downbeat to occur at a particular time (or at least within a certain window). A temporal prediction is thus formed. And when the downbeat does, in fact, occur at the time I predicted, I experience satisfaction. This, says Huron, is because it is biologically advantageous for me to make good predictions; and so I am rewarded for doing so, such that the heuristic I used to make that prediction will be used again in the future. However, it is not only when events are successfully predicted that I can experience satisfaction: I can also experience pleasure when my expectations are thwarted. After all, it is often when the music seems ‘surprising’ to me, rather than predictable, that I experience satisfaction as I listen. Whether those expectations are thwarted or met, in other words, Huron thinks that expectations always lie at the heart of emotional responses to music.

Most discussions of musical expectations focus on unconscious processing. Characterisations of the role of expectation in conscious experience are less common. We may hence ask: do any conscious expectations feature in musical experience? Or do we have conscious access only to their effects— experiences of tension and relaxation[[2]](#footnote-2), for example? There is ambiguity in the literature on this front, as we will now see.

II. Musical expectations: Conscious or unconscious?

The likelihood that unconscious ‘expectations’ exist, and that they affect experience, has long been acknowledged. Nevertheless, there has been ambiguity concerning the role of conscious expectations in musical experience at least since Hanslick’s *On the Musically Beautiful* (1854):

The most significant factor in the mental process which accompanies the comprehending of a musical work and makes it enjoyable … is the mental satisfaction which the listener finds in continuously following and anticipating the composer’s designs, here to be confirmed in his expectations, there to be agreeably led astray. It goes without saying that this mental streaming … occurs unconsciously and at the speed of lightning. (*ibid.*, p. 64)

For Hanslick, musical experience owes its aesthetic appeal to automatic, unconscious expectations. However, these expectations are evidently not *entirely* unconscious: after all, the listener is sometimes able to explicitly ‘anticipate’ the composer’s intentions, sometimes getting it right, and sometimes erring. Even if unconscious expectations that proceed automatically are undergirding musical experience, how should we understand the way an unconscious mental state (the expectation) influences a conscious mental state (the musical experience)? Does this involve the conscious representation of some of these normally-unconscious expectations, or aspects thereof? Or is conscious musical experience confined to feelings of, say, tension and release, without the subject’s really experiencing any expectations at all?

Meyer proposed that perceptual stimuli create unconscious expectations, which give rise to tendencies to respond in certain ways to those stimuli. These tendencies to respond, he says, can be conscious or unconscious (Meyer *op. cit.,* p. 24). Normally, when an expectation is met and a tendency runs its course, the process is unconscious. It is when the response tendencies are inhibited that they become conscious:

Countless reaction patterns, of which the responding individual is unaware, are initiated and completed each hour. The more automatic behavior becomes, the less conscious it is. The tendency to respond becomes conscious where inhibition of some sort is present, when the normal course of the reaction pattern is disturbed or its final completion is inhibited. Such conscious and self-conscious tendencies are often thought of and referred to as “expectations.” (*ibid.,* p. 24)

So, for Meyer, it is the inhibition of a tendency to respond (which results from a thwarted expectation) that is crucial when it comes to conscious musical experience. However, one may ask here: is the thwarted expectation *itself* (that X will occur, say) represented in conscious experience? Or is it merely the experiential *effect* of such a thwarted expectation that is so represented? Sometimes Meyer implies that he means the latter; he says, for example, that musical passages can give rise to a feeling of tension or suspense, which is ‘essentially a product of ignorance as to the future course of events.’ (*ibid.,* p. 27) It seems like no expectation at all is being experienced here. However, at other times, he implies that we do entertain at least some expectations as we listen, and are aware of them before they have either been confirmed or thwarted. He says, for example:

Sometimes a very specific consequent is expected. … [The] consequent chord is expected to arrive at a particular time, i.e., on the first beat of the next measure. … At other times expectation is more general: that is, though our expectations may be definite, in the sense of being marked, they are non-specific, in that we are not sure precisely how they will be fulfilled. (*ibid.*, p. 25–26)

Even though the expectations may be more or less specific, that is, Meyer seems to be suggesting here that those expectations do, in fact, feature in the content of musical experience, and that it is not merely the upshot of unconscious processing that plays a role. Which interpretation should we prefer?

This ambiguity survives in present-day discussions of musical expectations. For instance, Carol Krumhansl, whose research on musical expectations in harmonic and pitch perception is extensive, implies in places that we only have conscious access to feelings of tension and relaxation when we listen to music. These are to be thought of as the by-products of unconscious, implicit perceptual expectations, which are formed on the basis of exposure to regular sounds as well as particular stylistic conventions; the expectations themselves do not themselves feature in conscious experience. In a 1997 review article, she seems to interpret Meyer as making this sort of argument, which she also endorses:

Expectations *produce* waves over time of tension and release from tension. Expectations are *derived from* both general psychological principles (such as Gestalt principles of perceptual organization) and knowledge of the style (such as tonality, harmonic progressions, and musical form). (Krumhansl 1997, p. 338, emphasis added)

Here, she is implying that it is just the waxing and waning of tension that features in the conscious experience. The substrate of expectations is not itself consciously accessible, or at least, does not normally protrude into consciousness.

This argument is echoed by Brattico and Pearce (2013), who describe the mainstream view of expectations in music psychology (as emanating from Meyer, and foreshadowed in Hanslick) thus:

musical enjoyment is linked with patterns of tension and resolution resulting from the confirmation and violation of perceptual expectations of which we are usually unconscious. These expectations might concern, for example, the pitch of the next note in a melody, the next chord in a pattern of harmonic movement, or the timing of the next note in a solo percussion performance. (*ibid.*, p. 53)

These expectations are unconscious; they are implicitly acquired through statistical learning ‘in which listeners construct implicit probabilistic models of the next element in a musical sequence’ (*ibid.,* p. 53). It sounds as though like Krumhansl, these authors also endorse a view whereby it is only the by-products of implicit, unconscious expectations that feature in experience.

Elsewhere, however, Krumhansl suggests that listener expectations—which can be explicit expectations and predictions about tonal hierarchies, for instance—can, and do, feature in ongoing musical experience. In her 1990 book, she asserts that a central line of thought informing her research is the following idea, which she attributes to Meyer:

Of greater psychological import […] is Meyer’s suggestion that through experience, listeners internalize the complex system of probability relationships, and, when listening to a particular piece of music, relate the sounded elements to this knowledge. This process gives rise to dynamically changing expectations about subsequent events, which may or may not be satisfied, or may be satisfied only partially, indirectly, or with some delay. In these expectations reside what to Meyer is the syntax, the meaning, and the aesthetic experience of music. (Krumhansl 1990, p. 63)

Krumhansl here at least seems to be hinting that the expectations one may form and articulate about music really do feature in the ongoing conscious experience of music, that one actively relates them to what one hears as one listens to a piece.

This idea—that we have conscious access to the expectations that drive our experience—underpins much of the methodology of the research into expectations. Models of expectation are tested according to whether they correlate with subjects’ expectation judgments. For instance, in a series of studies by James Carlsen and colleagues (Carlsen et al. 1970, Carlsen 1981, Unyk and Carlsen 1987), listeners were presented with two successive tones, and asked to sing what they believed would be the continuation of the melody, had it been allowed to continue. A related method, known as the ‘probe-tone method’, has also been widely used (e.g. Krumhansl and Kessler 1982, Cuddy and Lunney 1995, Krumhansl 1995). Here, various musical contexts—like the two-tone contexts above—are followed by a probe tone. Listeners are asked to rate how well the probe tone matches their expectations as to what was likely to have followed from the previous two tones. This methodology seems to require that listeners have some access to their expectations. But if this is the case, then these experiments say nothing about our unconscious expectations unless we posit that unconscious experiences can somehow be consciously accessed—a blatant contradiction on at least some accounts of consciousness.

David Huron provides a taxonomy of expectations which might help to resolve the ambiguity. He suggests that there are four different types of expectations: schematic, veridical, dynamic and conscious expectations (Huron 2006). First of all, we have schematic and veridical expectations (this distinction is borrowed from work by Bharucha and colleagues, e.g. Bharucha and Stoeckig 1986, Bharucha 1994). *Schematic expectations* are those that arise from statistical learning. They are general: they concern, for example, the kinds of regularities that are usually apparent in music of a given style. *Veridical expectations* are specific to a given piece: when I am familiar with a tune, it is thought, I have precise veridical expectations about what will happen next. I ‘know how it goes’, in other words. The third kind of expectations are ‘*dynamic expectations*’, which are formed as a piece goes along: in cases of improvisation, for example, a listener can come to expect a note that sounds ‘wrong’ (because it conflicts with their schematic expectations about what is permissible in the current key, for example) if the performer keeps repeating it with apparent intent. These three types of expectations—veridical, schematic and dynamic—Huron tells us, are ‘preverbal or unconscious in origin’ (*op. cit.,* p. 235). The final class of expectations—what Huron calls ‘*conscious expectations*’—arise from explicit deliberation and reflection. For example, reading ‘Lento’ in the programme notes at a concert might cause me to expect the music to be slow.

Huron’s taxonomy is problematic in a number of respects. First, it is unclear whether it is intended to be exhaustive. Second, it is also unclear if the categories may overlap or not (for instance, whether a dynamic expectation could also be either veridical or schematic). Third, and more importantly, there is no indication as to whether these expectations are meant to feature in conscious musical experience or not. Being ‘unconscious or preverbal in origin’ does not necessarily mean that the schematic, veridical or dynamic expectations cannot be conscious. One can entertain many explicit thoughts, such as ‘I would like a ham sandwich’, that are no less conscious for having originated in the preverbal state of hunger—an unconscious registration of which physiological state could well have prompted my intention to make myself a snack. So, when we are trying to characterise conscious experience, Huron’s taxonomy does not overly help matters. So how are we to proceed?

Given that there seems to be a consensus that (a) there are unconscious expectations as well as conscious ones; (b) implicit perceptual expectations, or at least their effects, play some role or other in musical experience; and (c) it seems reasonable that explicit theoretical expectations (such as those arising from theoretical knowledge about music) and/or their effects could play a role in musical experience too, then we have at least four possible scenarios when it comes to characterising conscious musical experience. The first is as follows:

Case 1: All implicit perceptual expectations are unconscious during an episode of music listening; it is only their effects (such as experienced tension and release) that feature in conscious experience. All explicit theoretical expectations (such as those arising from general musical knowledge) are also unconscious; it is only their effects (such as experienced tension and release) that feature in conscious experience. Hence, conscious musical experience includes the effects of implicit and explicit expectations only; no expectations of any kind are consciously entertained.

We can hold the situation fixed for implicit expectations, and allow that at least some theoretical expectations could be represented in conscious musical experience, yielding a second scenario:

Case 2: All implicit perceptual expectations are unconscious during an episode of musical listening; it is only their effects that feature in conscious experience. *At least some* explicit theoretical expectations however, are consciously experienced. Hence, conscious musical experience represents the effects of implicit perceptual expectations, the effects of unconscious theoretical expectations, *and at least* some of our conscious theoretical expectations themselves.

Holding the situation fixed for explicit expectations yields the following two scenarios:

Case 3: *At least some* implicit perceptual expectations become conscious during an episode of musical listening. All explicit theoretical expectations are unconscious. Hence, conscious musical experience includes the effects of implicit expectations, *at least some* of those implicit expectations themselves, and the effects of explicit theoretical expectations.

Finally, we have the following scenario:

Case 4: *At least some* implicit perceptual expectations are consciously experienced during an episode of musical listening. *At least some* explicit theoretical expectations are consciously experienced as well. Hence, conscious musical experience includes some implicit perceptual expectations, the effects of other implicit perceptual expectations, at least some explicit theoretical expectations, and the effects of other explicit theoretical expectations.

The above four cases assume, however, that conscious musical experience is homogeneous: that there is only one kind of musical listening. In Section IV, we will suggest that it is not. Before we do that, however, we need to address another group of questions about musical expectations. What kind of mental state is a musical expectation? What kind of content could it have, and how finely-grained might this content be? And lastly, what reason do we have to suppose that *any* expectations could influence perceptual experience in the first place?

III. musical expectations: Structure and content

Discussions in music psychology and music theory are surprisingly vague about what expectations are. Turning to the philosophy of perception is not very helpful either, however, because while references to the importance of expectation are rife in philosophy of perception, there is no agreement about how we should think about expectations themselves.

The ‘poverty of stimulus’ argument proposes that expectations are a necessary ingredient of any kind of perception, since the sensory stimulation associated with any perceived object seriously underdetermines that object. The same retinal image, for example, could be the projection from a vast number of three-dimensional objects (Marr 1982). How, then, does the perceptual system arrive at the ‘correct’ reconstruction of the perceived scene in spite of the poverty of stimulus? A series of ‘expectations’ about the perceived scene—or, assumptions about what is present, derived from past experiences—are thought to guarantee the delivery of a stable percept.

There are more and less radical versions of this claim. According to the currently popular ‘predictive coding’ approach, perception is by its very nature prediction: what our sensory systems do is that they predict what we will perceive in the next moment, on the basis of expectations arising from contextual clues as well as past experience (Clark 2013, Friston 2005). It is only those parts of the stimulus that fail to conform to our expectations that are processed by our perceptual systems. In this picture, all perception is expectation-driven. Even if we don’t endorse this radical claim, it seems clear that if we hold the sensory stimulation fixed and vary the expectations, our conscious perceptual experience will also vary, given the underdetermination of the stimulus. Note that this claim is more substantial than saying that our *overall conscious* experience will change if we hold the sensory stimulation fixed and vary the expectations. The claim we are interested in is about the influence of expectations on *perceptual* phenomenology, not *overall* phenomenology[[3]](#footnote-3). Influences on overall phenomenology are cheap: as long as the expectation is conscious (something presupposed in Cases 2, 3 and 4 above), expectations would automatically alter overall phenomenology. But what is at stake in understanding the role of expectations in musical *listening* is whether and how expectations influence *perceptual* phenomenology.

Returning to the musical case, the very general picture that emerges from the philosophy of perception is as follows. We have expectations about music, based on contextual clues arising from the music itself as well as our past experience of listening to, reading about and talking about music. These can be thought of as constituting some kind of non-perceptual background mental state. Let us call this state A. The influence of A on perceptual experience could be direct or indirect.

If the influence is direct, then our expectations, considered to be non-perceptual mental states, directly influence our perceptual experiences. One advantage of this view is that it doesn’t put significant constraints on what expectations really are: it is consistent with considering expectations to be beliefs or dispositions, for example. Nevertheless, this framework presents something of an interface problem. Our musical expectations can be very complex: they can be theoretical, or even historical, social or cultural. It is unclear how non-perceptual mental states with such rich content could manage to influence our perceptual experience (which, presumably, has a much less rich content) without any mediation. It seems that the ‘direct influence’ view approach will put serious constraints on the richness of the content of expectations, since there are constraints on the richness of the content of perceptual states they influence directly.

One way out of this conundrum would be to beef up the content of the perceptual experience itself, by claiming that our perceptual experience of music has rich content. If perceptual experience has rich content—if it is already ‘as of’ the social, historical and theoretical dimensions of music, for instance—then this would make for a more straightforward interface between complex expectations and perceptual experience. Philosophical opinions diverge as to the range of properties that is perceptually experienced, however. Some argue for the sparse view that only very low-level properties like shape or colour, or perhaps pitch and loudness in the auditory case, are perceptually experienced (Dretske 1995). If this is right, then the interface problem remains, and it becomes difficult to see how expectations can be part of perceptual experience. But some philosophers of perception argue that we perceptually experience higher-level properties, such as sortal properties, causal relations, action-properties and so on (Siegel 2006, 2009, Bayne 2009, Nanay 2012a). In the musical case, then, one might argue that we perceptually experience rich musical features like tonal relations, metrical hierarchies and perhaps even formal features, such as being in sonata form). If this is right, the influence of rich expectations is easier to defend—but arguments are required to establish that the perceptual experience of music does, indeed, have such rich content. Moreover, even this richer way of thinking about perceptual content will make it difficult to explain how expectations with highly sophisticated content could interact with perceptual experiences that have much less sophisticated content.

The second option would be to say that the influence is indirect, by allowing for some kind of mediation between the non-perceptual mental states and the perceptual experience. This yields a two-step picture. First, we have A: the non-perceptual mental state that is comprised of our knowledge about music (including our expectations), which is in turn built from our previous exposure to music and music theory. A is a relatively static background state. Second, we have B: a mediating mental state, whose content is at least partly determined by A. B is a dynamic, fleeting mental state that parallels our perceptual experience without being identical with it. And thirdly, we have C: our perceptual experience of musical listening, which is influenced directly by B (and thus indirectly by A). Now, unless we appropriately specify what kind of state B is, the interface problem will remain; some account is owed as to how it is that B can ‘translate’ between expectations with rich content, and perceptual experience with sparse (or, at least, sparser) content.

Given we can identify some candidate states for B, a preliminary question we need to ask is whether the crucial expectations are located in A, or B, or both. Recall that B is a fleeting mental state that changes as our perceptual experience changes, while A works in the background. B seems to be a better bet for those who want to understand how our expectations may alter in real-time, and may thus influence our musical experience as it unfolds in a dynamic fashion. Since the content of B is itself determined by A, however, A still needs to be part of any such explanation.

Now if we are happy to endorse this sort of ‘indirect influence’ picture, we still need to figure out what kind of mental state B is. What kind of mental state can mediate between non-perceptual mental states and perceptual experiences? Here, we can help ourselves to the conceptual apparatus of the cognitive penetration debate, which is concerned with whether and how higher cognitive states influence our perceptual experiences. We will suggest that attention and mental imagery are apt candidates for mental state B, given ample evidence that both of these states frequently do mediate between higher-order mental states and perceptual experience, and also have the appropriate dynamic character that seems desirable in the musical case.[[4]](#footnote-4)

‘Cognitive penetration’ is, roughly speaking, the view that higher-order cognitive processes can, and do, influence lower-level perceptual processes. The view that perception is cognitively penetrable (to at least some extent) seems more and more plausible in the light of recent empirical findings (e.g. Goldstone 1995; Hansen et al 2006; Lupyan & Spivey 2008; Lupyan et al 2010; Lupyan & Ward 2013; Nanay 2013a, 2013b; Siegel 2011; Macpherson 2012; but see also Firestone & Scholl 2014 for a critical analysis). Top-down processes influence perceptual processing as early as the primary sensory cortex (Gandhi et al. 1999) or the thalamus (O’Connor et al. 2002). Such top-down influences become more numerous and more diverse as the perceptual processing continues.

In a classic experiment (Delk and Fillenbaum 1965), subjects were asked to match objects to colour samples. They tended to match a picture of an orange heart to colour samples closer to the red end of the spectrum than they did other, non-heart-shaped orange objects. This shows that the recognition of the object in question (the heart, in this case) influences the colour one experiences it as having. In a more recent experiment (Levin and Banaji 2006), two pictures of identical (mixed-race) faces were shown to subjects. The only difference between the stimuli was that under one picture, the subjects read the word ‘white,’ while under the other they read ‘black’. When asked to match the colour of the face, subjects chose a significantly darker colour for the face labelled ‘black’,[[5]](#footnote-5) suggesting that higher-level cognitive processes (such as the comprehension of written text, in this case) can influence one’s perceptual experience. Both of these experiments could be interpreted as demonstrating the influence of expectations on perceptual experiences: our expectation that hearts tend to be red, in the first case, and the expectation that people identified as ‘black’ have darker skin in the second. Mounting evidence suggests that experience is not just determined in a bottom-up manner by the perceptual stimulus: it depends on language, attention and one’s expectations as well, as well as one’s cultural context (see Hansen et al. 2006, Lupyan and Ward 2013, Lupyan et al. 2010).

However, the possibility of ‘cognitive penetrability’ has been extensively debated in recent decades. Depending on how one defines the concept (see Siegel 2011, Macpherson 2012 for summaries), it may not be too farfetched to retain some sense in which perceptual experiences are not cognitively penetrable. One could maintain (as Pylyshyn (1999) does, for instance) that top-down influences only come into play before and after ‘early vision,’ which is itself impenetrable. For this reason, we will not assume in what follows that perception is cognitively penetrable in the strong sense. The reason why we rely on the cognitive penetrability literature is because it has a lot to say about what mental states may mediate between non-perceptual mental states and perceptual experiences.

We are trying to understand what the mediating mental state B—the state that translates between rich musical expectations and the real-time perceptual experience of music—might look like. We will consider two candidates for the mediating mental state: mental imagery and attention[[6]](#footnote-6). In each case, we will argue that (a) the evidence suggests that imagery and attention do, in fact, mediate between higher-order mental states and perceptual experience; (b) this fact counts as evidence that the interface problem does not exist for those states; and hence (c) the states in question are good candidates for our ‘mediating state’ B.

Let us take attention first. Depending on how we deploy our attention, perceptual experiences involving the same sensory stimulation can differ radically. This phenomenon is demonstrated by inattentional blindness experiments (Mack and Rock 1998, Simmons and Chabris 1999). To take a famous (or infamous) example, if subjects viewing a short video clip of a ball game are asked to count the number of times a basketball is passed amongst a group of players, they will reliably fail to see the man in the gorilla costume strolling casually across the screen. If the subjects do not attend to the counting task, they immediately notice the gorilla figure (Simmons and Chabris 1999). Attention has, in this case, a very significant impact on our perceptual experience: depending on where we focus our visual attention, we can fail entirely to perceive some components in the scene. The proposal would then be that something like this is happening in the musical case. The non-perceptual background mental state A, partly constituted by expectations (whether explicit or implicit), is directing our attention (mental state B) to some musical features in the stimulus over others. It is this directing of our attention that is responsible for the change in our perceptual phenomenology. On this view, then, background expectations guide our attention, and it is attention that directly influences our perceptual experiences. Here is a toy example: suppose that one is listening to the beginning of Beethoven’s Fifth Symphony. The experience of the first three notes (‘TaTaTa’) makes us attend to the temporal point at which the fourth note (‘Taaaam’) is expected. If the note that we end up hearing at at the attended time diverges from the way we anticipated that it would be (if it is delayed, for example, or if its pitch, loudness or timbre is different to the way we anticipated it would be), we notice this difference and experience it as salient because we were attending closely at that moment in time (see Kok et al. 2012 for empirical support for this claim).

The other candidate is mental imagery (Macpherson 2012). If I ask you to visualize an apple, the colour of the visualized apple will be determined by your non-perceptual background mental states about the colour of apples, which are determined by your previous encounters with apples. The phenomenal character of your mental imagery is very much informed by your background knowledge and expectations about the object you are imagining. Moreover, there is evidence that mental imagery can influence ongoing perceptual experience in some cases. A key finding here is the Perky experiment, where (under some interpretations of the experiment, at least) subjects’ perceptual experience was significantly influenced by their visual imagery (Perky 1910, Segal 1972). In general, there is plenty of empirical evidence that perceptual processing interacts with mental imagery at various levels (see Kosslyn et al. 2006 for a thorough summary), and also that mental imagery and perception share a vast proportion of neural circuitry. Given that there appears to be a rich network of influences from mental imagery to perceptual experience, the proposal would be that it is auditory imagery—formed in the light of our background expectations about music—that is influencing our perceptual experience. Take, again, the toy example of Beethoven’s Fifth. On this view, the listener forms a mental imagery of the fourth note (‘Taaam’) on the basis of the experience of the first three (‘TaTaTa’) (there is a lot of empirical evidence that this is in fact what happens – see Yokosawa et al. 2013, Kraemer et al. 2005, Zatorre and Halpern 2005, Herholz et al. 2012, Leaver et al. 2009). This mental imagery may or may not be conscious. And now, if the actual ‘Taaaam’ diverges from the way our mental imagery represents it (again, if it is delayed, or altered in pitch or timbre for example), we notice this divergence and experience its salience in virtue of a noticed mismatch between the experience and the mental imagery that preceded it. The same effect is achieved—we notice that things have turned out differently to the way we anticipated they would turn out, and we experience this difference as salient—but on this view of the way in which expectations affect experience, the effect has come about through a process involving mental imagery, rather than one involving attention to a specific point in time.

Two points need to be emphasized. The first is that we are assuming that both attention and mental imagery can be conscious or unconscious. In the last section we made a distinction between conscious and unconscious expectations, and claimed that both plausibly play a role in musical experience. This is consistent with a role for attention and imagery when it comes to musical expectations, given that both can be unconscious. Unconscious attention has been at the forefront of empirical research in perceptual psychology (Cohen et al. 2012, Kentridge et al. 1999, 2008). One impressive finding is the following: if an attention-grabbing image is presented below the awareness threshold in one part of the visual field, we are quicker to recognize new stimuli in that part of the visual field—a phenomenon thought to arise from the fact that our ‘unconscious attention’ has been drawn there (Jiang et al. 2006). Perceptual psychologists also discuss mental imagery as a process that can be conscious or unconscious (Kosslyn et al. 1995, see also Phillips 2014 for a philosophical argument for unconscious mental imagery).

The second important point is that these two ways of accounting for musical expectations—by means of attention and auditory imagery—are by no means exclusive. In fact, if musical expectations involve the exercise of auditory imagery (as a lot of empirical findings suggest: see, for example, Yokosawa et al. 2013, Kraemer et al. 2005, Zatorre and Halpern 2005, Herholz et al. 2012, Leaver et al. 2009), the question about what fixes the content of such auditory imagery remains. Attention may play a crucial role here (see Kok et al. 2012), as may the findings about how the musical expectations of expert musicians involve different exercise of mental imagery (see Brodsky et al. 2003, Lotze et al. 2003).

We now have a pluralistic picture of how expectations can influence perceptual experiences. We may also endorse a pluralistic view regarding what expectations are. Expectations may be considered to be beliefs that influence our perceptual experiences via imagery or attention; this is consistent with the view that the relevant expectations are located in A (the background state) rather than B (the mediating state). They may also be considered to be auditory images, or patterns of attention, by means of which our background knowledge manages to influence perceptual experiences. This view sits well with a picture whereby the relevant expectations are located in the dynamic mediating state B, rather than in the static background state A.

These different ways of understanding what expectations are and how they influence perceptual experiences become especially important when we consider different ways in which we can engage with musical listening, and the different roles expectations play therein. This is the topic to which we now turn. In the next section, we distinguish between four modes of musical engagement, considering the role of expectations in each .

IV. Four Modes of Musical Engagement

To begin to investigate whether and how expectations feature in *conscious* musical experience, rather than in unconscious processing, a preliminary taxonomy of conscious experiences of music is useful.

First, we may distinguish between online and offline experiences of music. Roughly, ‘online’ experiences of music happen when one is listening to music as it is playing. Offline experiences of music occur when there is no music playing, and one is either imagining or mentally replaying it, or perhaps just discussing it, or reading about it. Within each category, we may make the further distinction between *engaged* and *disengaged* experiences. ‘Engaged’ experiences track the music closely, whether that music is being perceived or imagined. ‘Disengaged’ experiences, on the other hand, involve a stepping back from the musical texture (where that texture is either perceived or imagined) in order to evaluate it in some way, or formulate thoughts about it, or indeed to evaluate or form thoughts about the engaged *experience* that one has just had. This pair of distinctions yields a fourfold taxonomy of musical experience: online engaged, online disengaged, offline engaged, and offline disengaged.

A few caveats before we continue. To begin with, we will focus only on online experiences in what follows, for two reasons: firstly, it is the perceptual experience of music that is our main concern; and secondly, the vast majority of experiments target the perceptual experience of sounding music rather than the experience of imagined music. Furthermore, this taxonomy is intended as a first pass: it is not necessarily exhaustive. And though we are presenting the categories as mutually exclusive—we will argue that one cannot simultaneously have, for example, an online engaged and an online disengaged experience of music—we are aware that some arguments could be made to the contrary; while we do think it most likely that online and offline engaged experiences cannot co-occur, our main purpose here is to argue for the need to hold the different modes apart conceptually.

The ‘online engaged’ experience of music is the kind of experience that most people have in mind when they refer to an episode of musical listening: it is the experience of actively tracking or following a piece of music as it unfolds, keeping up with it at every moment. This kind of active tracking, or following, is discussed at length in Levinson (1997)[[7]](#footnote-7). The ‘online disengaged’ experience of music, on the other hand, occurs when one evaluates the music, predicts what will happen next, or formulates some other explicit thought about it. Suppose that we are listening to music together, and I say to you: ‘There’s a really good part coming up.’ When I do so, I can still hear the music, but I am no longer actively tracking it; I am forming thoughts about the music rather than just immersing myself in it. It seems that one must exit the online engaged mode in order to enter the online disengaged mode. I don’t just *happen* to briefly stop tracking the music when I entertain a prediction or evaluation about it. It is a *condition* on my being able to engage with it in this way that I distance myself to some extent from the unfolding texture. Formulating higher-order thoughts about music, or making predictions about it, takes time and effort; it requires some minimal distance from the musical texture.

Expectations function differently in these two states. Take online engaged experience first. Again, the phenomenology suggests that when I am attentively tracking music in the ‘online engaged’ mode, I am not consciously expecting anything. I am just following the music. I am not consciously entertaining any expectations whatsoever, be those explicit theoretical ones, or more implicit perceptual ones. If I were to go about making predictions, or entertaining expectations, I would lose the thread of the music entirely (an argument that is, incidentally, already made by Zuckerkandl (1969/1956) as an objection to Meyer’s positing of expectations as a feature of conscious musical listening). I must enter the online disengaged mode in order to be able to consciously entertain an expectation about the music I am hearing. Now of course, it is still possible to *retrospectively* attribute the experience of expectation to myself, once I have exited the online engaged mode. I can say something like, ‘That cadence was unusual; I was expecting something else.’ But even if such a statement seems natural, it does not imply that I was really consciously expecting anything at all in the moment before the cadence resolved.

The expectations that may become conscious during online-disengaged experience can be more or less explicit, depending on the level of knowledge or expertise of the listener. If I know what sonata form is, for example, I may expect that the expository material that I am currently hearing will be followed by a development section. In online-disengaged mode, I can consciously formulate this expectation, or communicate it to a fellow listener. However, even if I do not have any theoretical knowledge about music, I can still formulate and express some expectations about the music that I hear, where these expectations are derived from my acquaintance with the piece in question, or my acquaintance with music of a similar type. In the former case, I could say something like ‘A high melody will come in shortly,’ which does not rely on my having any theoretical knowledge about music whatsoever. In the latter case, I may expect (for instance) that the verse of the song I am hearing for the first time will be followed by a chorus. I can also entertain negative expectations: even if I know nothing in particular about baroque music, I can legitimately expect that the harpsichord piece I am enjoying will not suddenly be interrupted by a solo on electric guitar. There is thus a very wide range of explicit and implicit musical expectations that I can entertain and express, no matter what my level of musical knowledge or expertise is, once I enter the online-disengaged mode of experiencing music.

Recall the fourfold distinction we made in section II. The online-engaged mode is an instance of ‘case 1’: we do not consciously entertain any expectations as we listen to music. We experience some of their effects, certainly—feelings of tension and release, or a general sense of anticipation or arousal—but we do not experience the expectations themselves. The online-disengaged mode, on the other hand, is an instance of either case 2, 3 or 4, depending on the types of expectations that a given listener is in a position to entertain, on the basis of her knowledge and expertise. A listener might consciously entertain both explicit theoretical expectations and implicit perceptual expectations (case 4), or perhaps entertain some perceptual expectations without entertaining theoretical ones (case 3). It seems less likely that, if one is in the online-disengaged mode, one might be able to experience explicit expectations without experiencing implicit ones (case 2), given that explicit expectations seem to rely on implicit, internalised patterns for their recognition—but depending on how one defines explicit and implicit expectations in the first instance, and where one judges the division between the two to be in terms of richness of content, this might be possible too. Settling this issue does not bear on the argument here.

Let us take stock. We have two modes of engagement with an unfolding musical texture: online-engaged experience and online-disengaged experience. While evidence suggests that expectations, both implicit and explicit, affectexperiences inboth modes, they are only consciously *experienced as expectations* in online-disengaged experience. We now must address what kind of mental state could serve as the mediator that we require (between the expectations, whether those are conscious or unconscious, and the respective perceptual experiences). How does this fit with the proposal about the importance of auditory imagery and attention in musical expectations?

At first blush, the literature seems to point to attention as the most obvious candidate state. Various authors (e.g. Large and Jones 1999, London 2012) have suggested that rhythmic experience (and the experience of rhythmic stresses or ‘beats’ in particular) constitutively involves attention. When we hear music, it is suggested, we attend to patterns of events in time, and our attention waxes and wanes with regularities in the music: our attention peaks at anticipated onsets of events (strong beats, for example), and slackens during periods where nothing of significance is expected to occur (in the period between beats, for instance). It is the dynamic waxing and waning of attention, goes the argument, that is responsible for our sense that certain moments in the music are accented, or ‘stronger’ than other moments. This focus on attention in the psychology literature might lead us to select attention as the most likely candidate for our mediating state.

It is not obvious, however, that attention and imagery are competing explanations, as we saw in the last section. Further, it could be argued that they cannot be easily dissociated, at least in the auditory case. While it seems apparent that one is in a position to tell, say, whether one is perceiving a rhythm or imagining one, auditory attention—or, at least, attention to periodic events—may nevertheless involve imaginative activity as a fundamental component. When I listen to a rhythm, I am only in direct perceptual contact with one sonic event at a time, and yet the entire rhythmic pattern seems to be present to me. There is a sense in which my attention reaches outwards from the present moment: I have the impression of being in perceptual contact not only with what is present at time *t*, but also with what was just present at *t-1*, as well as what I am about to hear at t+1. Given that neither the event at t-1, nor the event at t+1, is actually present to me at time t, my strong sense that I am able to attend to the overall pattern—which encompasses all three events—could plausibly involve imaginative activity.[[8]](#footnote-8) Hence, when one considers attention to rhythmic patterns, it seems difficult to distinguish between attention and imaginative capacity on either the phenomenal level or the functional level, because (1) the phenomenology of attention to periodic events seems to overlap with the phenomenology of an episode of imagination of events no longer occurring, and (2) the proper functioning of attention to periodic events seems to implicate imaginative capacity.

This, again, points to a pluralistic conclusion: that both attention and mental imagery are implicated in musical expectations, given the apparent phenomenal and functional co-involvement of attention and imagination in the auditory case. Crucially, both can be, and normally are, unconscious. In the online-engaged mode of musical listening, it is the mediation of expectations by normally unconscious imagery and attention that shapes our conscious musical experience.

V. Conclusion

Philosophers and music psychologists rarely interact. Given the overwhelming dominance of discussions of vision in the philosophy of mind, philosophers have much to gain from consideration, not only of other modalities such as hearing, but also musical listening. Musical experience is ubiquitous: most of us listen to music, hum tunes and absent-mindedly tap rhythms every day. Multiple sensory modalities interact and combine when we attend concerts, or play music ourselves (see Nanay 2012c). Moreover, musical experiences involve the exercise of some perceptual capacities that are, it is thought, almost exclusively human[[9]](#footnote-9). We also think, however, that music psychologists have much to gain from increased interaction with philosophers, when it comes to clarifying the content of musical experience itself. We hope, in this paper, to have contributed to the clarifying of some conceptual issues, in the attempt to determine what the implications of psychological explanations are for conscious musical experience.

References

Barnes, R. and M. R. Jones (2000). ‘Expectancy, Attention, and Time.’ *Cognitive Psychology*, 41: 254–311.

Bayne, Tim (2009). ‘Perception and the reach of phenomenal content.’ *Philosophical Quarterly*, 59: 385-404.

Bharucha, J. J. and K. Stoeckig (1986). ‘Reaction time and musical expectancy: Priming of chords.’ *Journal of Experimental Psychology: Human Perception and Performance*, 12(4): 403–410.

Bharucha, J. J. (1994). ‘Tonality and expectation.’ In R. Aiello (ed.), *Musical Perceptions.* Oxford: Oxford University Press, 213-239.

Brattico, E. and M. T. Pearce (2013). ‘The neuroaesthetics of music.’ Psychology of Aesthetics, Creativity, and the Arts, 7(1): 48-61.

Brodsky, W., Henik, A., Rubinstein, B. S., & Zorman, M. (2003). ‘Auditory imagery from musical notation in expert musicians.’ *Attention, Perception, & Psychophysics,* 65(4), 602-612.

Clark, A. (2013). ‘Whatever next? Predictive brains, situated agents, and the future of cognitive science.’ *Behavioural and Brain Sciences* 36: 181-204.

Cohen, M. A., P. Cavanagh, M. M. Chun and K. Nakayama (2012). ‘The attentional requirements of consciousness.’ *Trends in Cognitive Sciences* 16: 411-417.

Cuddy, L. L. and C. A. Lunney (1995). ‘Expectancies generated by melodic intervals: Perceptual judgments of melodic continuity.’ *Perception & Psychophysics*, 57: 451–462.

Delk, J. L., & Fillenbaum, S. (1965). ‘Differences in Perceived Color as a Function of Characteristic Color.’ *The American Journal of Psychology*, 78, 290–293.

Dretske, Fred (1995). *Naturalizing the Mind.* Cambridge, MA: The MIT Press.

Firestone, C. and Scholl, B. J. (2014). ‘“Top-down” effects where none should be found: The El Greco fallacy in perception research.’ *Psychological Science*, 25(1), 38–46.

Friston, K. (2005). ‘A theory of cortical responses.’ *Philosophical Transactions of the Royal Society London B Biological Sciences*, 29, 815-836.

Gandhi, S.P. et al, (1999) ‘Spatial attention affects brain activity in human primary visual cortex.’ *Proceedings of the National Academy of Sciences* 96 (1999): 3314-3319.

Goldstone, R. L. (1995). ‘Effects of Categorization on Color-Perception.’ *Psychological Science,* 6(5), 298–304.

Hansen, T., Olkkonen, M., Walter, S., & Gegenfurtner, K. R. (2006). ‘Memory modulates color appearance.’ *Nature Neuroscience*, 9(11), 1367–1368.

Hanslick, E. (1986/1854). *On the Musically Beautiful*. Indianapolis, Indiana: Hackett Publishing Company.

Herholz, S. C., Halpern, A. R., & Zatorre, R. J. (2012). ‘Neuronal correlates of perception, imagery, and memory for familiar tunes.’ *Journal of cognitive neuroscience*, 24(6), 1382-1397.

Honing, H., H. Merchant, G. P. Háden, L. Prado and R. Bartolo (2012). ‘Rhesus Monkeys Macaca mulatta Detect Rhythmic Groups in Music, but Not the Beat.’ *PLoS ONE*, 7(12): e51389.

Huron, D. (2006). *Sweet anticipation: music and the psychology of expectation*. Cambridge, MA: MIT Press.

Jiang, Y., Costello, P., Fang, F., Huang, M, and He, S. (2006). ‘A gender- and sexual orientation-dependent spatial attentional effect of invisible images.’ *PNAS*, 103(45): 17048−17052

Jones, M. R. (1981). ‘Only Time Can Tell: On the Topology of Mental Space and Time.’ *Critical Inquiry*, 7(3): 557–576.

Kentridge, R.W., Heywood, C.A., and Weiskrantz, L. (1999). ‘Attention without awareness in blindsight.’ *Proceedings of the Royal Society London B*, 266: 1805-1811

Kentridge, Robert, W., Tanja, C. W. Nijboer, and Heywood, C. A. (2008). ‘Attended but Unseen: Visual Attention Is Not Sufficient for Visual Awareness.’ *Neuropsychologia* 46: 864–869.

Kok P, Jehee JF, de Lange FP 2012 Less is more: expectation sharpens representations in the primary visual cortex. *Neuron* 75: 265-270.

Kosslyn, S. M., Behrmann, M., & Jeannerod, M. (1995). ‘The cognitive neuroscience of mental imagery.’ *Neuropsychologia*, 33, 1335-1344,

Kosslyn, Stephen M., Thompson, William L. and Ganis, Giorgio (2006). ‘The Case for Mental Imagery.’ Oxford: Oxford University Press.

Kraemer, D. J. M., Macrae, C. N., Green, A. E. & Kelley, W. M. (2005). ‘Musical imagery: Sound of silence activates auditory cortex.’ *Nature* 434: 158.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.

Krumhansl, C. L. (1995). ‘Music psychology and music theory: Problems and prospects.’ *Music Theory Spectrum*, 17: 53–80.

Krumhansl, C. L. (1997). ‘An exploratory study of musical emotions and psychophysiology.’ *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 51(4): 336–353.

Krumhansl, C. L. and E. J. Kessler (1982). ‘Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys.’ *Psychological Review*, 89(4): 334−358.

Large, E. W. and M. R. Jones (1999). ‘The dynamics of attending: how people track time-varying events.’ *Psychological Review*, 106(1): 119–159.

Larson, S. (1997–1998). ‘Musical forces and melodic patterns.’ *Theory and Practice*, 22–23: 55–71.

Leaver, A. M., Van Lare, J., Zielinski, B., Halpern, A. R., & Rauschecker, J. P. (2009). ‘Brain activation during anticipation of sound sequences.’ *The Journal of Neuroscience*, 29(8): 2477-2485.

Lerdahl, F. (2001). Tonal pitch space. Oxford: Oxford University Press.

Levin, D.T., & Banaji, M.R. (2006). ‘Distortions in the perceived lightness of faces: The role of race categories.’ *Journal of Experimental Psychology: General*, 135: 501-512.

Levinson, J. (1997). *Music in the moment.* Ithaca, N.Y.; London, Cornell University Press.

Levinson, J. (2015). *Musical Concerns: Essays in the Philosophy of Music*. Oxford: Oxford University Press.

London, J. (2012). *Hearing in time: psychological aspects of musical meter* (2nd edition). New York: Oxford University Press.

Lotze, M., Scheler, G., Tan, H. R. M., Braun, C., & Birbaumer, N. (2003). ‘The musician’s brain: functional imaging of amateurs and professionals during performance and imagery.’ *Neuroimage*, 20(3): 1817-1829.

Lupyan, G. (forthcoming). ‘Cognitive penetrability of perception in the age of prediction: Predictive systems are penetrable systems.’ *Review of Philosophy and Psychology.*

Lupyan, G., & Spivey, M. J. (2008). ‘Perceptual processing is facilitated by ascribing meaning to novel stimuli.’ *Current Biology,* 18(10): R410–R412.

Lupyan, G., Thompson-Schill, S. L., & Swingley, D. (2010). ‘Conceptual penetration of visual processing.’ *Psychological Science*, 21(5): 682–691.

Lupyan, G., & Ward, E. J. (2013). ‘Language can boost otherwise unseen objects into visual awareness.’ *Proceedings of the National Academy of Sciences*, 110(35): 14196–14201.

Mack, A. & Rock, I. (1998). *Inattentional Blindness*. Cambridge, MA: MIT Press.

MacPherson, F. (2012). ‘Cognitive penetration of colour experience.’ *Philosophy and Phenomenological Research*, 84: 24-62.

Margulis, E. H. (2005). ‘A model of melodic expectation.’ *Music Perception*, 4: 663–714.

Marr, David (1982). *Vision*. San Francisco: Freeman.

Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago; London: University of Chicago Press.

Meyer, L. B. (1967). *Music, the arts and ideas: patterns and predictions in twentieth-century culture.* Chicago, London: University of Chicago Press.

Nanay, Bence (2012a). ‘Action-oriented perception.’ *European Journal of Philosophy* 20: 430-446.

Nanay, Bence (2012b). ‘Perceptual Phenomenology.’ *Philosophical Perspectives* 26: 235-246

Nanay, Bence (2012c). ‘The multimodal experience of art.’ *British Journal of Aesthetics* 52: 353-363.

Nanay, Bence (2013a). *Between Perception and Action*. Oxford: Oxford University Press.

Nanay, Bence (2013b). ‘Is action-guiding vision cognitively impenetrable?’ In: *Proceedings of the 35th Annual Conference of the Cognitive Science* Society (CogSci 2013). Hillsdale, NJ: Lawrence Erlbaum, pp. 1055-1060.

O’Connor, D. H., Fukui, M. M., Pinsk, M. A., & Kastner, S. (2002). ‘Attention modulates responses in the human lateral geniculate nucleus.’ *Nature Neuroscience* 5(11): 1203–1209.

Patel, A. D., J. R. Iversen, M. R. Bregman and I. Schulz (2009). ‘Experimental Evidence for Synchronization to a Musical Beat in a Nonhuman Animal.’ Current Biology, 19(10): 827–830.

Perky, C. W. (1910). ‘An Experimental Study of Imagination.’ *American Journal of Psychology*,21: 422-52.

Philips, Ian (2014). ‘Lack of imagination: individual differences in mental imagery and the significance of consciousness.’ in J. Kallestrup and M. Sprevak (eds.), *New Waves in Philosophy of Mind.* New York: Palgrave Macmillan.

Pylyshyn, Zenon (1999). ‘Is Vision Continuous with Cognition? The Case for Cognitive Impenetrability of Visual Perception.’ *Behavioral and Brain Sciences*, 22: 341-423.

Schachner, A., T. F. Brady, I. M. Pepperberg and M. D. Hauser (2009). ‘Spontaneous Motor Entrainment to Music in Multiple Vocal Mimicking Species.’ Current Biology, 19(10): 831–836.

Schellenberg, E. G. (1997). ‘Simplifying the implication-realization model of melodic expectancy.’ Music Perception, 14: 295–318.

Segal, S. J. (1972). ‘Assimilation of a stimulus in the construction of an image: the Perky effect revisited.’ In: Peter W. Sheehan (ed.), *The Function and Nature of Imagery*. New York: Academic Press, pp. 203-230.

Siegel, Susanna (2006). Which properties are represented in perception? In T. Gendler & J. Hawthorne (eds.), *Perceptual Experience* (pp. 481-503). Oxford: Oxford University Press.

Siegel, S. (2007). ‘How can we discover the contents of experience?’ *Southern Journal of Philosophy* (Supp), 45: 127-142.

Siegel, Susanna (2009). ‘The visual experience of causation.’ *Philosophical Quarterly*,59: 519-540.

Siegel, Susanna (2011). ‘Cognitive penetrability and perceptual justification.’ *Noûs*, 46: 201-222.

Simmons, D. J. & Chabris, C. F. (1999). ‘Gorillas in our Midst: sustained inattentional blindness for dynamic events.’ *Perception*, 28: 1059-1074.

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

Winkler, I., G. Háden, O. Ladinig, I. Sziller and H. Honing (2009). ‘Newborn infants detect the beat in music.’ *PNAS*, 106(7): 2468–2471.

Yokosawa, K., Pamilo, S., Hirvenkari, L., Hari, R. & Pihko, E. (2013). ‘Activation of auditory cortex by anticipating and hearing emotional sounds: an MEG study.’ *PLoS ONE* 8(11): e80284.

Zatorre, R. J. and Halpern, A. R. (2005). ‘Mental concert: Musical imagery and auditory cortex.’ *Neuron* 47: 9-12.

Zuckerkandl, V. (1969/1956). *Sound and symbol: music and the external world.* Princeton, N. J.: Princeton University Press.

1. Meyer’s essays in *Music, the Arts, and Ideas* (1967), which deal with expectations in musical listening, have also been very influential. [↑](#footnote-ref-1)
2. Walton (1999) discusses such experiences in ‘Projectivism, Empathy, and Musical Tension.’ [↑](#footnote-ref-2)
3. See Siegel 2007 and Nanay 2012b on the distinction between perceptual and overall phenomenology. [↑](#footnote-ref-3)
4. Mental imagery can be distinguished from imagination: the former is a mode of the latter (see Nanay 2015). It is the role of mental imagery that we want to emphasise in this context, rather than that of imagination considered broadly. [↑](#footnote-ref-4)
5. There has been some controversy about the Levin and Banaji 2006 findings, especially about their first experiment (see, e.g., Firestone and Scholl 2014, Lupyan forthcoming). However, the experiment we want to use here is not their first but their second experiment, where the two faces presented are identical in all respects apart from the label displayed under them. [↑](#footnote-ref-5)
6. We follow the psychological and philosophical literature in taking attention to be a mental state, rather than an aspect of conscious experience (see also the remarks below about unconscious attention). [↑](#footnote-ref-6)
7. See also Levinson’s ‘Concatenationism, Architectonicism, and the Appreciation of Music,’ reprinted in *Musical Concerns* (2015). [↑](#footnote-ref-7)
8. See Levinson’s *Music in the Moment* for more on this phenomenon, as well as discussion of Husserl’s reflections on the grasp of melody as it relates to time consciousness. [↑](#footnote-ref-8)
9. To take an example, evidence suggests that the human ability to perceive the ‘beat’ in music is present from birth (Winkler et al. 2009). It has proven difficult to clearly identify other animals that can react to a musical beat as easily and accurately as we can. While some ‘vocal learning’ species can, it has been suggested, entrain to a musical beat (e.g. Patel et al. 2009, Schachner et al. 2009), EEG evidence suggests that primates—or at least some species of primates, such as rhesus monkeys—cannot (Honing et al. 2012). [↑](#footnote-ref-9)