Musicing, Materiality, and the Emotional Niche
Joel Krueger
University of Exeter

Introduction

Elliott and Silverman have produced a remarkable book. Although it claims to offer a philosophy of music education, it touches on nearly every important dimension of philosophy of music, generally construed. I cannot engage in a similarly comprehensive discussion. Instead, I will consider some themes at the heart of Music Matters. These themes are encapsulated in the following remark, which can be read as a mission statement for the book: “[T]o understand musical understanding, we need to investigate the invisible embodied-enactive scaffolding, the embodied infrastructure, that underpins and fuels all forms of musicing and listening” (Elliot and Silverman 2015, p.201). I agree with Elliot and Silverman that “musicing”—their term (which I will adopt) for all music-related activities, solitary and social—is best understood by assuming a multidimensional and interdisciplinary perspective on its embodied, enactive, and scaffolded nature.

I will focus especially on this last point: the idea that music can function as scaffolding, a persistent environmental resource supporting the development of various experiences and embodied practices. In light of this focus, I will talk about the materiality of music. This term is meant to emphasize two things: first, that music shows up for us, experientially, as something we use, something we do things with; second, this is because music is mediated by artifacts and environments that afford different uses. And one of the central ways we use music, I will argue, is to actively manipulate social space—and in so doing, manipulate our emotions. Acts of musicing, seen as processes of environmental space manipulation, might thus be thought of as examples of what I will call “emotional niche construction”. The materiality of music is what makes these practices possible.

None of what I say below is inconsistent with Elliot and Silverman’s views. In fact, I see my claims as congruent with their analysis. I simply hope to enrich their already wide-ranging discussion
by introducing some additional concepts to help bring aspects of these “invisible” dimensions of musicing into sharper relief.

The space of musicing

The dimension of musicing I am concerned with is music listening. It is tempting to think of music listening as a passive affair. Whether stopping to listen to street performers, reclining into a plush velvet seat while taking in a symphony orchestra, or streaming music at home while preparing dinner, it’s not immediately clear that these episodes involve much music-directed activity or environmental manipulation—beyond, perhaps, pressing buttons on our listening device. There is a musical event happening, one might say, but as listeners we simply absorb this event. However, as Elliot and Silverman convincingly argue, this passive characterization of musicing is misleading. For, even seemingly passive cases of musicing—such as playing music softly in the background while making dinner—still involve a deliberate use of music. We actively manipulate various features of the music and listening context: the genre, order of the tracks, volume, length of our listening episode, how closely we attend, which device we use to listen, where we position it, etc. Musicing-in-listening is thus an active process. We use music as a resource for manipulating our environment and ourselves.

This manipulative dimension of musicing becomes even clearer when we focus on the world-making power of music. At one level of description, music is nothing but vibrations moving through air. But this description has no traction at the phenomenological level. When we perceive music qua music, we perceive it as meaningful—that is, as exhibiting an organized structure, a phenomenologically rich sounding environment (Reybrouck 2015; see also Clarke 2005, Krueger 2009). As Elliot and Silverman put it, we become “engaged with the musical world we’re creating with our musical, bodily-perceptual “knowing hows”” (Elliot and Silverman 2015, p.206). One of the reasons this is possible is because, phenomenologically, music bears spatial content—and our
perceptual encounter with music involves the experience of action-soliciting musical space. Phenomenologically, we encounter music as a spatially-structured soundworld comprised of qualities that beckon for further exploration and bodily engagement (e.g., attentive focus, rhythmic entrainment, etc.).

This is apparent in the way we routinely invoke spatial notions to describe music. It is very difficult to describe music *without* invoking spatial notions. When talking about pitch organization, for example—a feature which, in contrast to other elements such as timbre, texture, or tempo, is unique to music—we rely upon spatial contrasts such as “up and down”, “high and low”, “small and large”, etc. (Morgan 1980). This is also the case when talking about our emotional *responses* to music. For example, Peltola and Saresma (2014) found that listeners commonly use spatial and movement metaphors when asked to describe their experience of listening to sad music.

Another line of support comes from studies of amusia, a severe deficiency in processing pitch variation despite normal speech perception and intact sense of rhythm (Ayotte et al 2002; Peretz et al 2002). For the total amusiac, music is experienced as unstructured noise; it sounds like a screeching car or banging bots and pans (McDonald and Stewart 2008; Sacks 2007). But amusia is more than a sensory-musical deficit. There is evidence it is also a *spatial* deficit—an inability to represent space (Särkämö et al 2009). For example, people with amusia perform significantly worse than non-amusiac controls on mental rotation tasks (Douglas and Bilkey 2007; see also Cupchik et al 2001). They also have difficulty synchronizing bodily movements with rhythmic elements and pitch variations in music despite normal ability to synchronize with non-musical sounds (Dalla Bella and Peretz 2003; Foxton et al 2006).

Phenomenologically, amusiacs lack the ability to perceive music as a spatially-structured soundworld. Amusiacs can still perceive music’s *locational* space; they can point to music as an

---

1 This appears to be the case from birth—and probably even earlier. See Krueger (2013).
2 Although there is some debate about this claim. See Tillmann et al (2010).
irritating sound sequence coming from over there, for example. However, they cannot experience the inner space of the musical piece: the syntactical structure of the music established by the way its elements (tones, melody, rhythmic progressions, textures, etc.) hang together, lending the piece its coherence as a composed object. Perceiving this inner space of music is what makes music show up as meaningful.

It thus seems that music perception involves spatial content. For most of us, the spatial content of music means that it shows up as a richly-structured soundworld inviting further exploration; it is phenomenologically manifest, in other words, as a sounding environment that draws us toward and into itself, presenting a topography of sounds, textures, and rhythmic elements to attentively unpack, explore, and manipulate.

Musical manipulations and the emotional niche

These considerations affirm Elliot and Silverman’s portrayal of music listening as a dynamic, exploratory process. In virtue of its spatiality, music is an environmental resource listeners can bring into use. And one of these uses, Elliot and Silverman note, is to “actively construct their emotional lives by deliberately making and listening to specific kinds of music, at specific times, and specific places in their lives” (Elliot and Silverman 2015, p.308). In other words, music is material for constructing an emotional niche: a soundworld deliberately used to modify, regulate, and sustain (i.e., scaffold) particular emotional episodes. How does this process occur?

I now want to expand Elliot and Silverman’s claim and consider three dimensions of this question. Constructing a musically-scaffolded emotional niche, I suggest, consists in the manipulation of (1) soundworlds deliberately engineered (2) to provide self-stimulating feedback (3)

---

3 See Clarke (2013) and Krueger (2011) for further discussion of the spatial content of music perception.
4 There is also evidence that the intensity of our emotional responsiveness to music is correlated with the feeling of being spatially immersed within the music (Vastfjall 2003).
affecting functional gain. Collective, these three dimensions help clarify how music, as material resource, is brought into use in the service of emotional action and experience.

Musical engineering

As we’ve seen, music has an inherently spatial character. The spatial content of music is one of the features that specifies the musical event as meaningful. But music is a flexible resource that can also fit into—and in so doing, reconfigure—pre-existing spaces, too.

For example, we play background music to shape the atmosphere of gatherings such as a dinner party, athletic event, or meditation session. The music need not dominate; its materiality means that it can be selectively positioned. At a formal dinner, for example, music can be deliberately played at a low volume and allowed to supplement the proceedings, providing unobtrusive background color. Alternatively, if a sedate dinner party is interrupted by up-tempo pop music blasting from speakers in the next room, guests might respond by getting up, moving to a new space, and dancing. The music in this way reconfigures the space of that environment by introducing novel sonic elements as well as novel action possibilities. Music is selectively used to manage the organizational dynamics of different environments—including an environment’s practical and normative space, as the dinner party example demonstrates. As Clarke observes, this spatial manipulation is possible because “music is inextricably bound up with that wider auditory world, since it sounds within it, incorporates environmental sounds into its own material, and (with the development of recording, broadcast and listening technologies) takes on fluid relationships with the physical and social spaces that it occupies—from practical and normative to provocative and paradoxical” (Clarke 2013, p.90).

5 There is also an important political dimension to the way that music can create and reconstruct social-spatial boundaries. See, for example, Smith (1997) on the role of jazz and rap in post-civil rights Black American culture and Cusick (2013) on music as an instrument of detention and torture.
But music can be employed in a more intimate and emotionally salient way. Music is mediated by material culture, including instruments, environments, listening technologies, and people. This means that, as Elliot and Silverman put it, music is a *product*; it is something created and consumed in private and public contexts (Elliot and Silverman 2015, pp.282-305). Technological advances have rendered music more portable—and thus *personalizable*—than ever before. Thanks to the development of the MP3 format and cheap, ultra-portable digital music players—along with streaming services that push music through mobile phones, tablets, and computers—music is now ubiquitous. We construct musical soundworlds anywhere and at any time with the push of a button or tap of a touchscreen.

Crucially for our purposes, these portable soundworlds are *self-styled* environments. We selectively engineer them to accommodate emotional requirements in real-time: we create playlists, juxtaposing artists, genres, and individual tracks in nearly limitless combinations. Alternatively, we can offload some of the decision process onto the technology. Streaming services provide personalized soundtracks based on the time of day, our desired mood or activity, previous listening habits, genre likes and dislikes, etc. The technology adapts to our interactions and adjusts itself accordingly. And the musical content we now have access to via these services is immense: vast libraries, from famous classical pieces to indigenous music or obscure Finnish Black Metal. The materiality of modern musicking affords unprecedented control over *what* we listen to, *when* we listen to it, and *how* we listen to it.

Our portable soundworlds have a direct impact on our emotional engagement with the wider auditory and social environment. For example, we use them to emotionally detach from the environment (DeNora 2013). By using headphones and playing our music loudly, we occlude hostile or unpleasant features of our situation. A noisy subway ride full of raucous passengers becomes tolerable when it fades to the background as we immerse ourselves in a favorite musical soundworld. Alternatively, we can use music to reframe that situation and open up possibilities for
reengagement. While listening to a soothing tune, for instance, the people around us, once the source of irritating conversation, might now be experienced in a new light: as fellow travelers simply getting on with their business. We can also reconfigure the social and normative space of that environment by introducing the music as a public resource. Playing music so that others hear it might solicit conversation and connection—or irritation if they find it disruptive or unpleasant.

The point of these observations is that, as a material resource, we routinely use music to personalize the public spaces we move through, as well as the way we emotionally engage with these spaces. We do so on micro (e.g., solitary experiences, such as listening with headphones) and macro (e.g., playing music at a dinner party or political rally) levels. Our portable musical soundworlds are thus powerful examples of engineered niches, nested within the wider socio-cultural and acoustic spaces of a common world.

Musical feedback

In addition to its power to manipulate space, another reason music functions effectively as a tool for emotional niche construction is because it is a persistent source of bio-regulatory feedback. By selectively engineering our environments with music, it becomes part of a self-stimulating feedback loop that drives, structures, and regulates the character and development of various embodied processes responsible for emotional action and experience.⁶

To get a firmer grip on the dynamics of this process, it may be helpful to first consider how gestures and facial expressions play a similar self-stimulating role. A common deflationary view of gestures sees them as expressing fully-formed thoughts; gestures function as fleshy props, this line of thinking goes, helpful for communicating the meaning of thoughts and intentions but in no way essential to mental phenomena they communicate. However, from the perspective of embodied and

⁶ See Krueger (2014a, 2014b) for a more in-depth discussion of this idea.
extended cognition (e.g., Clark and Chalmers 1998; Menary 2010), gestures play a more significant role in the economy of cognition in that “the gesture, the actual motion of the gesture itself, is a dimension of thinking” (McNeill 2005, p.98). Gestures may even provide material resources that grant access to otherwise-inaccessible forms thought (Clark 2013; Goldin-Meadow 2003).

There is evidence, for example, that inhibiting gesture during mathematical and memory tests significantly reduces overall performance (Goldin-Meadow et al 2001; Wagner et al 2004). Instead of working through difficult problems exclusively in our head, gestures provide an expanded set of representational formats driving the process along. We offload certain computations onto the visuospatial dynamics of gestures; we use them to redundantly reflect verbal information or “augment that information, adding nuances possibly only through visual or motor formats” (Golding-Meadow 2003, p.186). This offloading frees up internal resources for other tasks and prompts real-time recognition of new problem-solving possibilities.

Gestures and facial expressions seem to play a similar self-stimulating role in the realm of emotions. There is abundant evidence that afferent feedback from muscles and skin when making emotion-specific facial expressions (smiling, frowning, etc.) shapes the qualitative intensity of the emotion (see Laird 2007 and Niedenthal 2007 for overviews). Simply making a facial expression is sufficient to bring about the associated experience; once present, further expressions modulate its ongoing development and felt character (Adelmann and Zajonc 1989). Other support comes from deficit studies. When this embodied feedback is missing—in the case of congenital facial paralysis, for example (Cole 1997), or in individuals who’ve voluntarily received muscle-inhibiting Botox injections (Davis et al 2010)—the phenomenology of the emotion is significantly diminished (Krueger 2012, pp.157-160). As with gestures and cognition, the expression appears to be a material dimension of the experience. Without it, we cannot access the experience, at least in its full qualitative intensity.
So how might music play an analogous self-stimulating role? Music is an ultra-portable resource we carry with us and use to construct on-demand soundworlds. As such, like gestures and facial expressions, it is a persistent material feature of our expressive repertoire. And by stimulating and directly modulating emotion-specific neural and physiological responses, music functions as a real-time emotion regulator, coaxing emotions out of us and shaping their dynamics as they unfold in real-time.

There is a great deal of empirical evidence to support this picture. As we’ve already discussed—and as Elliott and Silverman amply demonstrate—music listening is a richly multimodal, active experience (Janata et al 2012; Witek et al 2014). Even seemingly passive listening episodes involve brain activity in the supplementary motor area, mid-premotor cortex, and cerebellum (Chen et al 2008). The motor system is central to processing musical rhythms (see, e.g., Sakai et al 1999; Grahn and Brett 2007; Kohler et al 2002); other studies suggest that the experience of agency and bodily movement plays a crucial role in shaping how music shows up for the listener (e.g., Phillips-Silver and Trainor 2005). There is a sense in which we literally hear music through movement; we perceptually approach music as something that affords motor potentialities. These potentialities frame our experience of music as a resource that can be brought into use.7

But music also pulls motor and neurophysiological responses out of us, too, that directly modulate our emotional experience and action (Burger et al 2013). Studies using facial electromyography (EMG) found that listeners exhibit spontaneous facial expressions mimicking the expressive character of the music (e.g., smiling in response to upbeat music, frowning in response to sad music), including non-vocal music (Chan et al 2013; Lundqvist 2009; Witvliet and Vrana 1996). As noted above, these overt physical expressions elicit the associated emotional experience (Laird 2007; Niedenthal 2007). Music also exerts a direct modulatory impact on other components of the

7 Again, the experience of these motor potentialities are what seem to be missing from the amusic’s experience—hence their anomalous musical phenomenology.
emotional response system (Bartlett 1996; Lundqvist et al 2009; see Scherer and Zentner 2001 for an overview). When we listen to music judged to be happy, sad, or fearful, for example, an array of physiological responses unfold that are very similar to responses that occur when we experience these emotions in non-musical contexts (Krumhansl 1997). Nyklicek et al (1997) found evidence of cardiorespiratory differentiation of musically-induced emotions (happy, sad, serene, agitated, etc.). And as far as the brain is concerned, music perception and emotional responsiveness are integrated processes. Various neuroimaging studies indicate that our emotional responses to music recruit core brain structures involved in initiating, generating, detecting, maintaining, and regulating emotions (Koelsch 2010; see also Blood et al 1999; Koelsch 2014; Overy and Molnar-Szakacs 2009).

These studies (and many more like them) present an emerging picture of music as a potent resource for driving, organizing, and sustaining emotional experience and behavior. We play music—we selectively engineer a self-styled soundworld—and then by temporarily inhabiting this soundworld let music take over some of the bio-regulatory work on our behalf. In taking over this regulatory role, music functions much the way that input from a skilled dance partner extends, transforms—and in an important sense, completes—the shape of our own responses. In fact, our engagements with music are, in many ways, perceptual dances. They are reciprocal interactions constituted by the sensorimotor dynamics of how we both approach the music as well as how the music, in turn, approaches and engages with us (Windsor and de Bézenac 2012). Musical feedback thus scaffolds emotional action and experience. Its dynamic qualities (periodic modulations of melody, rhythm, volume, intensity, etc.) and temporal character make it well-suited for this task. Elliot and Silverman put this last point well when they write that music listening “is experiential in the full-blooded sense of phenomenological couplings between our entire beings (not just our brains) and the meaning-making opportunities or “affordances” of pieces of music that flow in time and across time” (Elliot and Silverman 2015, p.205).
**Musically-driven functional gain**

So far, I’ve considered the materiality of music and its role in constructing an emotional niche. I’ve argued that music is something brought into use—a material resource that can be selectively manipulated in various user-specific ways. These manipulations, I’ve argued further, are important in this context because they exert a downstream modulatory impact on the listener’s ongoing patterns of emotional action and experience. In manipulating music, we simultaneously manipulate not only the environment but ourselves. The music we manipulate loops back onto us and, as active and responsive listeners, shapes our further manipulations and embodied responses. Something like this is what Elliot and Silverman seem to have in mind when they speak of “phenomenological couplings” between music and listener.

But there is still a third dimension to this process that helps explain why we are driven to enact these music-listener “phenomenological couplings”. Simply put, when we “couple” with music, we access emotional capacities and experiences that, without the music’s ongoing input, remain otherwise inaccessible—again, much like a novice dancer can only perform advanced moves with a skilled partner. To borrow a term of art from cognitive science: as musically-coupled listeners we realize functional gain.⁸

Once more, there are various lines of empirical evidence to support this picture. From birth, we are motivated to enact musical couplings (Adachi and Trehub 2012; Dissanayake 2013; Zentner and Eerola 2010). Newborns recall musical stimuli perceived during their fetal period (e.g., their mother singing “Twinkle, Twinkle, Little Star”) (Partanen et al 2013) and, after birth, actively seek out music. They appear to recognize it as a resource that can be brought into use. For example, infants perceptually discriminate musical from non-musical sounds, coordinating their responsive behavior with the former but not the latter (Trainor and Heinmiller 1998; Zentner and Kagan 1998; Nawrot 2003). Even very young infants, including preterm infants, appear to experience music as a resource

---

⁸ See Wilson (2010); see also Krueger (2014b).
affording bodily engagement. They entrain various embodied responses—respiratory patterns, sucking (both rhythm and intensity), tongue and mouth protrusion, eye opening and closing, vocalizations, etc.—with rhythmic and melodic patterns in lullabies and consonant music (Haslbeck 2014; Teckenberg-Jansson et al. 2011).

Within these early episodes of musicing, infants realize functional gain. This is clear when we consider that, for the first months of life, infants lack endogenous resources to self-regulate attention and emotion (Posner and Rothbart 1998). What they attend to and experience is largely determined by what is going on around them. Accordingly, in noisy or disruptive environments—such as the neonate intensive care unit (NICU), where young infants are confined to incubators, devoid of touch, and continually surrounded by beeping and whirring machines, bottles clanking, voices of nurses and doctors rushing to and fro, the crying of other babies, etc.—infants are often in a chronically distressed state, both physiologically and behaviorally.

This changes when they become phenomenologically coupled with music. For example, when a caregiver sings lullabies while stroking the infant or plays gentle music with a simple, predictable structure, the infant realizes an elevated bio-regulatory competence. The music functions as a stabilizing environment, “modulating the array of physiological states and micro-behaviors associated with instability into an array associated with stability—stable heart rate, blood pressure, colour, feeding, changes in posture, muscle tone, less frantic movements, rhythmic crying, cessation of grimacing, and an ability to sleep or become animated and intent” (DeNora 2000, p.81). By latching onto the music as it unfolds in time, musically-scaffolded infants realize cognitive and emotional capacities that, outside of this transient soundworld, remain otherwise inaccessible.

What about mature perceivers? One important dimension of this functional gain is access to an *expanded phenomenological repertoire*. In using music as a resource to construct our emotions in real-time, we gain access to an enriched palate of expressive capacities lifting us beyond our normal modes of experience and expression.
Again, DeNora’s work is instructive, particularly her analysis of music as an on-demand resource for emotional “venting”, as she terms it. One woman DeNora interviewed says that listening to specifically-chosen tracks while sad is like “looking at yourself in a mirror being sad” (DeNora 2000, p.57). The music provides self-stimulating resources enabling the woman the articulate and explore her sadness in a new way, she says, guiding her into a qualitatively deepened state before leading her out of it (DeNora 2000, p.57). Another woman reports using music as a tool for working through her grief upon losing a young child: “The Verdi Requiem is one of my favorites. That is associated with losing a baby. And I’d got to know it through my husband and it was really quite a way of grieving—I’d shut myself in my room [she begins to cry]...It’s cathartic, I think” (DeNora 2000, p.58). Reports of using music to construct and explore emotions this way are not uncommon. As DeNora puts it—in language similar to Elliot and Silverman—venting with music is to use music “as a virtual means of expressing or constructing emotion...to define the temporal and qualitative structure of that emotion, to play it out in real time and then move on” (DeNora 2000, p.58).

These musical-emotional constructions are possible, I suggest, because music possesses expressive qualities that are more agile, evocative, and nuanced than are their behavioral counterparts (e.g., gestures, facial expressions, etc.). In contrast to a facial expression or gesture, say, musical expressions possess increased complexity, temporal range, subtlety, and force (Cochrane 2008, p.338). When we construct and inhabit our self-styled soundworlds, we thus (temporarily) gain access to these expressive capacities—and music functions as a pathway scaffolding access to new forms of experience and expression. In sum, we bring music into use in constructing a musically-scaffolded emotional niche.

**Conclusion**
As I said at the outset, nothing I’ve discussed, as far as I can tell, goes against the general orientation Elliot and Silverman adopt in *Music Matters*. Rather, I’ve simply tried to bring out some additional themes that supplement their rich embodied and situated approach to musicing. However, if the previous analysis is on the right track, it seems that there are additional conceptual resources in the 4E approach to cognition (i.e., cognition as *embodied, embedded, enactive*, and *extended*) that Elliot and Silverman might find theoretically fruitful. A growing number of researchers in philosophy of mind and cognitive science defend some version of the *hypothesis of extended cognition* (HEC) (Clark and Chalmers 1998; see also Hurley 1998, Menary 2010, Rowlands 2010).

According to HEC, some cognitive processes extend across brain, body and world. Environmental resources—when coupled with neuronal and bodily processes in the right sort of way—become constitutive parts of the associated cognitive process and, in so doing, open up access to otherwise-inaccessible forms of cognition and intelligent behavior.

Although discussions of HEC have generated much debate, few working in this area have discussed emotions—and even fewer aesthetic experiences. But this is a serious omission. Emotions and artistic practices are fundamental dimensions of human experience. As Elliot and Silverman so convincingly demonstrate, musicking—in all its variegated forms—is an embodied practice thoroughly interwoven into nearly every facet of everyday life. As I’ve tried to argue above, there are compelling reasons to think of music listening as a practice that’s not just embodied and situated but also environmentally extended. Elliot and Silverman’s analysis is well-placed to make an important contribution to this emerging approach. I therefore encourage them to explicitly embrace the final “E” and turn their considerable resources to an analysis of the musically extended mind.

---

9 There are some recent exceptions. See, for example, Colombetti and Roberts (2015), Krueger (2014c), Slaby (2014), Stephan et al (2014), and the essays in Manzotti (2011).
References


