

Music and Cognitive Extension

LUKE KERSTEN [1]

Institute of Cognitive Science, Carleton University, Ottawa, Canada

ABSTRACT: Extended cognition holds that cognitive processes sometimes leak into the world (Dawson, 2013). A recent trend among proponents of extended cognition has been to put pressure on phenomena thought to be safe havens for internalists (Sneddon, 2011; Wilson, 2010; Wilson & Lenart, 2014). This paper attempts to continue this trend by arguing that music perception is an extended phenomenon. It is claimed that because music perception involves the detection of musical invariants within an “acoustic array”, the interaction between the auditory system and the musical invariants can be characterized as an extended computational cognitive system. In articulating this view, the work of J. J. Gibson (1966, 1986) and Robert Wilson (1994, 1995, 2004) is drawn on. The view is defended from several objections and its implications outlined. The paper concludes with a comparison to Krueger’s (2014) view of the “musically extended emotional mind”.

Submitted 2014 February 22; Accepted 2014 October 10.

KEYWORDS: *music perception, extended cognition, wide computationalism, ecology*

EXTENDED cognition is the view that cognitive processes sometimes leak into the world (Dawson, 2013). A recent trend among proponents of extended cognition has been to put pressure on phenomena thought to be safe havens for internalists (Sneddon, 2011; Wilson, 2010; Wilson & Lenart, 2014). By arguing that topics like visual perception or moral cognition no longer mark a clear divide between internalist and externalist views of mind, proponents of extended cognition have sought to “loosen the screws on the individualist skullcap” (Wilson, 2010, p.277). This paper attempts to continue this trend. It is argued that music perception is part of a locationally wide or extended computational system. This extended view of music perception is a marriage of J.J. Gibson’s (1966, 1986) ecological theory of perception and Robert Wilson’s (1994, 1995, 2004) notion of “wide computationalism”. It shares with ecological approaches an emphasis on environment-perceiver interactions (Gibson, 1966, 1986; Neuhoff, 2004; Clarke, 2005), while it shares with computational approaches an allegiance to cognition as computation (Pylyshyn, 1984).

To motivate the present project a bit more, consider a brief example from the history of cognitive science. Problem solving was traditionally thought to involve a search through problem space (Newell, Shaw, & Simon, 1960). For many, this meant that cognizing relied exclusively on in-the-head resources (Dawson, 2013). However, such a view neglected certain important features of problem solving. For instance, that cognizers often interactively explored problems by physical manipulating external structures (Kirsh, 2009). Such actions are more than just pragmatic. They crucially help cognizers to simplify and solve problems. Internalist approaches revealed little about these epistemic actions, and therefore how people actually solved problems, because they failed to connect to real-world practices. Some phenomena strained internalist explanations, because what’s inside the head is often supplemented by what’s outside. An extended account can therefore sometimes offer a superior explanatory account (see, e.g., Wilson, 1995, 2014).

ECOLOGICAL ACOUSTICS

To arrive at an extended view of music perception, two questions need to be answered: How can an acoustic array carry information? How is it that computation relates to extended music perception? Answering these two questions will help guide discussion.

To begin, consider that at any given moment environments are filled with vibratory events, e.g.,

slamming car doors, crying babies, chirping birds, etc. Acoustic events are constantly emitted from mechanical disturbances, refracted from objects and surfaces, and diffused through gases, liquids, and solids. Sonic energy converges from all directions. In many ways, the arrangement of the environment and the medium of transmission shape the vibratory fields listeners exist within. Correspondingly, listeners also occupy a location or space within these sonic environments. They are localizable with respect to the propagating compression waves. It can be said, then, listeners are embedded within an “acoustic array”. The acoustic array is analogous to Gibson’s (1966, 1986) “ambient optic array”, which is a field of diffused and refracted light in which an observer might be positioned.

In studying the acoustic array it is most profitable to adopt an ecological approach to perception. Gibson offers a helpful description: “The environment described is that defined by ecology. Ecology is a blend of physics, geology, biology, archeology, and anthropology, but with an attempt at unification...what can stimulate a sentient organism” (1966, p.29). This approach can be contrasted with traditional psychoacoustics. Whereas ecological acoustics investigates the bidirectional relationship between the acoustic environment and the organism, psychoacoustics is often exclusively concerned with studying how sound stimulates the perceiving organism (Deutsch, 2013).

One notable feature of the ecological approach is that it rejects the cognitivist or constructivist assumption that stimuli are impoverished. Organisms do not perceive internal representations; they perceive invariant features of the environment (Neisser, 1999). Instead of using information carried within the organism (e.g., retinal images), perception uses information contained within the environment itself. For example, texture gradients remain constant as perceivers move through their environment. Because of this, equal amounts of texture represent equal amounts of terrain. As the density of optical texture increases, the scale of space is more easily defined. Texture gradients are therefore invariant features that offer information to the visual system (Gibson, 1986, p.67). From the ecological perspective, perception is made possible by organisms’ attunement to the invariants of the environment. In adopting an ecological framework, music perception is conceived of as the search for “musical” invariants (Balzano, 1986; Neuhoff, 2004, ch.10).

MUSIC PERCEPTION AND INVARIANTS

It is time for an answer to the first question. The acoustic array offers information about the identity of and location of vibratory events. These two kinds of information are delivered by two sonic structures: Wave trains and wave fronts. Wave fronts are the concentric spheres that eminent in all directions from a vibratory source; wave trains are the mixed frequencies that exist along the radius of the wave front. Whereas the former affords localization and orientation, the latter offers discrimination and identification (Gibson, 1966, p.81).

Consider localization first. If, for example, the head is at a 30° right angle to a vibratory event, a wave front will enter the right ear slightly earlier and stronger than it will the left. There is stimulus asymmetry. This will result in the auditory system immediately adjusting to compensate for the stimulus imbalance, for example, by turning the head. As the auditory system orients and adjusts toward the source, the intensity of the sound increases. There is reciprocal feedback. As Gibson remarks: “Each ear seems to orient by turning or twitching so as to funnel the maximum amplitude into its eardrum” (1966, p.83). It is the asymmetry of stimulus that explains how wave fronts carry information: through orientation of the auditory system, changes or transformations in the wave front specify information about sound location.

The story for musical identification and discrimination is a bit different. The standard mode of analysis in psychoacoustics takes sounds as abstract combinations of simple sinusoidal waves, whereas music perception involves the processing and construction of internal representations (Clarke, 2005, p.16). But as Gibson (1966) noted, one issue with this approach is that it neglects variables that relate to transitions and tempo. Pitch and amplitude dimensions of acoustic stimulation fail to correspond neatly to frequency and loudness, especially for more meaningful, non-artificial sounds. Because sounds can change in tonal and rhythmic quality, in duration, and in rate of change of loudness, these dimensions can be combined to produce higher order variables. It is these higher order variables that are crucial to the ecological study of auditory and music perception (Gibson, 1966; Balzano, 1986).

Consider a non-musical example. Speech sounds are remarkably constant across changes in pitch, loudness and duration. For example, the ratios in vowel patterns remain constant whether spoken in a low or high register. Because the phonemic units of speech have transposable sequential patterns across frequency, intensity and time, it can be said that speech sounds carry invariant information about phonemes

(Gibson, 1966, p.93). Wave trains specify the information about the identity of sounds because there are higher order constants embedded within sonic patterns. It's worth noting the notion of information being used here is one of "natural information". This is the view that mediums or vehicles carry information insofar as their variables reliably correlate with other variables (Piccinini & Scarantino, 2011, p.21). The acoustic array carries information because the front and train waves reliably correlate with variables of mechanical disturbances and transposable sonic patterns. Furthermore, note that whereas the localization requires bodily orientation, identification can be achieved while stationary. That is, it can be accomplished by either half of the auditory system, though this not to say that orientation cannot enhance identification.

The question now is how we identify music qua music, or what invariants the perceptual system uses to pick out music. Consider Balzano's (1986) answer: there are objective structural pitch-time constraints that allow listeners to perceive music as music. To get a grip on pitch-time constraints, consider the difference between speech and music sounds. Whereas there are an infinite number of values for pitch and time in speech (recall, for example, how vowels remain constant across changes in pitch), there are a relatively small number of values in musical pitch-time, and of those values there are specific relationships that hold. As Balzano puts it, "there is a difference in the presence of specific constraints in the global selection of pitch and time values" (1986, p.218). In the case of musical sounds, pitch-time constraints generate two types of invariant properties: "quantal" and "generative". *Quantal properties* refer to the relatively small number of values of pitch and time in music. *Generative properties* refer to specific relationships between pitch and time values (Balzano, 1986, p.218). Melodies that have diatonic pitches violating octave equivalence are examples of the former because they involve reductions in the continuous frequency domain of equally spaced elements. Melodies that have two octave diatonic scales abiding by octave equivalence are examples of the latter because they include specific constraints on the relationships between the elements of the reduced frequency domain—generative properties are built on top of quantal properties.

How do these pitch-time values figure in music perception? To answer this question, Balzano had subjects listen to 38 pseudomelodies and then attempt to decide whether the sounds were more or less musical. Participants either heard pseudomelodies that used pitch pairs with identical rhythm (time structure) and different pitch or time pairs with identical pitch and different rhythm (time structure). The pseudomelodies were 70-notes in length, varied in tone duration between .14 to .475 sec, and had a two octave range between A3(220Hz) and A5(880Hz). Balzano compared the pitch-time value relationships across several conditions. In a low condition, pitch-time values had unlimited range; in a medium condition, pitch-time values were quantized but without a generative relation; and in high condition, pitch-time values were quantized with a generative relation. This experimental design meant that the high constraint conditions included quantal and generative properties, the medium conditions included only the quantal property, and the low conditions included neither.

The underlying idea was that if listeners were not responsive to the quantal and generative properties (the musical invariants), then the pseudomelodies would have been identified as musical sounding to roughly the same degree across all three-constraint conditions. However, the pitch-time constraint conditions that included the quantal and generative properties showed significantly higher effects on musical identification. When pitch-time values were quantized with generative relations, as they were in the high and medium conditions, perception of pseudomelodies as musical increased—that is, the sounds were identified as more musical sounding (Balzano, 1986, p.227). These results are important for two reasons. First, they suggest that musically untrained participants resonate to higher order or invariant properties of musical sounds. Second, they suggest that music perception does not only involve the construction of musical information from impoverished stimuli. Music perception also fundamentally involves the "pick up [of] structural constraints of the sort that distinguish music from non-music" (Balzano, 1986, p.233).

MUSIC PERCEPTION AND WIDE COMPUTATIONALISM

So far an answer to the first question of how an acoustic array carries information has been offered, but the second question of how computation relates to extended music perception remains unaddressed. To answer it, appeal can be made to Robert Wilson's (1994, 1995, 2004) notion of "wide computationalism".

The basic idea is quite straightforward, and has been acknowledged, in principle, by several authors (Segal, 1997; Piccinini & Scarantino, 2011). Computational systems are wide just in case some of the computational units are not wholly instantiated in the head. "[W]hy think that the skull constitutes a

magic boundary beyond which true computation ends and mere causation begins? Given that we are creatures embedded in informationally rich and complex environments, the computation that occur inside the head are important part but are not exhaustive of the corresponding computational system” (Wilson, 2004, p.165). Wilson here is emphasizing the location neutrality of computational descriptions. If the method of computational analysis is locationally silent, then computational systems can be wide or extended. Nothing about computational individuation requires computational systems to be only in the head. The promise is that if cognition is computation, where computational systems are wide, cognitive systems are extended.

To illustrate, consider one of Wilson’s examples: the multiple spatial channel theory of form perception of Sekuler and Blake (1990). According to this theory, an organism’s visual system has channels that decompose any visual scene into four parameters: orientation, spatial frequency, contrast, and spatial phase. “On this conception of form perception, part of the task of the perceptual psychologist is to identify formal primitives that adequately describe the visual environment” (Wilson, 1994, p.363). For Wilson, the multiple spatial channel theory can be thought of as involving a wide computational system because the inputs to the computational processes involve environmental elements. However, even granting the *prima facie* coherence of Wilson’s example, the question is how to identify wide computational systems, and, more pertinently, whether it can be determined that music cognition implements such a system.

How is it that one can know when the relations between an organism and environment qualify as computations? Consider how this is normally done for an internal psychological process. To computationally model some in-the-head process, one first must identify and formalize the relevant primitive states. This allows the process to be broken down into component features. Then, one must describe the changes between the states in terms of transition rules; they must be given a function-theoretic account. For example, in Marr’s (1982) theory of vision the retinal image and the internal 3-D representation of the environment are the relevant input and output states, while the transition rules are computational processes that govern the transformations of the information contained within the retinal image to the 3-D representation. Identifying wide computational systems requires: (i) identifying and formalizing specific properties of the environment that an organism is sensitive to, (ii) decomposing natural scenes into the parameters set out by the formal primitives, and (iii) specifying the algorithms or rules that apply to the identified primitives governing an organism’s behaviour. The important point to note is that where the informational primitives of the system could be internal, they could also be environmental.

Consider, then, what happens when musical invariants of the acoustic array are taken as the primitive states of the computational analysis for music perception. First, the musical invariants can be identified and formalized as computational units given that they are external information-bearing states. This means they can supply the inputs to the computational processes. This is requirement (i). Second, as Balzano’s research showed, the auditory system is directly responsive to musical sounds; it responds as a function of the presence of external music information structures. This means that musical events or scenes can be decomposed into the invariant constraints. This is requirement (ii). Finally, there seems to be a lawful causal relationship between the physical structures of sounds, i.e., musical invariants, and the stimulation of the auditory system. This is requirement (iii). Thus, all three of the above requirements for computational characterization are met. It’s worth mentioning, though, that describing the causal relationship in (iii) is not an easy task. All that’s been done here is gesture at its plausibly. Nonetheless, given the direct responsiveness of listeners’ auditory systems to the presence of musical invariants, it does seem plausible that such a lawful relationship exists and that it could be described in more detailed terms. Given all this, it seems that music perception can be said to involve computational processes ranging across environmental and in-the-head elements. Because the musical invariants are used by downstream parts of the computational system they become integrated into, incorporated as parts of, a wide computational system. Music perception can be considered part of an extended computational system.

However, more needs to be said about why the in-the-head plus musical invariants can be given a computational characterization and why the wide computational system can be thought of as genuinely cognitive or perceptual. The former of these worries can be dealt with now, the latter postponed until later. There are two reasons to think that the acoustic array is a good candidate for computational analysis. First, its invariants persist through time; second, its invariants have structure in virtue of which they carry information. As Piccinini and Scarantino (2011, p.30) point out, as long as a medium or vehicle has persisting informational structure it can be given a computational characterization.

To illustrate, consider again the analogy to texture gradients. Light is constantly diffused and reflected throughout the environment. Because of this, there are textures overlaying surfaces. As perceivers

move through environments, the amount of texture corresponds to the amount of terrain. In effect, this means that as the density of optical texture increases the scale of the space is specified or revealed. Optical texture provides information about the environment. Analogously, musical invariants such as pitch-time constraints have order and equivalence relations that remain constant through flux or transformations (see Balzano, 1986, p.227 or Trehub et al., 1984, p.828); they have persistent structure. What's more, music perceivers, as has been shown, resonate or detect the information contained or carried by the invariants' persistent structure. The musical invariants of the acoustic array are apt to be included within computational analysis because they are the right kind of external information-bearing states.

Having answered the two orienting questions with which this discussion began, it's time to outline the argument for extended music perception in full. Schematically, it looks like this:

- (i) Cognition is a species of computation.
- (ii) When the elements of a computational system include parts of the organism and parts of the environment, the computational system extends beyond the individual and into the world.
- (iii) Music perception involves the detection of musical invariants within the acoustic array.
- (iv) The interaction between the auditory system and musical invariants of the acoustic array is characterizable as a wide computational system.
- (v) Therefore, music perception is extended.

One brief caveat, Shapiro (2011) notes that extended and embodied views of cognition are sometimes seen as replacing the need for the internal representations. Since organisms interact with their environments in deep and sustained ways, it is suggested that representations can be lessened if not dropped—enactive and dynamical views of cognitions sometimes make these claims (see, e.g., Chemero, 2011). Thus, it might be thought that an extended view of music perception similarly removes the need for internal representations. This is not the case. An extended view of music perception does not replace representational views. Rather, it augments them. Internal representations can exist within wide computational systems; nothing in what has been said denies this. Some computational systems will be fully instantiated within the individual; others will cross into the world. What has been shown is that the music perception does not *solely* have to be conceived of as an entirely internal process.

OBJECTIONS AND IMPLICATIONS

There are three places the present argument can be challenged: First, premise (ii) might be challenged by arguing that the informational invariants are, in fact, not external to the perceiver; second, premise (iii) might be challenged by questioning the coherence of wide computationalism; third, premise (iv) might be challenged by putting pressure on the connection between computation and cognition.

First, it might be suggested that Balzano's results can be given an "internalist" reading. One might claim that the pseudomelodies remind perceivers of music because they involve the retrieval of internally stored musical knowledge. Construed this way, the participants do not detect musical invariants in sound; rather, they retrieve and compare the pseudomelodies to pre-existing information. Thus, there is no reason to posit pitch-time constraints as objective, invariant features of the acoustic array. However, this objection holds little water. As Balzano points out, "this [familiarity] notion will not work; even the most highly constrained pseudomelodies were strongly 'atonal' in their pitch structure and exhibited no longer-range temporal periodicities of the sort that permit measures to arise. In both these respects, the pseudomelodies were grossly unfamiliar to subjects" (1986, p.227). For the internalist account to be true, there would have been no difference between the three conditions, yet there is. For example, in the 'low' condition—the condition unstructured by the pitch-time constraints—there was a notable drop in participants' discrimination of pseudomelodies as musical. This means that the perceivers were responding to the invariant information generated by the pitch-time constraints in the medium and high conditions.

Second, Segal (1997) has suggested that wide computationalism problematically fails to specify whether the inputs to functions are environmental properties or internal representations. For example, of the original example Wilson says that spatial navigation (what's known as 'dead reckoning' for some animals), takes as its inputs "solar heading, forward speed, and a representation of the solar azimuth, producing as output a representation of the creature's position relative to some landmark" (1994, p.366). Whereas the first two (solar heading and forward speed) are properties of the environment, the "representation of solar

azimuth” is a property of the organism. Thus, it is not clear how properties of the environment *and* representations can both be used in the computations of a system. However, noticing that the informational invariants are wholly external environmental structures can allay this concern. It can be conceded that computations cannot involve both representations and environmental properties, but still held that in the present case the computations take as their inputs only environmental elements. If informational invariants are external structures, characterizations of them as inputs do not require further reference to internal representations. Wilson hints at this response when he says: “Once we take this step [of acknowledging the importance of arrays] the interaction between the information processing structure inside organisms and information bearing states outside of them becomes central to a computational account” (2004, p.171). The worry that wide computationalism problematically slides between inner and outer information-bearing states ceases to be a problem when the invariant information in the acoustic array serves as the input to the extended music perceptual processes.

Third, and finally, while it may be the case that cognition is computation, it certainly isn’t the case that all computation is cognition. Almost anything can be given a computational description (Stufflebeam, 1999; Edelman, 2008). Though it’s possible that music perception can be modeled computationally, this does not necessarily imply that music perception is a wide computational system. Here the question is not so much whether physical systems can be given computational descriptions, but whether the physical systems themselves can be thought as performing computations. In other words, there is a distinction between computational *explanation* and *description*. Finding out whether a physical system can be formalized and captured by transition rules is a difficult task, but it is precisely part of the task undertaken here. Balzano’s investigation helps to show that important aspects the musical environment *can* be formally described. Because the invariant information is delivered via the wave train to the auditory system, this shows that physical components of the environment plus in-the-head components can be thought of as performing computations. One way of looking at the present argument is as evidence for viewing music perception as a computational explanation of a wide computational system. This response also addresses the second question postponed earlier. When computational analysis goes beyond description and offers explanation, phenomenon, in this case music perception, can be thought of as genuinely cognitive. Though these responses have not definitively defended extended music perception, the view should be beginning to look a little more plausible.

What, then, are the implications of adopting an extended view of music perception? First, an extended view of music perception begins to reorient the study of music perception toward the acoustic array and its musical invariants. Unfortunately, Balzano’s account is one of only a handful of studies that attempt to specify and formalize musical invariants—Trehub et al. (1984) offers another example. If other aspects of music cognition are to be given extended accounts, further theoretical and empirical work will need to be done. Part of the problem is that invariants are quite complex. They are difficult to isolate and study. What’s particularly unfortunate about the current state of affairs is that there is wealth of formal musical analysis, but little that looks at music as it constructed and perceived in non-artificial contexts (for an example of musical analysis in artificial contexts see Deutsch (2013, ch.2)). Second, as hinted at earlier, by adopting a hybrid computational, neo-Gibsonian approach an explanatory pluralism can be taken with regard to music cognition. Little about what’s been said requires calling into question previous or current internalist research. It may well be that much music cognition is instantiated solely in the head, but what this approach does offer is a way of understanding how music cognition can be both partly instantiated internally and in the world.

MUSIC AND COGNITIVE EXTENSION

In rounding off discussion of music and cognitive extension, it is helpful to relate the current view to Krueger’s (2014) view of the “musically extended emotional mind”. Krueger’s account begins with the observation that one of the main reasons people listen to music is to regulate emotions. For Krueger, this is achieved by “resonating” with musical structures (2014, p.3). Music offers or affords particular kinds of entrainments or synchronizations, Krueger calls these “musical affordances”. Krueger claims that because of entrainment listeners “scaffold” their endogenous capacities onto musical affordances. He writes: “If we take seriously the possibility that certain environmental resources can scaffold the emergence of extended emotions [...] we ought to take seriously that music is a particularly powerful example of an emotion-extending resource” (Krueger, 2014, p.9). Through scaffolding, the emotional mind becomes extended. Krueger nicely summarizes his position when he writes: “[M]usical affordances provide resources and

feedback that loop back onto us and, in doing so, enhance the functional complexity of various motor, attentional, and regulative capacities responsible for generating and sustaining emotional experiences. It is thus sensible to speak of the musically extended (emotional) mind” (2014, p.4).

Though Krueger never explicitly lays out his argument, I parse it as the following:

- (i) Music offers particular cognitive and motor engagements (musical affordances)
- (ii) Listeners become entrained to musical affordances such that they form an integrated system
- (iii) The integrated music-listener system actively drives and augments various emotional and attentional regulatory processes
- (iv) Therefore, music - or more specifically, musical affordances - extend emotional and attentional regulatory processes

Much of Krueger’s focus is on establishing premises (i) and (ii). An impressive array of studies is marshaled in support. However, while much is offered in defense of premises (i) and (ii), little is said about premise (iii). This is worrisome because without premise (iii), which is implicit in the above quotation, it seems implausible that Krueger move to (iv), namely, the view that the emotional mind is extended. Without premise (iii), the link between coupled systems and extended processes is severed. To illustrate the problem, consider how Adam and Aizawa’s (2001, 2008) “coupling-constitution fallacy” might apply—this presentation of argument is borrowed from Wilson (2010, p.286). In this instance, y = emotional process and x = musical affordance.

- I. Y is a cognitive process
- II. X is causally coupled to Y
- III. X and Y form an integrated system (with functional gain)
- IV. X is part of a cognitive process

Construed in this way, Krueger’s argument plausibly exhibits a version of the coupling-constitution fallacy. It moves from a claim about a coupled system to a claim about an extended process. To do so, however, problematically attributes constitution status to external processes that might only be causally related.

Though a defense of premise (iii) is needed, one isn’t forthcoming. The closest Krueger comes to offering a defense of (iii), in fact, almost seems question begging. He writes: “Within certain circumstances, artifacts, tools, technologies, cultural institutions can become part of a spatially extended cognitive system in virtue of the active role they play in driving various cognitive processes” (Krueger, 2014, p.6). However, it cannot be assumed that because external resources allow for scaffolding that they get to count as cognitive. An account of musically extended cognition must defend such a claim. Krueger’s argument is lacking with respect to premise (iii). Briefly then, consider some buttressing comments for premise (iii).

What Krueger seems to have in mind with premise (iii) is Wilson’s (2010) notion of a “functionally integrated gainful system”. For example, Krueger often writes about “reciprocal causation” and “integration” (2014, p.7). For present purposes, only the basics of the idea need to be introduced. To identify something as a functionally integrated gainful system, three conditions need to hold. First, there needs to be reliable causal connection between two (or more) components. If so, there is a coupled process. Second, when two (or more) coupled processes reciprocally affect each other, there is an integrative system. Third, when an integrated system produces functional gain—that is, when the coupled processes enhance the system’s overall effectiveness in some interesting way—it becomes a functionally gainful integrated system. The promise of functionally gainful integrated systems is that they can serve as the basis of extended systems. The coupled components that form the basis of the system’s processes can include environmental elements. Defending Krueger’s premise (iii) requires substantiating functionally gainful integrated systems in the natural world.

To illustrate, consider two examples—both examples are taken from Wilson (2010, p.285). First, the human digestive system involves causal couplings between human body parts and microorganisms. For example, *E.coli* and the human stomach often perform digestive processes in concert. This is achieved through reciprocal interaction—*E.coli* help to break down simpler substances that are then absorbed by the intestines. The *E.coli* and other components of the stomach form an integrative system. Because the human digestive system does not work as well without both components acting in concert, there is functional gain produced by this integrated system. The result is a functionally gainful integrative system of

microorganismic and bodily digestion. Next, consider the giant water bug *Lethocerus*. *Lethocerus* injects its prey with enzymes, which liquefy the contents of the prey's innards. *Lethocerus* then sucks up and ingests the liquefied contents. This digestive process is constituted by an integratively coupled system of external and internal components. In coupling its digestive process to external resources, *Lethocerus* gains added functionality. *Lethocerus*' extended digestion seems to exhibit a functionally gainful integrative system. Though brief, these two examples suffice to show that functionally gainful integrative systems are common occurrences within the natural world. What's more, given the scientific respectability of extended biology, it seems plausible for Krueger to suggest that cognition might also exhibit something like functionally gainful integrated systems. Thus, assuming Krueger appeals to this notion in premise (iii), the potential problem raised can be resolved.

Consider, then, a brief comparison between the two views. First, while Krueger's account makes use of the notion of affordances, the current view refers only to informational invariants; both notions are borrowed from Gibson. The distinction amounts to this: Invariants are properties of the environment that remain constant through change, and affordances are invariant combinations of invariants; higher order invariants specify affordances. One potential drawback of conceiving of music as offering musical affordances is that it places some burden on Krueger to explain how perceivers detect affordances. As it stands, Krueger's answer leaves something to be desired. To use what's "synchronized with" to define what affordances music offers is somewhat circular. Even if affordances are bi-directional entities, this doesn't really suffice to explain *why* music should offer the affordances it does. This is not offered so much as a criticism, but as a cautionary flag. One of the deficits of Gibson's theory is that it is notoriously difficult to even specify what the invariants are (Goldstein, 1981, p.193). Providing an *informative* answer to what affordances are is an even bigger challenge, and it remains to be seen if Krueger's view can do this. One advantage the current view might have, then, is that it does offer an informative answer about both what musical invariants are and how they figure in music perception.

Second, Krueger's approach suggests sympathy with more "dynamic" approaches to cognition. One trend in the 3E (embodied, embedded, and extended) cognition literature has been to put pressure on static views of cognition. As Clark (1997, 2003, 2008) and others have pointed out, cognition is a real-time, pressured activity. Approaches to studying cognition need to reflect this by incorporating dimensions of feedback and emergence as much as possible. This push for dynamicism has led many, including Krueger, to look at how external resources play an integral part in cognitive feedback loops. Thus, we find Krueger saying things like "[m]usical dynamics thus provide external scaffolding supporting the synchronic emergence" (2014, p.7). This view contrasts in some ways to the one on offer. In viewing cognition as computation, emphasis is put on cognition as information processing. Such views do not necessarily exclude dynamic features, but they do not put them front and center. At some point, views of extended music cognition may have to come to terms with each other and adopt a common framework. However, given the relative youth of combining music studies with 3E cognition, it seems safe to say that the study of music cognition can pursue both approaches. Like so much of cognitive science the proof may be in the pudding. Whether either framework is more profitable depends on the research that comes out of it. For now, there seems ample room for both approaches.

NOTES

[1] Address correspondence to Luke Kersten, Institute of Cognitive Science, Carleton University, Ottawa, Ontario, Canada, K1S 5B6; e-mail: lukekersten@cmail.carleton.ca

REFERENCES

- Adams, F., & Aizawa, K. (2001). The bounds of cognition. *Philosophical Psychology*, 14(1), 43-64.
- Adams, F., & Aizawa, K. (2008). *The bounds of cognition*. Oxford: Blackwell Press.
- Balzano, G. (1986). Music perception as detection of pitch-time constraints. In G. Balzano & V. McCabe (Eds.), *Event cognition: An ecological perspective*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

- Chemero, A. (2011). *Radical embodied cognitive science*. Cambridge, MA: MIT Press.
- Clark, A. (1997). *Being there: putting brain, body, and world together again*. Cambridge, MA: MIT Press.
- Clark, A. (2003). *Natural-born cyborgs*. New York: Oxford University Press.
- Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. New York: Oxford University Press.
- Clark, E. (2005). *Ways of listening: An ecological approach to the perception of musical meaning*. Oxford; New York: Oxford University Press.
- Dawson, M. R. W. (2013). *Mind, body, world: Foundations of cognitive science*. Edmonton, AB: Athabasca University Press.
- Deutsch, D. (2013). *The psychology of music* (3rd Ed.). San Diego: Elsevier.
- Edelman, S. (2008). On the nature of minds, or: Truth and consequences. *Journal of Experimental and Theoretical Artificial Intelligence*, 20(3), 181-196.
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.
- Gibson, J. J. (1986). *The Ecological Approach to Visual Perception*. East Sussex; New York: Psychology Press.
- Goldstein, B. (1981). The ecology of J. J. Gibson's perception. *Leonardo*, 14(3), 191-195.
- Jones, M., & Hahn, J. (1986). Invariants in sound. In G. Balzano & V. McCabe (Eds.), *Event Cognition: An Ecological Perspective* (pp.255-256). Hillside, New Jersey: Lawrence Erlbaum Associates, Publishers.
- Kirsh, D. (2009). Problem solving and situated cognition. In P. Robbins & M. Aydede (Eds.), *The Cambridge handbook of situated cognition* (pp. 264-306). New York: Cambridge University Press.
- Krueger, J. (2014). Affordances and the musically extended mind. *Frontiers of Psychology*, 4(1003), 1-9. Retrieved from: <http://www.frontiersin.org/Journal/10.3389/fpsyg.2013.01003>
- Krumhansl, C. L. (1990). *Cognitive Foundations of Musical Pitch*. New York: Oxford University Press.
- Majita, J., & Schavio, A. (2013). Enactive music cognition. *Constructivist Foundation*, 8(3), 351-357. Retrieved from: <http://www.univie.ac.at/constructivism/journal/8/3/351.matyja>
- Neisser, U. (1999). Ecological psychology. In R. Wilson & F. Keil (Eds.), *The MIT encyclopedia of the cognitive sciences*. Cambridge, MA: MIT Press.
- Neuhoff, J. G. (2004). *Ecological Psychoacoustics*. New York: Academic Press.
- Newell, A., Shaw, J. C., & Simon, H. (1960). Report on a general problem solving program. In S. de Picciotto (Ed.), *Information processing: Proceedings of the International Conference on Information Processing* (pp. 256-264). Paris: UNESCO.
- Piccinini, G. & Scarantino, A. (2011). Information processing, computation, and cognition. *Journal of Biological Physics*, 37(1), 1-38.
- Pylyshyn, Z. (1984). *Computation and cognition*. Cambridge, MA: MIT Press.
- Segal, G. (1997). Review of R. A. Wilson, 'Cartesian psychology and physical minds: Individualism and

- the sciences of mind?. *British Journal for the Philosophy of Science*, 48(1), 151-156.
- Sekuler, R. & Blake, R. (1990). *Perception* (2nd Ed.). New York: McGraw- Hill.
- Shapiro, L. (2011). *Embodied cognition*. New York: Routledge.
- Sneddon, A. (2011). *Like-minded: Externalism and moral psychology*. Cambridge, Mass.: MIT Press.
- Stufflebeam, R. (1999). Computation and representation. In W. Bechtel & G. Graham (Eds.), *A companion to cognitive science*. Malden, Massachusetts: Blackwell Publishers.
- Trehub, S. E., Bull, D., & Leigh T. A. (1984). Infants' perception of melodies: The role of melodic contour. *Child Development*, 55(3), 821-830.
- Wilson, R. (1994). Wide computationalism. *Mind*, 103(4), 351-372.
- Wilson, R. (1995). *Cartesian psychology and physical minds: Individualism and the sciences of the minds*. Cambridge: Cambridge University Press.
- Wilson, R. (2004). *Boundaries of the mind: The individual in the fragile sciences*. Cambridge: Cambridge University Press.
- Wilson, R. (2010). Extended vision. In N. Gangopadhyay, M. Madary, & F. Spicer (Eds.), *Perception, action and consciousness* (pp.277-290). New York: Oxford University Press.
- Wilson, R. (2014). Ten questions concerning extended cognition. *Philosophy Psychology*, 27(1), 19-33.
- Wilson, R., & Lenart, B. (2014). Extended mind and identity. In J. Clausen & N. Levy (Eds.), *Handbook of Neuroethics*. London: Springer.