

Oxford Handbooks Online

Improvisation, Action Understanding, and Music Cognition with and without Bodies

Vijay Iyer

The Oxford Handbook of Critical Improvisation Studies, Volume 1 (*Forthcoming*)

Edited by George E. Lewis and Benjamin Piekut

Online Publication Date: Jun
2014

Subject: Music, Music Cognition

DOI: 10.1093/oxfordhb/9780195370935.013.014

Abstract and Keywords

A posited definition of improvisation encompasses such a broad range of human actions that it is helpful to consider both improvisation and rhythm in terms of embodied cognition and a notion of bodily empathy. This suggests a possible (though unstable and inconclusive) connection to action understanding, empathy, and mirror neurons, while acknowledging the latter's disputed status. With or without mirror neurons, the concept of action understanding offers a reconsideration of improvisation and music cognition with or without bodies (i.e., live or recorded). The relationship of improvisation, rhythm, and embodiment to contemporary theories of expectation, speech, and the evolution of music are considered. Action understanding is posited as the foundation of both music cognition and the perception of improvisation, marking both processes as inherently intersubjective, even whether the other's body is absent or fantasized (as is the case with recorded music).

Keywords: improvisation, music, neuroscience, mirror neurons, action understanding, empathy, race, perception, embodied cognition, experience, rhythm, the body

What are we referring to when we use the word "improvisation"? The term is used in innumerable ways, but always with the implicit assumption that there are acts that are improvised and acts that are not, and that those two kinds of acts are distinguishable. Two main aspects of that class of acts we call "improvised" seem to be (1) a real-time process of making choices and acting on them, and (2) the sense of temporal embeddedness: the fact that these actions take time, and that the time taken matters. With this understanding, we might take improvisation to denote that semi-transparent, multi-stage process through which we sense, perceive, think, decide, and act in real time.

But when construed this broadly, improvisation seems to encompass most of our behavior, including acts as disparate as walking through a forest or an airport, hunting and gathering, conversational speech, sport, climbing, driving, courtship, parenting, social dancing, and surfing the web. The class of improvisational behaviors is so vast that it may be easier to list behaviors that are *not* improvised—the carrying out of routines, plans, checklists, pre-routed or pre-ordained actions, well-rehearsed songs and dances, rituals, recitations, pageants, ceremonies, scripted performances of fully composed works—these last few exemplifying what Edward Said called "extreme occasions" (Said 1991). It seems that this class of non-improvised behaviors are the overall exception, a relatively small (but important) subset of human behavior as a whole.

Improvisation would also seem to encompass the noisy processes by which we acquire most skills. Babies learn to talk by babbling; they learn to walk by staggering, finding their balance, stumbling, finding something to hold onto; they learn to eat efficiently by first making a lot of messes. Improvisation is also the means by which we solve problems, by resorting to a repertoire of skills and adapting them to the situation at hand: putting out a fire, fixing a leak, doubling back to catch a missed turn, or building a shelter.

As we expand and refine these lists, we realize that most behaviors include improvised and non-improvised

components. For improvisation also seems to govern the ways in which we do things; even if we go to a store to buy the items on a grocery list, we might find ourselves making moment-to-moment decisions in how we navigate the store, how we choose specific tomatoes, how and with what (and indeed whether) we decide to pay. We make choices based on what's at hand, what's allowed, and what's desired, and also based on what we are taught, trained, forced, or empowered to do, or on what we are experienced in doing.

And similarly, acts of improvisation can readily incorporate patterns of behavior. Seemingly spontaneous speech acts can easily lapse into routine exchanges, or the re-narrating of previously told stories and jokes; a politician may answer a question from a reporter with a previously crafted statement; an improvising musician may develop a "personal sound" that might include hallmark melodic ideas, specific techniques, a habitual way of producing a certain sound. These facts do not deny their improvised or real-time quality; rather, they reveal how decades of choosing can lead to patterned responses to similar conditions.

In light of these observations, it becomes more and more problematic to identify moments of "pure" improvisation, or to disambiguate them from the execution of pre-ordained programs. We might think that we can recognize improvised acts in extreme moments—uncontrolled facial expressions, slips of the tongue, non-grammatical formulations, or graceful witticisms on an unscripted television show, for example—but it is still difficult to prove their unscriptedness; we merely trust or believe that they are so.

What we seem to be doing, instead of literally *identifying* improvisation according to some intrinsic attribute, is allowing cultural and contextual factors to *regulate* the presence or absence of improvisation. To attend a play or a narrative film or a symphony, for example, is to witness what one knows to be a series of carefully scripted and sculpted human actions, while to watch emcees in a street corner "cipher," to hear a performance of Hindustani classical music, or to attend an "improv comedy" event or a jazz club is to knowingly witness individual and collective acts of improvisation, and to parse them in those terms. We may also use visual and auditory cues to signal an event's composedness, its un-improvised character: a moment of ensemble synchrony in music or dance, for example, might seem statistically unlikely to be improvised, so we take it to be somehow planned. But in this realm, we can also be tricked: the improvisational dance technique known as "flocking" can create the illusion of choreographed group movement; similarly, systems of cues are frequently embedded in improvised musical settings, whereby moments of synchrony can emerge from apparent disorder. We therefore arrive at a crucial question for music cognition: when construed this broadly, does improvisation "sound like" anything? We will return to this question.

Embodied Music Cognition

A central theme in my work is the trace of the body in music. My dissertation and subsequent publications (Iyer 1998, 2002, 2004a, 2004b, 2009) explore the role of physical embodiment and sociocultural situatedness in music cognition. Drawing from recent advances in cognitive science, I claimed that music cognition should be understood as intimately tied in with the body and its physical and sociocultural environment—a perspective that was previously neglected in the music cognition literature.

The paradigm of embodied cognition emerged in the late 1980s as a corrective response to the Cartesian "dualist" theories of mind that had prevailed in cognitive science since the field's inception in the mid-twentieth century. Dualism held that the mind exists in a realm separate from the brain—that is, that the mind could be understood as "the software" and the brain and body as "the hardware." The dualist paradigm known as "cognitivism" thereby presupposed that cognition was a kind of rule-based computation that could happen in any machine using the same rules, and that there was therefore nothing special about the bodies that housed our brains. The cognition-as-computation view then influenced the field of music cognition, where music was treated primarily as a disembodied flow of forms in an abstract space. Early research in the field focused on perception of timbre, harmony, and pitch, largely neglecting subjects such as rhythm, performance, or the physical act of making music, not to mention the cultural forces that might lead one to prioritize harmony at all.

Theories of embodiment hold that the body, the brain, and the mind must be understood as one system, and that the brain is an organ optimized for producing motor (i.e., bodily) output in response to sensory stimuli. This "sensory-motor loop" becomes the basis for what we call cognition. Rather than seeing thought as a process separate from sensation or action, we understand the faculties of perception, thought, and action as codependent,

having developed together both ontologically (from birth through childhood and into adulthood) and phylogenetically (in the evolution of the species).

It should be noted that adopting this framework does not inherently refute the idea of abstraction or “concepts.” Intermediate theories such as the “grounding by interaction” framework (Mahon and Camerazza 2008) allow for abstract or symbolic concepts to be instantiated in the context of specific sensory and motor information. Such a framework offers a view of cognition that is neither fully embodied nor fully disembodied, but contains aspects of both.

From the embodiment paradigm, or some intermediate form of it, we can develop a body-based view of musical cognition. This view is borne out, for example, by brain imaging studies that have highlighted a fundamental identification between rhythm and human movement. It is understood that to perceive rhythm is to “imagine movement”; musical rhythm provokes the brain to prepare the body to move, facilitating or activating a physical response (Todd 1999, Todd et al. 1999). In consideration of some recent claims about mirror neurons (Kohler et al. 2002), this might be considered as a kind of “aural mirroring.” In this view (discussed critically below), we tend to respond *in kind* to what we think we hear another body doing, imagining or actually generating an action that is suggested by the rhythmic character of the sounds we hear.

Music is then understood as the sound of human bodies in motion; to listen to music is to perceive the actions of those bodies, and a kind of sympathetic, synchronous bodily action (i.e., dance) is one primary response. Of course, this is mediated by culture. Certain cultural settings may foreground bodily responses to music, while others conceal or suppress them; these variations express “situated cognition” (Robbins and Aydede 2008, Clancey 1997)—the interrelationship of mind and world, the interdependence between knowledge and its context.

The notion of music cognition as an embodied, situated phenomenon ties in well with contemporary ethnomusicological accounts of African diasporic musics, in which embodiment, performativity, and cultural context play a crucial role in the production of meaning. The idea of embodiment can also bring the field of music perception and cognition into a healthier dialogue with the music humanities, which has in recent decades seen robust critical engagement with “the body” in terms of race, gender, and sexuality. When we hear bodies but do not see them, we instead fantasize about them; listening to music (especially in the disembodied way that it circulates today) is deeply informed by that same process of racialized, sexualized fantasy-formation about the virtual bodies that made those sounds. While interdisciplinarity can, at its worst, provoke an unproductive confrontation of epistemological incongruities between paradigms, bodies can still offer a strong focus for dialogue, critique, and new productions of knowledge across many fields of inquiry. Although bodies are described from many different standpoints, somehow we can all agree on their sheer, stubborn presence (or absence, as the case may be).

Embodied Rhythm Perception

The central idea that music is an embodied, situated activity means that music depends crucially on the structure of our bodies, and also on the environment and culture in which our musical awareness emerges. Rhythm, especially, is a complex, whole-body experience, and its role in music makes use of the embodied, situated status of the participant. Such claims have a variety of implications; they lead us to appreciate traces of the embodiment in instrumental and vocal music, to notice how musical cultures and individuals variously deal with the role of physicality in music-making, and to understand music perception as an active, culturally contingent process.

The claim that music perception and cognition are embodied activities also means that they are actively constructed by the listener, rather than passively transferred from performer to listener. This active nature of music perception highlights the role of culture and context. For example, the discernment of qualities such as pulse and meter from a piece of music is not perceptually inevitable; rather, the music may offer perceptual ambiguities whose resolution depends on an observer’s culturally contingent listening strategies (Iyer 1998: 83–104). In addition, I have argued that rhythmic expression is often directly related to the role of the body in making music, and to certain cultural aesthetics that privilege this role. In particular, certain subtle microrhythmic variations in rhythmic performance display striking systematic structure, carrying an encoded trace of the culturally situated music-making body. (For a detailed explanation, see Iyer 2002.)

The salience of fine-grained musical rhythm at this level of detail is borne out by more recent neuroscientific studies of timing perception (see Patel 2008, 96–154; Levitin 2009). In particular, microrhythmic perception appears to take advantage of our ability to differentiate between phonemes (Patel 2008), and, crucially, our ability to aurally locate and track human movement around us (Changizi 2011). The assertion of the existence in the microrhythmic realm of meaningful musical structure, activating our faculties for perception of human movement and speech, runs counter to more common descriptions of microrhythmic variation as “discrepancies” (Keil 1987), “imperfections” or “being slightly off” (Hennig et al. 2012), or, for *The New York Times*, “essentially mistakes” and “error[s]” (Belluck 2011).

Rhythm—in music, speech, bodily acts, or some interleaved combination thereof—offers one productive line of inquiry for theories of embodied cognition. Another such category is improvisation. Rhythm and improvisation are not mutually exclusive categories in this regard, but two overlapping fields of behavior, both treating *time* as a central parameter.

Time and Temporal Situatedness

A fundamental consequence of physical embodiment and environmental situatedness is the fact that *things take time*. Temporality must ground our conception of physically embodied cognition. Smithers (1996) draws a useful distinction between processes that occur “in-time” and those that exist “over-time.” The distinction is similar to that between process-oriented activity, such as speech or walking, and product-oriented activity, such as writing a novel or composing a symphony.

In-time processes are *embedded* in time; not only does the time taken matter, but in fact it contributes to the overall structure. The speed of a typical walking gait relates to physical attributes like leg mass and size, and shoulder-hip torsional moment; this is why we cannot walk one-tenth or ten times as fast as we do. Similarly, the rate at which we speak exploits the natural timescales of lingual and mandibular motion as well as respiration. Changizi (2011) further argues that human speech is made of vocalized imitations of real-world solid-object events—“hits, slides, and rings”—with speech’s rhythmic profiles derived from this physical “grammar.” Accordingly, we learn to process speech at precisely such a rate. Recorded speech played at slower or faster speeds rapidly becomes unintelligible, even if the pitch is held constant. The perceived flow of conversation, while quite flexible, is sensitive to the slowdown caused by an extra few seconds taken to think of a word or recall a name.

Over-time processes, by contrast, are merely *contained* in time; the fact that they take time is of no fundamental consequence to the result. Most of what we call computation occurs over time. The fact that all computing machines were originally considered computationally equivalent regardless of speed suggests that time was not a concern in the original theory of computation, and that the temporality of a computational process was theoretically immaterial. Though computational theory is more nuanced today, “real-time” computer applications make use of the speed of modern microprocessors, performing computations so fast that the user doesn’t notice how much time is taken. However, this is not what the mind does when immersed in a dynamic, real-time environment; rather, it exploits both the constraints and the allowances of the natural timescales of the body and the brain as a total physical system. In other words, Smithers (1996) claims, *cognition chiefly involves in-time processes*. Furthermore, this claim is not limited simply to cognitive processes that require interpersonal interaction; it pertains to all thought, perception, and action.

The Temporality of Performance

In intersubjective activities, such as speech or music making, one remains aware of a sense of mutual embodiment. This sense brings about the presupposition of “shared time” between the listener and the performer. This sense is a crucial aspect of the temporality of performance. The experience of listening to music is qualitatively different from that of reading a book. The experience of music requires the listener’s “co-performance” within a shared temporal domain (Schutz 1964). While the essentially solitary act of reading a book also takes time, the specific amount of time is of little consequence. (Literary notions of co-performance, such as Roland Barthes’ idea of “writerly texts” [1975], do not fundamentally incorporate the temporality of experience.) The notion of musical co-performance is made literal in musical contexts primarily meant for dance; the participatory act of marking time with rhythmic bodily activity physicalizes the sense of shared time, and could be viewed as embodied listening.

The performance situation itself might be understood as a context-framing device. In his study of the music of a certain community in South Africa, ethnomusicologist John Blacking wrote, “Venda music is distinguished from nonmusic by the creation of a special world of time. The chief function of music is to involve people in shared experiences within the framework of their cultural experience” (Blacking 1973: 48). There is no doubt that this is true to some degree in all musical performance, and we can take this concept further in the case of improvised music.

What Does Improvisation Sound Like?

Can one *hear* the fact that a sound was just decided in its moment of creation? Does music have any characteristic that announces itself as such? My sense, from my own performing, teaching, and listening experiences, is that most listeners can’t tell the difference in isolation; the perception of the relative presence or absence of improvisation is largely imagined and profoundly contextual, based on cultural factors and assumptions. And analytically, it is difficult to find any such traits on the musical surface; the tendency in the West is to focus on so-called “mistakes” as an indicator of improvisation, as if they somehow verify the fragile and risky process.

Limb and Braun (2008), in their fMRI brain imaging studies of skilled musicians in the act of improvising, found that in a focused state of improvisation, a soloist has lowered self-correcting inhibitions and enhanced activation of a self-narrative function. “Just relax and be yourself,” might be another way to put it; it’s the kind of advice given before a date or a job interview, both of which are also instances of heightened, carefully framed improvisation. But Limb (private communication, 2012) further suggests that there is something else happening in moments of highly skilled musical improvisation, something like what Csikszentmihalyi (1990) calls a “flow state,” a mental and physical state of utmost relaxation, focus, and concentration, known to be conducive to creative thinking. Such a criterion (if it is indeed a criterion, and not merely a correlated phenomenon or perhaps an outcome of the “overlearning” involved in improvisational skill) may enable us to qualitatively set apart sustained creative musical improvisation, for example, from our more everyday spontaneous acts.

However, that doesn’t tell us anything about what improvisation “sounds like” to an observer. So another question is, does the distinction matter? Why are we, in music, so bound up with the issue? The reason is that the experience of listening to music that is understood to be improvised differs significantly from listening knowingly to composed music. The main source of drama in improvised music is the sheer fact of the shared sense of time: the sense that the improviser is working, creating, generating musical material in the same time in which we are co-performing as listeners (Iyer 1998, 80). As listeners to any music, we experience a kind of *empathy* for the performer, an awareness of physicality and an understanding of the effort required to create music. This empathy is one facet of our listening strategies in any context. In improvisational music, this embodied empathy extends to an awareness of the performers’ coincident physical and mental exertion, of their “in-the-moment” processes of creative activity and interactivity, the risks taken in the face of unbounded possibilities, the inherent constraints of the mind deciding and the body acting in time. Perhaps it can be said that improvisational music magnifies the role of embodiment in musical performance. The perception of improvisation seems to involve the perception of another body or bodies engaged in embodied, situated, real-time experience. If so, then a sense of mutual embodiment, of the shared space, time, and bodily presence of performers and observers, would seem to open the door to specific kinds of *empathy*.

A Science of Empathy

A flurry of recent, intensely debated findings in neuroscience have suggested a neural basis for empathy that seems to reinforce this view. The body of research on so-called “mirror neurons” (Gallese et al. 1996; Kohler et al. 2002; Rizzolatti and Sinigaglia 2010; Gallese et al. 2011) promotes the idea of *action understanding*, a kind of empathy at the neural level, and claims the existence of a “mirroring mechanism” for action understanding, in which the perception of certain familiar actions in another body can trigger the activation of similar motor programs in the observer’s brain: “[E]ach time an individual observes another individual performing an action, a set of neurons that encode that action is activated in the observer’s cortical motor system” (Rizzolatti and Sinigaglia 2010: 264). These activations may manifest as analogous bodily motion, action or stance to that of the observed body, or they may just remain at the level of an “imagined movement.” There is evidence in primates for the

existence of such a neural system—as famously depicted in photographs of a baby macaque sticking its tongue out in response to a scientist doing the same (Gross 2006)—suggesting something quite fundamental about this process, even applying across species in primates (Buccino et al. 2004).

Unfortunately, we cannot allow these findings to paint too rosy a picture of universal “understanding.” Recent studies on mirror neurons and racial identification (Gutsell and Inzlicht 2010) actually suggest that the perception of racialized difference may inhibit or constrain empathy. It was observed that test subjects (all whites from North America) displayed a greater mirror neuron-type response to images of other whites than they did to non-whites. In some cases, whites displayed practically zero empathy-like mental simulation of actions of non-whites. This finding has been extended to more fluid “in-group/out-group” affiliations, suggesting a profound neuroplasticity in this “mirroring” mechanism associated with empathy. For reviews of such findings, see Eres and Molenberghs (2013) and Matusall (2013).

That science might find empathy to be instinctively possible across real species boundaries, and yet also suppressed across imagined racial boundaries, suggests that there could be both innate and learned aspects to action understanding, and that it can be informed by both structural and superficial qualities visually perceived in the other. It would appear that there is no such thing as “clean” mirroring; there is perhaps always some distortion of the metaphorical mirror, since the problematic visual “perception” of racial difference can seemingly interfere with action understanding.

However, crucially for music cognition, some recent research suggests that mirror neurons might be involved both visual *and* auditory processing. Whereas early research in this mirror neuron system focused on action understanding activated through visual stimuli (Rizzolatti et al. 1996), subsequent work has revealed a similar mechanism at work through auditory channels (Kohler et al. 2002). It is claimed that “these audiovisual mirror neurons code actions independently of whether these actions are performed, heard, or seen” (Kohler et al. 2002). The notion that this process could occur through sound—that we may undergo a kind of empathetic action understanding when we merely *hear* someone do something, without *seeing* that person do it—offers quite radical implications for how we listen to music (and especially what happens when we hear music without seeing).

This offers a tantalizing reading of the history of recorded music, as bound up as it is with race and twentieth-century American history. Is it possible that music-heard-and-not-seen (which, of course, was a rarity before the advent of recorded music, Pythagorus’s “acousmatic” scenario notwithstanding) might have overridden the visual, racialized, culturally imposed constraints on empathy? Could the essential humanity of African Americans been newly revealed for white American listeners in the twentieth century through the disembodied circulation of “race records,” by activating in these listeners a neural “understanding” of the actions of African American performers? These were, after all, among the first recordings to circulate on a mass scale in the United States. Could a new kind of cross-racial empathy, or at least a new quasi-utopic racial imaginary, have been inaugurated through the introduction and sudden ubiquity of recorded sound? As the above line of inquiry suggests, this very idea—that disembodied human sound can elicit in the listener a mirroring or empathic understanding of the imagined movements of an imagined other—carries the disruptive potential to restructure our knowledge of what music is, why it exists, and how it works.

Mirror neurons and their identification with action understanding have received intense scientific scrutiny and critique, especially in the last few years (Hickok 2009, Hickok and Hauser 2010, Hutto 2013). Contributing to the issue has been the fact that “the concept of ‘action understanding’ has been evolving” (Hickok 2009) due to the persistent, irresolvable question of what it means to “understand” the action of another. To confirm “understanding,” must one reproduce the other’s action identically, or does it suffice simply to “know how” to do the action, or perhaps to understand its intended goal? Can there be generative mis-“understandings” of the actions of others, and how are they distinguished from “true” action understanding?

Still, underlying most instances of the phrase “action understanding” is the idea that, in certain cases, “self-generated actions have an inherent semantics and that observing the same action in others affords access to this action semantics” (Hickok 2009). A recent review by Rizzolatti and Sinigaglia (2010) concedes that there may be types of action understanding might be attributable to non-mirror mechanisms in the brain, and then turns its focus to a specific type of mirror mechanism “that allows an individual to understand the action of others ‘from the inside’” (264), which seems to specifically mean the knowledge of *how* to perform an action that one is observing.

A tidy resolution of the rapidly evolving mirror-neuron debate is beyond the scope of this article (or indeed of any article, as of this writing). What is of direct relevance to our discussion here is *the concept of action understanding*—not its exact neural mechanism, but its very existence, and its explanatory power as an intersubjective framework for music cognition. The last century's global cultural transformation—from humankind's longstanding identification of music with embodied action, to the sudden propagation of recorded music and its concomitant abundance of music without bodies—offers us a productive conceptual space to consider the role of action understanding (or its commodified 20th century replacement—which can only be called *fantasy*) in the act of listening to music.

Expectations

If improvisation and rhythm are central to embodied musical cognition, are these claims borne out in the current literature in music cognition? Curiously, recent “big-picture” treatises in music cognition have avoided discussions of improvisation, despite its seeming primacy in musical and cognitive experience. David Huron's work on expectation (2008) considers the evolutionary advantages of humankind's ability to predict events based on cues: “Those who can predict the future are better prepared to take advantage of opportunities and sidestep dangers ... Accurate expectations are adaptive mental functions that allow organisms to prepare for appropriate action and perception” (Huron, 3). These functions, he continues, are entangled with emotional response: “[T]he emotions accompanying expectations are intended to reinforce accurate prediction, promote appropriate event-readiness, and increase the likelihood of future positive outcomes ... [M]usic-making taps into these primordial functions to produce a wealth of compelling emotional experiences ... including surprise, awe, “chills,” comfort, and even laughter” (Huron, 4). Building on Meyer's (1956) proposed correspondence between expectation, emotion, and meaning in the perception of musical form, Huron develops a perspective on music perception grounded in the science of expectation.

It becomes apparent that Huron's theory of expectation, while building on Meyer's composer-centered theory, is fundamentally similar to embodied and situated cognition. Expectation is a capacity that guides our understanding of real-world, real-time events in a way that helps us make efficacious, life-sustaining actions, to “predict the future” and “take advantage of opportunities.” This view would seem completely compatible with, and indeed nearly identical with, our working understanding of improvisation. It is therefore ironic and unfortunate that Huron's sole discussions of improvisation focus on how improvisers cover for “wrong” notes (234–235, 291). Indeed, the improvisational orientation of Huron's entire theory—grounded in an understanding of perception as optimized for interacting in real time with an information-rich environment—is somehow repressed in his discussions of music. The reason seems to be that improvisation is not compatible with his working model of music, which is characterized by a division between music and its listeners; his and Meyer's basic thrust is that music makers (which for Huron and Meyer essentially means *composers*) make choices to manipulate the expectations of a passive audience of listeners. Music, in his view and in the views of many other researchers in the field, is something that *happens to* listeners, or something that they perceive without very much direct engagement; music is rarely framed as an *activity* that listeners *coexist with* as well as *participate in* throughout their entire lives, and are always already acculturated to.

The presupposition of a division between music and listener, between performer and audience, stems from a fundamentally non-participatory understanding of music, which runs counter to most anthropological evidence about how music tends to function in culture. That kind of separation is of course a widespread paradigm in the West, and in the court musics of many non-western cultures, but that does not make it meaningful in evolutionary terms. We stand to learn more about music's origins by attention to humankind's vernacular and folk musics, which are participatory almost by definition. Just as we humans have not evolved very much in the millennia since writing was introduced, we certainly haven't evolved significantly in the century since recordings became popular, or even in the last few centuries since composers started writing for orchestras.

The point here is that the embodied improvising agent, situated in a real-world physical and cultural environment, is most often the listener *and* the doer in the equation. Expectation is perhaps best understood as a capacity of the improviser—that is, all of us—to take in information, make predictions, and carry out informed, situated actions based on those predictions, with real-world consequences. Just as the theory of musical expectation is a consequence of a more general theory of expectation, this view of expectation as an improvisational skill has both

“real-world” and musical implications. For example, it has been observed in simulations (Friston, Mattout, and Kilner 2011) that “mirror neurons will emerge naturally in any agent that acts on its environment to avoid surprising events”—a startling conclusion that brings together notions of situated cognition and improvisation (the agent acting on its environment), expectation (learning to reduce predictive error, i.e., avoiding surprises) and action understanding—with repercussions for cognition in intersubjective situations, musical or otherwise.

Music and Speech

A recent treatise by Patel (2008) considers fundamental connections between music and language from a neuroscience perspective. Drawing from a huge range of research in music and speech perception, Patel presents a thorough view of the state of our current understanding of the connections between these two systems.

Given the exhaustive nature of this work, the absence of any substantial consideration of improvisation is again striking. But certain conceptual biases about what music *is* soon reveal themselves, which help explain this strange gap. In a discussion of linguistic meaning in relation to musical meaning, Patel imagines composers trying to write short pieces that “mean” common nouns or verbs (“school,” “eye,” “know”). Would listeners be able to “hear” their meanings? Most probably they would not, he answers. However, “lacking specificity of semantic reference is not the same as being utterly devoid of referential power ... Instrumental music lacks specific semantic *content*, but it can at times suggest semantic *concepts*. Furthermore, it can do this with some consistency in terms of the concepts activated in the minds of listeners within a culture” (Patel 2008: 328).

This uncovers a certain assumption about music and speech—that it is all received, never generated; that composers make music, and others learn to hear it. The assumption is that language is both simply created by others for us to learn to use, that it has inherent meanings that we can “hear,” and so forth—as if language were not itself a vast, improvisational, arbitrary, and continually evolving system of signs. The inherent bias is that music is not something that we do, but instead something that we merely accept from those who have the authority to do it for us. This removes music from the realm of action into the passive realm of “reception.”

Meanwhile, is the meaning of a speech act simply a question of processing—of decoding sounds and hearing their meaning? There is also something very important going on in real time, in the realm of expectation. A speech act comes into being in the void, in a sense; it not only conveys a meaning, but it also fills up an experiential space where there might just as easily not have been such an act. And once it is done, it cannot be undone. So the very fact of it having been *decided* in those moments under those constraints—decided often not even as a complete thought but *word by word*—marks it undeniably as improvisation. In other words, to speak is necessarily to improvise. At some level, to listen to speech is always to bear this fact in mind; the improvisational nature of speech is essentially axiomatic, seemingly a precondition for our ever communicating at all. “Speech acts” are performative in the sense that they represent a filling in of shared time with an improvisation that aims to construct meaning. Certainly in retrospect speech acts also “are” their semantic meanings, but before they acquire meaning, they are, first and foremost, *acts*.

Perhaps one reason that listeners cling to “mistakes” in music as evidence of improvisation is that “gaffes” do exactly the same thing for speech. Such “mistakes” (and of course it is difficult to name anything as such) underscore the fragility of the improvisational act. So we must bear in mind also the similarity between acts of musical improvisation and speech acts on this level: they always move *forward in time*, and they always are in some way a replacement for their own absence—their existence is always the result of a set of choices: that of whether to say anything at all, of what to say, and of when to say it. The gravity of an improvisational act is the very fact that it happened at all, as opposed to anything else that could have happened, including nothing.

Those choices exist within a dynamic web of interacting constraints—particularly the more social considerations of what is appropriate, what is expected, what is individually desired, and what is “right.” One of Charles Limb’s (2008) discoveries is that when we improvise, questions of what is “right” diminish in importance. In this way a set of constraints is relaxed, perhaps making it easier to choose, or allowing more choice.

Harnessing

The latest body of research to support the embodied and situated view of music is summarized in Changizi (2011). Here it is argued that music takes advantage of the existing skills we have of recognizing and decoding audible traces of human action. Instead of emphasizing pitch, harmony, and the other hallmarks of music perception research, the author focuses on our perceptual attunement to the specific sounds of human motion.

Since much of the time in everyday life we hear our fellow human beings in our peripheries without seeing them, Changizi suggests that we have evolved to decode and respond to these stimuli—to hear everyday human moving-around noises not as abstract sounds but as markers of bodies in motion. From an evolutionary perspective we are optimized to communicate with and ‘read’ our fellow humans. Building on this idea, the author argues that the details of music take advantage of our aural grounding in the perception of specifically *human* motion: the sound and rhythmic profile of footsteps as a marker of locomotive behavior; small Doppler shifts as an indicator of direction of motion (Oechslin et al. 2008); the correspondence between loudness and distance; and other such sonic hallmarks. Rather than suggest that humans evolved to hear music, he argues that humans “harnessed” an existing perceptual apparatus, which had evolved for the perception of human motion, to develop music, which, he suggests, mimics human action.

Changizi argues that music takes advantage of our aural ability to notice human action in the same way that written language takes advantage of our visual ability to notice contours, edges and joints, the building blocks of human vision. Furthermore, human movement is emotionally evocative; we can recognize the emotion from someone’s gait. Music “can often sound like *contagious* expressive human behavior and movement, and trigger a similar expressive movement in us” (Changizi 2011: 116). But we can take this argument a step further, since from the perspective of embodiment, music is more than a mere sonic imitation of human action; indeed it was never anything *but* human action. Until the last century, music was only ever made by *bodily* engagement with available sound-producing technology; whether it was mediated by objects adapted from the natural world (gourds and logs, animal skins and bones) or pure bodily acts (stomping, clapping, and singing), the sound of music *always* was the sound of people in motion—perhaps a stylized, synchronized, or sustained kind of motion, but never disconnected from bodily presence or action, nor ever outside of the realm of plausible human actions. This means that we don’t perceive rhythms in implausible frequency ranges of 1000 Hz or .001 Hz, because they do not correspond to any human action. We don’t readily integrate a stream of stimuli from physically separated sources as if it were a single source, but, rather, we notice harmonic tones because we are sharply attuned to the harmonicity of human voices.

In this perspective, crucially, music is inherently social; it taps into parts of our brains that connect us to other people. We can hear and immediately understand what people are doing in our midst; their gaits indicate their behaviors and emotions; the direction of their motion is indicated by changes in volume and pitch. Working with these very perceptual ingredients, music recreates for us the sensation and emotional thrill of people in our midst. Such reasoning falls in line with what J. J. Gibson called the “ecological mode of perception,” in which our perceptual systems are tuned to apprehend real-world *sound sources* in an environment, rather than only to pure sound itself (Gibson 1979; Shove and Repp 1995). It also aligns squarely with the views on embodied music cognition—the idea of bodies listening to bodies—as well as with the neural foundations of empathy and the notion of auditory action understanding.

Conclusion

To summarize the extravagant claims in this rather speculative article: Improvisation “matters” in music because a knowing listener experiences some kind of empathy for the embodiment of the performer, or some kind of understanding of the effortfulness of real-time performance. There is evidence that this phenomenon possibly has a neural basis, which is grounded in our ability to perceive, recognize, and decode the sights *and* sounds of bodies in motion. And this phenomenon is also linked to the foundations of rhythm perception, since the sound of a humanly generated rhythm (i.e., the sound of a body in motion) can activate an analogous body motion in a listener.

Our skill at perceiving, “understanding,” and/or imitating the sonorous actions of another enables us to synchronize our actions, to operate in rhythmic unison or in sustained antiphony, to move, sing, dance, or work together. Improvisation and rhythm, two foundational elements of music and creativity, both have at their core some kind of embodied perception and cognition of the other, and therefore they seem to be what enable us, as

human beings, to do anything together in the same time and space.

Music is born of our actions—its ingredients are the sound of bodies in motion—and therefore *music cognition begins as action understanding*. This does not mean that we cannot process musical information without bodies, but it does mean that our sensations and actions provide the context for abstraction, symbolic music cognition, and the fantasies brought about by music-without-bodies.

Coda

What do we mean when we use the word “experience”? It refers to both (1) the stream of sensation and perception and (2) the accumulation of cognition through sustained immersion in this stream. (Let us set aside any notion of “consciousness.”) Our very language suggests an assumed relationship between the first and second meanings: an “experienced” person is someone who has “experienced” enough to gain knowledge from these experiences. So the second sense of “experience” would seem to encompass the first; to be “experienced,” to have field knowledge, to know how to handle things in a given situation, is to have undergone prior “experiences” of a similar sort. In the embodiment perspective, we understand the first sense of “experience”—sensation and perception—as connected to and dependent upon embodied action; cognition becomes an umbrella term for all of these processes, as well as the mental structures that connect these different stages of the “sensorimotor loop”—sensation, perception, thought, action.

We might reconsider whether we have cast too wide a net with the expansive conception of improvisation posed in this chapter. But if we were to refuse to dignify an infant’s babbling with so exalted a term as improvisation—if we were to insist instead, for example, that a true act of improvisation first requires a coherent self, or some threshold level of “creativity,” or some move away from what is “habitual,” or that it must display some sort of “resistance” or “non-normativity” or “soul”—then we might limit the more radical implications of this overall perspective. For I am suggesting instead that improvisation, in this broad sense, might be considered as the means by which we *acquire selfhood*. It is not only a means of self-transformation, as Arnold Davidson (2005) has eloquently described it, but of self-generation. In other words, I am positing a relationship—or more to the point, an *identity*, a *sameness*—between what we call “improvisation” and what we call “experience.” A corollary is that an observer’s perception of improvisation is contingent upon an “understanding” of its status as experience—which, again, underscores the essential intersubjectivity of music cognition.

This broad view does not reject the possibility of political action or engagement. Indeed, political struggles for selfhood have been advanced by concurrent transformative improvisations in culture, whether it was the possible problematization of race brought on by the circulation of music-without-bodies, or the improvised musical expressions of an African American subculture defiantly asserting its collective humanity. It is in our interest to consider these perspectives on the origins of music, so that we may better understand how fundamental it is to the origins and elaborations of the self.

Improvisation is a human response to necessity.

—Muhal Richard Abrams (2007)

It seems to me what music *is* is *everything that you do*.

—Cecil Taylor (Mann 1981)

Bibliography

Abrams, M. R., M. Jefferson, Y. Komunyakaa, G. Lewis, and P. Williams. “Improvisation in Everyday Life: A Conversation.” Public lecture, Columbia University, New York, September 25, 2007. Available at <http://www.worldleaders.columbia.edu/events/conversations-series-improvisation-everyday-life>. Accessed December 31, 2013.

Barthes, R. *S/Z*, translated by Richard Miller. London: Cape Publishers, 1975.

- Belluck, Pam. "To Tug at the Heart, Music First Must Tickle the Neurons." *New York Times*, Vol. 160 (April 19, 2011): D1.
- Blacking, J. *How Musical Is Man?* Seattle: University of Washington Press, 1973.
- Buccino, G., F. Lui, N. Canessa, I. Patteri, G. Lagravinese, et al. "Neural Circuits Involved in the Recognition of Actions Performed by Nonconspecifics: An fMRI Study." *Journal of Cognitive Neuroscience* 16, no. 1 (Jan.-Feb. 2004): 114–126.
- Changizi, M. A.. *Harnessed: How Language and Music Mimicked Nature and Transformed Ape to Man*. Dallas, TX: BenBella Books, 2011. (Kindle edition)
- Clancey, W. *Situated Cognition: On Human Knowledge and Computer Representation*. New York: Cambridge University Press, 1997.
- Csikszentmihalyi, Mihaly. *Flow: The Psychology of Optimal Experience*. New York: Harper and Row, 1990.
- Davidson, Arnold I. "Introduction." In *Michel Foucault: The Hermeneutics of the Subject, Lectures at the College de France, 1981–1982*, translated by Graham Burchell, and edited by Arnold I. Davidson, xix–xxx. New York: Palgrave Macmillan, 2005.
- Eres, R., and P. Molenberghs. "The Influence of Group Membership on the Neural Correlates Involved in Empathy." *Frontiers in Human Neuroscience* 7 (May 2013): 176.
- Friston, K., J. Mattout, and J. Kilner. "Action understanding and active inference." *Biological Cybernetics*, Vol. 104, issue 1–2 (February 2011): 137–160.
- Gallese, V., L. Fadiga, L. Fogassi, and G. Rizzolatti. "Action Recognition in the Premotor Cortex." *Brain* 119, Issue 2 (1996): 593–609.
- Gallese, V., M. A. Gernsbacher, C. Heyes, G. Hickok, and M. Iacoboni. "Mirror Neuron Forum." *Perspectives on Psychological Science* 6, vol. 4 (2011): 369–407.
- Gibson, J. J. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin, 1979.
- Gross, L. "Evolution of Neonatal Imitation." *PLoS Biology* 4, no. 9 (2006): e311.
- Gutsell, J. N., and M. Inzlicht. "Empathy Constrained: Prejudice Predicts Reduced Mental Simulation of Actions during Observation of Outgroups." *Journal of Experimental Social Psychology* 46, issue 5 (2010): 841–845.
- Hennig, H., R. Fleischmann, and T. Geisel. "Musical Rhythms: The Science of Being Slightly Off." *Physics Today* 65, no. 7 (2012), Quick Study.
- Hickok, G. "Eight Problems for the Mirror Neuron Theory of Action Understanding in Monkeys and Humans." *Journal of Cognitive Neuroscience* 21, no. 7 (2009): 1229–1243.
- Hickok, G., and M. Hauser. "(Mis)understanding Mirror Neurons." *Current Biology* 20, no. 14 (2010): R593–R594.
- Huron, D. *Sweet Anticipation*. Cambridge, MA: MIT Press, 2008.
- Hutto, D. D. "Action Understanding: How Low Can You Go?" *Consciousness and Cognition* 22, Issue 3 (September 2013): 1142–1151.
- Iyer, V. "Microstructures of Feel, Macrostructures of Sound: Embodied Cognition in West African and African-American Musics." Ph.D. dissertation, University of California, Berkeley, 1998.
- Iyer, V. "Embodied Mind, Situated Cognition, and Expressive Microtiming in African-American Music." *Music Perception* 19, no. 3 (2002): 387–414.
- Iyer, V. "Exploding the Narrative in Jazz Improvisation." In *Uptown Conversation: The New Jazz Studies*, edited by R. O'Meally, B. Edwards, and F. Griffin, 393–403. New York: Columbia University Press, 2004.

- Iyer, V. "Improvisation, Temporality, and Embodied Experience." *Journal of Consciousness Studies* 11, no. 3–4 (2004): 159–173.
- Iyer, V. "Improvisation: Terms and Conditions." In *Arcana IV: Musicians on Music*, edited by John Zorn, 171–175. New York: Hips Road/Tzadik, 2009.
- Keil, Charles. "Participatory Discrepancies and the Power of Music." *Cultural Anthropology* 2, no. 3 (1987): 275–283.
- Kohler, E., C. Keysers, M. A. Umiltà, L. Fogassi, V. Gallese, and G. Rizzolatti. "Hearing Sounds, Understanding Actions: Action Representation in Mirror Neurons." *Science* 297 (2 August 2002): 846–848.
- Levitin, Daniel J. "The Neural Correlates of Temporal Structure in Music." *Music and Medicine* 1, no. 1 (2009): 9–13.
- Levitin, Daniel J., and Anna K. Tirovolas. "Current Advances in the Cognitive Neuroscience of Music." *The Year in Cognitive Neuroscience 2009: Annals of the New York Academy of Sciences* 1156, no. 1 (2009): 211–231.
- Limb, C. J., and A. R. Braun. "Neural Substrates of Spontaneous Musical Performance: An fMRI Study of Jazz Improvisation." *PLoS ONE* 3, no. 2 (2008): e1679. doi:10.1371/journal.pone.0001679.
- Mahon, B. Z., and Caramazza, A. "A Critical Look at the Embodiment Hypothesis and a New Proposal for Grounding Conceptual Content." *Journal of Physiology—Paris* 102, issue 1-3 (January-May 2008): 59–70.
- Mann, R. *Imagine the Sound* (documentary feature film). New York: Janus Films, 1981.
- Matusall, S. "Social Behavior in the 'Age of Empathy'? A Social Scientist's Perspective on Current Trends in the Behavioral Sciences." *Frontiers in Human Neuroscience* 7 (May 31, 2013): 236.
- Meyer, Leonard. *Emotion and Meaning in Music*. Chicago: University of Chicago Press, 1956.
- Oechslin, M., M. Neukom, and G. Bennett. "The Doppler Effect—An Evolutionary Critical Cue for the Perception of the Direction of Moving Sound Sources." In *Proceedings of International Conference on Audio, Language and Image Processing*, Shanghai, China, July 2008. ICALIP 2008.
- Patel, A. *Music, Language, and the Brain*. Oxford: Oxford University Press, 2008.
- Rizzolatti, Giacomo, Luciano Fadiga, Vittorio Gallese, and Leonardo Fogassi. "Premotor cortex and the recognition of motor actions." *Cognitive Brain Research* 3, Issue 2 (March 1996): 131–141.
- Rizzolatti, G., and C. Sinigaglia. "The Functional Role of the Parieto-frontal Mirror Circuit: Interpretations and Misinterpretations." *Nature Reviews Neuroscience* 11, no. 4 (2010): 264–274.
- Robbins, P., and M. Aydede, ed. *The Cambridge Handbook of Situated Cognition*. Cambridge, UK: Cambridge University Press, 2008.
- Said, E. W. *Musical Elaborations*. New York: Columbia University Press, 1991.
- Schutz, A. "Making Music Together." In *Collected Papers II: Studies in Social Theory*, edited by Arvid Brodersen, 159–178. The Hague: Martinus Nijhoff, 1964.
- Shove, P., and B. Repp. "Musical Motion and Performance: Theoretical and Empirical Perspectives." In *The Practice of Performance*, edited by J. Rink, 55–83. Cambridge, UK: Cambridge University Press, 1995.
- Smithers, T. "On What Embodiment Might Have to Do with Cognition (Technical Report FS-96-02)." In *Embodied Cognition and Action: Papers from the 1996 AAAI Fall Symposium*, edited by M. Mataric, 113–116. Menlo Park, CA: AAAI Press, 1996.
- Todd, N. P. M. "Motion in Music: A Neurobiological Perspective." *Music Perception* 17, no. 1 (1999): 115–126.
- Todd, N. P. M., D. J. O'Boyle, and C. S. Lee. "A Sensory-Motor Theory of Rhythm, Time Perception, and Beat Induction." *Journal of New Music Research* 28, no. 1 (1999): 5–28.

Vijay Iyer

Vijay Iyer, Harvard University

