



LEEDS
BECKETT
UNIVERSITY

Citation:

Ballan, H and Abraham, AG (2016) Multimodal imagery in music: Active ingredients and mechanisms underlying musical engagement. *Music and Medicine*, 8 (4). ISSN 1943-863X

Link to Leeds Beckett Repository record:

<https://eprints.leedsbeckett.ac.uk/id/eprint/3129/>

Document Version:

Article (Accepted Version)

The aim of the Leeds Beckett Repository is to provide open access to our research, as required by funder policies and permitted by publishers and copyright law.

The Leeds Beckett repository holds a wide range of publications, each of which has been checked for copyright and the relevant embargo period has been applied by the Research Services team.

We operate on a standard take-down policy. If you are the author or publisher of an output and you would like it removed from the repository, please [contact us](#) and we will investigate on a case-by-case basis.

Each thesis in the repository has been cleared where necessary by the author for third party copyright. If you would like a thesis to be removed from the repository or believe there is an issue with copyright, please contact us on openaccess@leedsbeckett.ac.uk and we will investigate on a case-by-case basis.

Title: Multimodal imagery in music: Active ingredients and mechanisms underlying musical engagement

Abstract: Clinicians and researchers have provided strong evidence for the efficacy of Guided Imagery and Music (GIM) and similar therapies across a wide range of clinical conditions. What is still lacking is a theoretical framework that would allow identification of the 'active ingredients' in this process. This paper seeks to introduce a new systemic framework for investigating such therapies by examining the biological roots as well as the role of music in the regulation of individual and social life to maintain homeostasis via multimodality by means of arousal, imagery, attentional engagement, emotion, memory and analogous processes. Taking the work of Edelman, Damasio and other leaders of modern neuroscience as a point of departure, homeostasis and multimodality are presented as essential not only to the human life process in terms of our active mental life but also to the fullness of Edelman's "primary consciousness" and Damasio's "core self." The implications of these intricate cross-connections are considered as well as the unique propensity for music to spontaneously and multimodally engage these connections. Proposals to evaluate these ideas and stimulate further research in both basic science and clinical practice are made.

Keywords: music, multimodal, imagery, homeostasis, consciousness, music therapy

Clinicians and researchers have presented strong evidence for the efficacy of Guided Imagery and Music (GIM) and similar or related therapies [1]–[4]. For such therapies to be successful, it has been argued that *balance* is necessary, balance of all kinds – among senses, among thoughts and feelings, between memories and current perceptions, between therapist and client. In the relationship between memories and current perceptions, both are necessary and indeed related; neither must overwhelm the other. Similarly, in the sensory realm the modalities must be integrated for the therapeutic effect to be enhanced rather than diminished by cross-modal influence.

Music therapy is an intensely interpersonal and multimodal experience. It often involves a great deal of movement, some of it small muscles around sensory portals, some subtle head turning, sometimes clapping, singing, vigorous playing of instruments, and even dancing. As described in more detail below, we propose that, in order for GIM therapy to be successful, multimodal neurons and neurons susceptible to cross-modal influences, in both subcortical and cortical areas, must cooperate to promote homeostasis (defined below), which we propose as the goal of therapeutic musical interventions. This variable is not an ideal candidate for studies based on neuroimaging studies, for instance, in which both the therapist – client relationship and actual movement are left out of the study.

A recent review article reported on the efficacy of GIM therapy stated that, “The studies reveal that GIM therapy reduces activation of the HPA stress axis, but they fail to isolate effects of music *per se*. The therapist-participant interaction is presumably a contributing factor, as well” [5]. So although GIM therapy acts on the Hypothalamic-Pituitary-Adrenal axis of the stress response by moving the client towards homeostasis, by reason of limitations within the experimental designs, the studies discussed in the review article fail to identify what it is about the multimodal experience of music that is therapeutic. What indeed are the ‘active ingredients’ that facilitate this therapeutic process? In this article, we address one candidate of such an active ingredient: multimodality, and in particular, the interaction of interoception with other sensory modes (vision, hearing, touch) and movement. The central point being made is that it is possible that the musical stimulus may be only one of many active ingredients in music therapy as it is practiced, others of which may be also critical for positive outcomes. That is one reason for the disconnection between studies of music cognition with its focus on music processing, particularly those studies based on neuroimaging, and the actual practice of music therapy. The article goes on to propose that the therapeutic effect of music (not just GIM) is likely to be its ability to regulate stress and arousal by modulating brainstem-mediated processes. The authors note that in addition to modulating stress, attentional processes, arousal and emotions, brainstem activation through music mediates sensory and motor function [5].

We will refer to the “balance” alluded to in the initial paragraph, in its biological aspect, as ‘homeostasis’ which is a common usage of the term [6]. In this paper we explore how homeostasis and related attentional processes can be promoted by multimodal experiences of music, how cross-modal influences may enhance the therapeutic effects achievable through efforts targeted at isolated modalities, how value-laden memories and current multimodal perceptions may be correlated to produce a “scene” of objects and events [7], and how the instantiation of embodiment combined

with the goal of homeostasis makes the study and practice of multimodal therapies a necessity [8], [9].

Homeostasis: From Single Cells to Nervous Systems

Homeostasis refers to the maintenance of an equilibrium or balance between interdependent elements by means of physiological processes. The loss of homeostasis (one way of describing stress) provokes a neurochemical response driving an organism, in principle any organism, to take steps to restore it. Choices by organisms that reduce stress (choices that increase homeostasis) may safeguard health and well-being. Indeed, one evolutionary function of music may be to restore homeostasis [5]. In surveying the world of living things, the drive to homeostasis should come as no surprise. Even single cells appear to seek a balanced state or internal milieu necessary to sustain life in the face of changing external circumstances, whether it is the nucleus and cytoplasm of a single cell constantly changing their internal balance or a multicellular organism doing the equivalent at a higher order of complexity [6], [9].

A single cell may be an island, entire of itself, attending to its own homeostasis without communicating with other organisms. Maintaining homeostasis in the “island” environment of single-cell organisms is complex enough. However, no neuron is an island, and the homeostatic demands of neuronal groups far exceed, in complexity and number, those of single cells [7], [9]–[11].

Neurons transmit signals, and are organized in networks or cellular societies. So much so that we even speak about “neuronal coalitions.” Neurons are cells that influence the function of other cells by special forms of signaling that are unique to neurons. So a neuron is a cell, but it is not like other cells in the body as it is, metaphorically speaking, a ubiquitously ‘social’ cell, and its homeostatic drives, singly and in groups, reflect that circumstance. So the nervous system is an organic body system unlike other body systems in the degree and kind of its intercellular communication [7]. This fact about the nervous system is at the heart of the relationship between music and imagery. It is where multimodality, cross-modal influence, the relationship between memory and perception, and the complex relationship among motor plans, efference (“feed forward”) copies of those plans in the sensory system, and motor action all take root. After a discussion of some unique challenges faced by complex organisms with integrated motor and sensory systems, we shall turn our attention to a remarkable structure in the brainstem that is crucial to multimodal contributions to homeostasis such as those that occur in GIM and related therapies – the superior colliculus – and then to cortical structures that integrate sensory modalities and to the attentional processes that are affected by the same.

Homeostasis and Multimodality: Drawing from Edelman and Damasio

Gerald Edelman and Antonio Damasio have given complex accounts of how internal homeostatic systems in the organism link up with signals from the world [7]–[9]. In the discussion that follows, we predominantly focus on Edelman’s account.

In Edelman’s influential model of the “primary self,” internal homeostatic systems operate in the brainstem, the hypothalamus and in other autonomic centers. These internal states are registered in the hippocampus, the amygdala and the septum, *limbic* regions that have access to multimodal input from primary and secondary cortical areas such as vision, hearing and touch. These limbic areas, in turn, are linked by pathways to special value-laden memories in frontal, temporal and parietal cortical areas, implying previous experience and neuronal group selection in those areas.

The model of internal homeostasis is complicated by the fact that internal homeostatic systems are not only a product of the relationship among internal states, but also are affected by current perceptions. The model gets very complex when input from the external world is fully taken into account. What Edelman refers to as “world signals” (which include proprioception), link to primary and secondary cortical areas such as vision, hearing and touch, and these areas have input to the limbic areas like the hippocampus, the amygdala and the septum. However, the processes as described thus far do not yet give rise to the here-and-now experience of primary consciousness. That experience requires what Edelman refers to as “reentrant loops”, pathways that connect multimodal primary and secondary cortical input to value-laden memory and, in turn, to current perception.

Edelman (1989) defines reentry as: “The dynamic ongoing process of recursive signaling across massively parallel reciprocal fibers connecting maps. This process results in binding and is the basis for the emergence of consciousness through the workings of the dynamic core. Allows coherent and synchronous events to emerge in the brain; that is, it is the basis for spatiotemporal correlation.”

This definition requires unpacking. The *recursive signaling* occurs between groups of neurons, which can vary tremendously in terms of size (50 to 100,000). Without neuronal groups, that dynamic ongoing process of recursive signaling would not occur. It is similar to the parallel processors in your computer. The signaling occurs between processors, not single elements of processors.

The *maps* referred to are neuronal groups that represent aspects of the external world in the brain. *Binding* refers to the manner in which various aspects of one sensory modality (form and color for vision, or pitch and rhythm for hearing, for example) are integrated. It also refers to the way the separate modalities are integrated. Reentry is thus critical to binding; binding is the basis of multimodality; and multimodality permits the integration of interoception with other senses, resulting in qualia (defined later).

The *dynamic core* is a neuronal group where the “present” or “here and now” is created in the brain. Because of the 100 to 250 milliseconds delay in reentrant signaling that creates the present, Edelman refers to our everyday experience of the present as “the remembered present.” To review, a stimulus enters the brain through sensory portals (eyes or ears, for example). Within 100 to 150ms that single stimulus has correlated with perhaps 1 million electrical and chemical signals within widely distributed and even remote neuronal groups.

The *coherent and synchronous events* to which Edelman refers can be understood as a form of “integration” of signals (as opposed to differentiation). In the 1970s, Edelman stressed the interplay of both integration and differentiation in reentrant signaling which aligns well with recent proposals that consciousness arises from a balanced combination of integration and differentiation in particular brain regions [12]. Another basic point of Edelman’s – the problem of the paucity of input from sensory portals to explain perception as compared to the richness of brain signaling in response to single external stimuli is very much a subject of contemporary discussion [13].

These mechanisms and explanations form the crux of reentrant processing.

So, in Edelman’s model, internally generated past signals and externally generated category signals synchronize, probably in thalamo-cortical and cortico-cortical loops, leading to memory and conceptual areas. It is this memory (both recent and remote) that is linked by reentrant paths to current perception of world signals. The result, primary consciousness, occurs only when multiple modalities give rise to a “scene” made up of objects and events which coexist in memory and perception, though they may not be causally connected. Notwithstanding the absence of causal connection, primary consciousness may connect the objects and events in the scene through memory and its previous, value-laden experience [7], [14]¹. In this view, homeostasis and multimodality, for both of which the brainstem contains important structures, are essential to the fullness of primary consciousness.

What so captured the imagination of Edelman’s contemporaries was his identification of integrative activity unique to complex nervous systems via the reentrant signaling loop connecting current input from all of the primary and secondary sensory cortices to great stores of value-laden memory in the frontal, temporal and parietal cortices. This seemed to connect the homeostatic process (generally associated with the brainstem) to widely distributed parts of the brain especially neocortex and the outside world; it linked current perception with memory, and sensory modalities with each other, as no previous theory had. In wonderment at the possibility of connecting practically all parts of the brain to each other with no homunculus controlling the process, in awe at the

¹ Crick and Koch (2003) present a similar theory of synchronized neuronal coalitions. However, their theory is focused on widely distributed cortical areas without Edelman’s implied self-process involving the subthalamic nuclei. In other words, Edelman implies a theory of consciousness and of self-process, whereas Crick/Koch present a theory of consciousness without including emotion-based signals (Damasio’s refers to these as “somatic markers”), that would be required for a theory of self. Others include Panksepp (2005) and Llinás (2002) who present theories of consciousness and of self.

theory's rigorous scientific exposition and at the same time its possibilities for illuminating the human condition, Oliver Sacks was moved to proclaim, "Thank God I have lived to hear this theory" [15]. In an interview with the BBC, Edelman found a metaphor that clarifies the thrust of the theory, so helpful for our purposes in explicating homeostasis and multimodality, but with much broader implications. He likened the synchronization of so many activated neuronal groups in widely separated parts of the brain to hidden signals going back-and-forth among members of the string quartet:

"Think: if you had a hundred thousand wires randomly connecting four string quartet players and that, even though they weren't speaking words, signals were going back-and-forth in all kinds of hidden ways [as you usually get them by the subtle non-verbal interactions between the players] that make the whole set of sounds a unified ensemble. That's how the maps of the brain work by reentry."²

Antonio Damasio has proposed a theory of the "self" that is important to the ensuing discussion. According to Damasio, input from all the senses but one is subject to a Jamesian stream of consciousness, which refers to the conscious flow of thought in our minds. The exception is interoception, our perception of internal body states, which is remarkably stable in a healthy person. Damasio posits that bidirectional signals to and from the body provide a stable basis through interoception for "primordial feelings," including the feeling of a self [8], [9]. As it stands, this is a highly persuasive theory of the self and of the self-mind distinction³.

How Damasio's views complement those of Edelman's can be gleaned from their view on *qualia*. Edelman defines qualia as follows, "Qualia; qualia: terms used to refer to the "feel" of conscious experience – "what it is like to be x," where x, for example, is a human or a bat. I use the term "qualia" as coextensive with conscious experience. Consciousness reflects the integration among vast numbers of qualia. Qualia are discriminations made possible by the activity of the reentrant dynamic core."

Qualia, in other words, are feelings (generated by interoception, as we shall see below) which, when integrated in vast numbers, make consciousness possible. Damasio carefully distinguishes between two qualia problems. The Qualia I problem is: what is a feeling and why do we have them. The Qualia II problem is: how do qualia arise biologically. The answer offered by Damasio to the Qualia II problem and argued in a series of four carefully written books and numerous articles is that feelings are generated by electrical or chemical signals from the body to the brain – interoception and proprioception -- arising from, among other things, movements, including micromovements (see also Rudolfo Llinás, 2002). These movements and other body-brain signaling provide massive amounts of information to the brain, which integrates with input from sensory portals to create feelings correlated with that sensory input.

² Quoted in Sacks (2015).

³ Others include Crick and Koch (2003), and Llinás (2002).

Feed Forward Loops and Working Memory

What has been less of a focus within the aforementioned theoretical accounts is how multimodal sensory experiences are coupled to movement and action. We (or our brains) have sensory and motor systems which feed forward to predicted and imagined movement (see below) and feed back to actual body states. We make motor plans and revise them through complex loops of Bayesian prediction and revision—prediction and revision that always involves a connection between motor and sensory systems and their link to imagery.

Feed forward loops have been extensively studied in the motor system [16] and explored in connection with multiple aspects of human creativity [17]. Voluntary movement begins with an idea (not a simple construct), which appears in the association cortex as a plan or program. The association cortex receives somatosensory input regarding the planned action before the association cortex communicates with the motor system (the basal ganglia, supplementary motor area, premotor cortex and lateral cerebellum). In other words, the plan to move is enough to evoke a somatosensory response; a non-sensory idea “feeds forward” to a response. Each part of the aforementioned motor system communicates with the motor cortex, whose job is to activate muscles for action. The motor cortex also receives somatosensory input before action. In addition, the intermediate cerebellum, which implements and corrects movement patterns, receives somatosensory input before the movement is implemented. In these multiple and complex ways, ideas about action give rise to somatosensory input regarding movement before the movement actually occurs. We posit that similar operations are in play with music and images in GIM Therapy. Ideas about music and/or images may feed forward to sensory input (from the auditory or visual system) before actual sensory input enters the brain through a sensory portal such as the eyes or ears [18].

Indeed, insights derived from the motor system via feed forward loops have even been applied in the context of understanding creativity to explain deliberate and spontaneous modes, as determined by explicit neural systems, as well as the mode of flow experience which comes about through implicit systems. The degree to which moments of insight are seen, “sighted” or present in working memory which involves maintaining information in the moment in service of a specific goal, differ across these modes, as does the prediction process in which mental imagery involves either simulation, a process in which mental imagery regarding future movement is sketchy and without detail, or emulation, in which mental imagery is step-by-step and precisely detailed [17]. Similarly, in GIM and related therapies, the explicit and implicit systems may be used (depending in part on the role of verbal instructions), predictive mental imagery may be more or less detailed (as in simulation and emulation) and the amount of time images spend in working memory may vary. “Sightedness” and working memory are closely related to issues of attention, which are addressed next.

Multimodality and Attention

Especially in more primitive organisms and self states, emotions may be understood in terms of action tendencies, particularly actions that are connected with restoring homeostasis and that are "felt" via interoception [9]. Interoception engenders images or neural patterns of internal body states and engages emotional systems (such as the limbic system) in higher animals, with chemical responses in neuroendocrine systems and subcortical nuclei. There may be movements associated with these reactions, such as facial expressions. Positively and negatively valenced memories and plans may be invoked. Mental processing also changes – with sadness it slows down, with happiness it speeds up, and with these changes in mental processing, attentional focus changes. Specifically, it narrows (for example, to the sad or happy feeling) [19]⁴.

Future research should investigate whether, as we have suggested, multimodality enhances or provokes these attentional effects by changing the speed of processing (from mode to mode, depending on the degree of sensory integration), and whether these changes alter attentional focus, perhaps broadening it from the fixation of excessive focus to a more flexible use of attention [20]⁵.

Emotion affects and is derived not only from the relatively stable, interoceptively originated self, but also from the Jamesian stream of consciousness. For example, the feeling of knowing an external object (a case where interoceptively derived feelings may integrate with other sensory modalities) creates salience for that object, a kind of emotionally colored, focused attention, attention colored by feeling [9]. The synchronization of a particular set of perceptions and feelings would presumably also be affected by the timing of presentation of stimuli [21]–[23].

As objects and the modalities used to engage them change, the salience of the objects changes, and likewise the feelings associated with those objects. It is possible (and seems probable based on anecdotal evidence) that these changes can be induced clinically by multimodal presentation of stimuli. We suggest this as an area of further research.

One mechanism that may be important to the clinical success of multimodal therapies is the modulation of the default mode network, the network of brain regions that is active when the brain is at rest [24]–[26]. In this case, there is little stimulus-directed attention [27]–[29]. It is possible this more diffuse attentional state can be facilitated by multimodal presentation of stimuli. Anecdotally, the mind wandering and daydreaming reported by patients in GIM therapy supports this hypothesis and suggests it may be an important area for future research.

⁴ The concept of a “somatic marker” refers to an emotion-based signal that affects perception and cognition, as basic to theories of the self. For similar ideas, see also Panksepp (2005), Llinás (2002) and Edelman (1989).

⁵ This might occur by “somatic marking” similar to that described by Damasio (2010).

Sensory Integration versus Sensory Interference

We have noted that for GIM and related therapies to be effective, the brain must integrate information from the different senses. One possibility is that in multimodal experience, one sense dominates the others, as in the McGurk effect in which visual input dominates auditory input [30]. In some disorders (autism, PTSD), it is sometimes difficult for auditory input to be integrated with other sensory input, and patients may exhibit exceptional sensitivity to sound. This kind of dominance or interference is not likely to be the source of positive therapeutic outcomes when music and images are combined. Another possibility is that the input from the different senses enhances the accuracy of perception in each. Rather than interfering with one another, the senses are integrated and enhance each other [31]–[34].

The subcortical area of interest for sensory integration is a portion of the brainstem called the superior colliculus [35]. This area creates superimposed maps (neural patterns) of visual, auditory and somatosensory inputs. Individual cells in the superior colliculus may be activated by more than one of those sensory modalities. As a consequence, the response of individual cells may be stronger when there is input from more than one sense. Specifically, responses to auditory, visual or somatosensory input may be enhanced. For example, if the response to an auditory input is too weak to be perceived or to have a therapeutic effect, the addition of a simultaneous visual or somatosensory input may enhance the auditory response and, presumably, the therapeutic efficacy in a clinical setting. Many cortical areas, too, especially the superior temporal sulcus, but also regions of the parietal and frontal lobes and the hippocampus, may exhibit similar amplifying or synergistic effects: input from one sensory modality may enhance the cortical response to simultaneous input from one or more other modalities [36].

To complicate matters, in other circumstances, input from different modalities may produce interference effects, as in the examples of autism or PTSD referred to above [37]. We recommend further study of sensory integration in GIM and related theories to discover the detailed mechanisms underlying both integration and interference between modalities in clinical practice.

Movement Changes Everything – The complexity of an integrated sensory and motor system

In the study of imagery and sensory integration, movement changes everything. The involvement of primary and secondary cortical areas when sitting and listening to music is well-established [38], [39]. When listening in naturalistic settings though, the motor system is also actively involved, and the complexity of processing increases enormously. Movement systems must interact with sensory systems. If music produces or is produced by movement (or is responded to by imaginary movement), then sensory systems,

predominantly visual and somatosensory as well as auditory, are likely to be involved. In such contexts, then, music cannot be separated from imagery. An auditory or somatosensory image activated by music is likely to trigger a visual image [40]. Just sitting comfortably and listening to music (even without the apparatus of GIM) likely involves rich and complicated bi-directional somatosensory and interoceptive processing and signaling as well as auditory (and probably visual) imagery.

We see, we hear, we touch, we move. The increase of complexity resulting from this last fact is only hinted at above; it is almost incomprehensible. Movement increases the number of activated neurons but, more importantly, the ever-larger circuits in which they are organized, from single neurons to complex brain networks, all set in motion (so to speak) by the evolutionary necessity to move—away from predators and towards prey and mates [16]. More neurons mean more complex patterns of internal and external communication.

Movement allows organisms to evolve from being passive recipients of the environment acting upon them (seeing, hearing, touching) to being intentional agents (selves) who act upon and shape this environment and cause other organisms to see, hear and feel the products of their action. Indeed, the biological systems giving rise to the subjective experiences of “music and imagery” may originate (or be an epiphenomenon of) the fact that we are creatures who move.

As a consequence of the complexity involved in the global workspace that is brought into play by music and imagery, which is not modular but rather networked and connected both within the brain and between the brain and body, we will never be able to approach the full richness of music and imagery by looking exclusively at individual locations in the brain. The problem involves a nonlinear complexity (multiple causes and effects at vastly different scales) deriving from the fact that we are animals that have highly developed systems and networks that are refined and interconnected to implement the problems posed by multisensory experience in general, and by music and the quest for homeostasis in particular.

In sum, a moving organism *requires* images or imagery, neural patterns derived from sensory experience, memory and prediction to assist movement, whether auditory, visual, somatosensory (including proprioception and interoception) vestibular, olfactory, gustatory or other.

Damasio posited that non-visual images of internal body processes may give rise to a “core self” of here-and-now patterns, complex emotions and eventually, social intelligence [9]⁶. The core self of here-and-now may expand to ever greater capacities for memory and prediction, for imagination and, ultimately, the development of an autobiographical self. Perhaps GIM therapy, by employing and enhancing sensory integration, facilitates this development in individual clients.

⁶ This has parallels with the concept of “primary consciousness” as described in Edelman (1989).

The Development of Images and of Mind: From Core Self to Autobiographical Self

Assuming the existence of a self, a stable process originating in perception and homeostasis and contrasted with a mind of more or less synchronized, flowing images, it remains to be seen how primitive forms of that self relate to more advanced forms in which such operations as autobiographical memory participate. As we saw with the superior colliculus in the brainstem, when smoothly integrated images, originating in different modalities and colored by somatic markers, enhance rather than interfere with one another, when vision increases auditory acuity and vice versa, homeostasis is enhanced; life can regulate itself through sensory integration, feedback and feed forward loops, through memory, prediction and imagination. Behavior assisted by integrated multisensory images (some raw and perceptual, some imagined, some remembered, some predicted) could become more flexible, even creative. The emergence of the neural patterns that we refer to as mental imagery reflects simultaneously the emergence of mind.

Because we see mind as rooted in embodiment, in body-brain interaction, in the senses and in multimodality, and, in particular, in the contribution of the somatosensory and interoceptive systems to vision and hearing at subcortical and cortical levels, our preferred definition links “mind” directly to images as neural maps of variously combined sensory modalities in both memory and perception. Without these images—of the body within and the world without— we suspect the successful quest for homeostasis would not have been possible.

Once neurons form “networks,” “coalitions,” and “societies” of communicating cells, the vast increase of complexity makes possible diverse behaviors among different organisms of a species. The loops between value-laden memory and current perceptual categorization are unique for each individual. We strive through sociality to appreciate and empathize with precisely that individuality. Perhaps this is the source of creativity, self-reflection and consciousness. Jill’s multi-modal experience of a musical concert is different from Jane’s, and their body-brain experiences and resulting somatic markers differ [41]. The memories that enter (or “re-enter”) their conscious experience differ. For example, they may see and hear more or less the same thing but feelings arising from viscera and internal milieu, the location of limbs and associated memories from all senses may be highly individual. Out of such differences, and our eventual ability to understand them, each other and ourselves, conscious experience may flourish. The multimodal and in particular interoceptive contributions to our sense of ourselves and our memory-perception loops are enriched by encounters with others, the attempt to understand their subjectivity as well as our own, and the development of theory of mind which is both a cause and an effect of this process.

The Autobiographical Self and Sociality

The autobiographical self (to the extent that it exists as a stable process apart from minds) requires that perceptions, constituted by groups of cooperating neurons, be stored systematically as memories [9]. At least two kinds of memory are implicated: a long term memory in which neuronal societies hold together neuronal records of (multimodal) events, and a working memory in which neuronal societies store memories temporarily in order to work with them in real time so they may be communicated to other persons who have their own memories. Movement, vocalization, language, music all require multiple forms of multimodal storage to communicate what is felt, thought and believed by other persons as their autobiographical selves interact with other selves and the world.

Societies of neurons may give rise to societies of persons who communicate through all sensory modalities, usually more than one at a time [42], [43]. This is not only true of ritual and ceremony. Even when speaking on the telephone, the listener feels the phone in his hand, the clothing on his body, the temperature in the room, various interoceptive processes (e.g., digestion of food) as well as strong sensory experience from imagined seeing, etc., in addition to hearing. And it is memory—sensory, motor, long-term, short-term, semantic, autobiographical—colored by feelings, that holds it all together to form what we call consciousness.

Multimodal Experience, Memory and Art

Consciousness has consequences, one of which is reflection. Human beings reflect on all manner of things, from the origins of life to its meaning, from pragmatic calculation to morality. What is it that guides truly deliberate reflection? One indispensable component is the autobiographical memory to which we just referred—what each of us has been, seen, thought, done—memory that is multimodal and no sooner recalled than reconsolidated under new cross-modal influences. Once theory of mind engenders societies of neurons; once our own sensorimotor-derived subjectivity is matched by a deepening understanding of similar patterns in other persons, our reflections on social life, too, are guided by multi-modal memories and their vicissitudes. Our telling, singing or chanting of stories filled with references to images in multiple modes both narrates and helps to create autobiographical minds (selves) that are tightly stitched together and come to our service in collective deliberation over questions great and small. We have learned from Edelman and Damasio, among others about the intricate ways in which value-laden memory and current, multimodal perception interact.

As one clinical illustration of the significance of Edelman's model, which implies a high degree of neural plasticity that would be facilitated by rich multimodal input, we take combat-related posttraumatic stress disorder (PTSD). In our experience, three cognitive distortions seem to be especially common in combat-related PTSD: others are bad, I am bad, the world is dangerous. The experience of making music in a group, a

multimodal experience that may evoke value-laden memories in conjunction with current multimodal perceptions, seems to have a profound effect on those distortions, an effect that may be mediated by the elevated levels of oxytocin that are produced during music making and that promote social affiliation [5]. In a recent study, higher oxytocin levels were shown to enhance compassion of people suffering from symptoms of PTSD [44]. Oxytocin may be only part of what promotes healing in therapeutic musical interventions. It may also be that self-understanding of the primary or core self changes under the influence of a rich and interesting multimodal stimulus that triggers memories, perceptual inputs, entrainment effects, as well as shared, mirrored and/or synchronized movements in a way that erodes the cognitive distortion that others are bad. When that cognitive distortion is reduced, it appears that clients will tend to see themselves as less bad and the world as less dangerous. Similar mechanisms, in which group music making facilitates social affiliation, may be at work in healthy individuals who synchronize their music making and movements, and to some extent their perceptions, in a way that may facilitate empathy and lead to greater social affiliation [45].

Added to the pleasure of our multimodal experience is the feeling that art, and particularly making or enjoying art together, is “a way to explore one’s mind and the mind of others” [9]. Multimodal artistic experience may be accompanied by feelings of well-being and sociality. None of it would be possible without those rich multimodal experiences that we refer to as “imagery.” If we are able to co-exist as social beings, if we can imagine a future and work towards its realization, it is likely because of the richness of multimodal experience, combined with the informational richness of our memory. These allow us to transcend the here-and-now through imagination, itself a kind of feed-forward loop from ideas to sensory inputs, then back to ideas and sometimes to actions.

Conclusion: Bridging the Gap

In sum, the purpose of the paper is to take some of the original insights that constitute the foundation of modern neuroscience and apply them to music therapy – as a theoretical framework, as a direction for new research, and as a set of suggestions for clinical practice.

The clinical success of GIM and related therapies is not in doubt [1]. We are reminded of it every day in hospitals, clinics and private practices around the world. In the midst of this success, we suggest a pause to consider the diverse and complex underlying mechanisms that account for it in the hope that this return to basic science will result in an even better clinical practice and will enrich our knowledge of some of the most basic processes of life.

Not very long ago, *musicking* was a verb. The current practice of listening to music without simultaneously dancing, singing and playing is a historical and cultural anomaly. For one million years, dating to our pre-hominid ancestors, music was a richly multimodal experience [46]–[48]. The “homeostatic drive” behind music manifests itself

even now, from churches to electronic dance clubs. What we need to understand better is its workings.

By placing music therapy within an empirical framework in which the active ingredients can be identified, we have highlighted in need for new directions of the research in the practice of music therapy. Future research should focus on investigating diverse variables, such as the therapist-client relationship, as well as movement and other ‘non-musical’ candidates for active ingredients in facilitating success in music therapy. Should the findings bear out, the clinical practice may need to be transformed into an even more movement – oriented one. The evolution of GIM therapy provides one example of such a change in direction. Helen Bonney, the inventor of GIM, introduced “body listening” into GIM late in her career, prominently adding systematic movement to GIM therapy. Body listening and other forms of GIM incorporating movement are now widely practiced. One might ask – are the music and images of GIM the main active ingredients in “body listening” or does the movement predominate? Most likely, a combination of auditory output and motor input is responsible for positive outcomes, as evidenced in Michael Thaut’s work [49], [50]. While such work has stressed the importance of actual movement in carefully designed studies, we propose that movement in music therapy is part of a multimodal experience that crucially involves feelings and action tendencies, mediated through interoception. Although different, this approach is consistent with and indeed complementary to that of Thaut.

Notwithstanding all of the work in neurologic music therapy emphasizing the auditory motor-connections mediated by rhythm [51], we need to develop a thorough understanding of all the active ingredients in music therapy, and these need to be evaluated in an unbiased manner. For instance, how do music, imagery and movement interact in facilitating therapeutic outcomes in PTSD? Are the best outcomes associated with a combination of all three factors? This would be evidence in favor of the additive effect of all three factors. Of the dual combinations – music and imagery/movement and imagery/music and movement – which has the most advantageous outcomes, and is it context dependent? Such findings would give us insights about which aspects of the interventions to emphasize in order to target specific therapeutic outcomes. These are examples of the basic questions and directions we hope to emphasize in our article particularly given their experimental and clinical implications.

Research on music therapy is challenging as we need to tease apart the therapeutic impact of (i) the therapist – client relationship, (ii) the auditory stimulus of music, (iii) the complex effects of speech during sessions, (iv) the mental images, (v) the sensorimotor experiences, including positioning of the body and limbs, proprioceptive, interoceptive and vestibular inputs, not to mention the micro-movements that accompany listening. The project we propose is a collaborative effort to design and carry out modest, reductionistic experiments using randomized controlled trials where possible. These would involve carefully selected active and passive control groups and strict inclusion/exclusion criteria, with the best available outcome measures. In this way, we can tease out each element in the music therapy experience in order to maximize the

context-specific effectiveness of our therapeutic interventions while enriching our understanding of what it is to be human, and how our needs, pleasures and sense of well-being are intricately connected.

References

- [1] K. E. Bruscia, Ed., *Case studies in music therapy*. Phoenixville, PA: Barcelona Publishers, 1991.
- [2] K. E. Bruscia and D. E. Grocke, *Guided imagery and music the Bonny method and beyond*. Gilsum, NH: Barcelona Pub., 2002.
- [3] C. H. McKinney, M. H. Antoni, M. Kumar, F. C. Tims, and P. M. McCabe, “Effects of guided imagery and music (GIM) therapy on mood and cortisol in healthy adults,” *Health Psychol*, vol. 16, no. 4, pp. 390–400, Jul. 1997.
- [4] M.-F. Lin, M.-C. Hsu, H.-J. Chang, Y.-Y. Hsu, M.-H. Chou, and P. Crawford, “Pivotal moments and changes in the Bonny Method of Guided Imagery and Music for patients with depression,” *J Clin Nurs*, vol. 19, no. 7–8, pp. 1139–1148, Apr. 2010.
- [5] M. L. Chanda and D. J. Levitin, “The neurochemistry of music,” *Trends Cogn. Sci. (Regul. Ed.)*, vol. 17, no. 4, pp. 179–193, Apr. 2013.
- [6] C. De Duve, *Blueprint for a cell: the nature and origin of life*. Burlington, N.C: N. Patterson, 1991.
- [7] G. M. Edelman, *The remembered present: a biological theory of consciousness*. New York: Basic Books, 1989.
- [8] A. R. Damasio, *The feeling of what happens: body and emotion in the making of consciousness*, 1. Harvest ed. San Diego, CA: Harcourt, 2000.
- [9] A. R. Damasio, *Self comes to mind: constructing the conscious brain*, 1st ed. New York: Pantheon Books, 2010.
- [10] R. R. Llinás, *I of the vortex: from neurons to self*, 1. paperback ed. Cambridge, Mass. ;London: MIT Press, 2002.
- [11] J. Panksepp, *Affective neuroscience: the foundations of human and animal emotions*. Oxford: Oxford Univ. Press, 2005.
- [12] C. Koch, M. Massimini, M. Boly, and G. Tononi, “Neural correlates of consciousness: progress and problems,” *Nat Rev Neurosci*, vol. 17, no. 5, pp. 307–321, May 2016.
- [13] M. A. Cohen, D. C. Dennett, and N. Kanwisher, “What is the Bandwidth of Perceptual Experience?,” *Trends in Cognitive Sciences*, vol. 20, no. 5, pp. 324–335, May 2016.
- [14] F. Crick and C. Koch, “A framework for consciousness,” *Nat. Neurosci.*, vol. 6, no. 2, pp. 119–126, Feb. 2003.
- [15] O. Sacks, *On the move: a life*, First edition. New York: Alfred A. Knopf, 2015.
- [16] D. M. Wolpert, Z. Ghahramani, and M. I. Jordan, “An internal model for sensorimotor integration,” *Science*, vol. 269, no. 5232, pp. 1880–1882, Sep. 1995.

- [17] A. Dietrich, *How creativity happens in the brain*. Houndmills, Basingstoke, Hampshire ; New York, NY: Palgrave Macmillan, 2015.
- [18] G. Plomp, A. Hervais-Adelman, L. Astolfi, and C. M. Michel, “Early recurrence and ongoing parietal driving during elementary visual processing,” *Sci Rep*, vol. 5, p. 18733, 2015.
- [19] A. R. Damasio, “The somatic marker hypothesis and the possible functions of the prefrontal cortex,” *Philos. Trans. R. Soc. Lond., B, Biol. Sci.*, vol. 351, no. 1346, pp. 1413–1420, Oct. 1996.
- [20] D. Talsma, D. Senkowski, S. Soto-Faraco, and M. G. Woldorff, “The multifaceted interplay between attention and multisensory integration,” *Trends Cogn. Sci. (Regul. Ed.)*, vol. 14, no. 9, pp. 400–410, Sep. 2010.
- [21] P. J. Laurienti, J. H. Burdette, M. T. Wallace, Y.-F. Yen, A. S. Field, and B. E. Stein, “Deactivation of sensory-specific cortex by cross-modal stimuli,” *J Cogn Neurosci*, vol. 14, no. 3, pp. 420–429, Apr. 2002.
- [22] B. E. Stein and M. T. Wallace, “Comparisons of cross-modality integration in midbrain and cortex,” *Prog. Brain Res.*, vol. 112, pp. 289–299, 1996.
- [23] M. T. Wallace and B. E. Stein, “Cross-modal synthesis in the midbrain depends on input from cortex,” *J. Neurophysiol.*, vol. 71, no. 1, pp. 429–432, Jan. 1994.
- [24] R. M. Braga, D. J. Sharp, C. Leeson, R. J. S. Wise, and R. Leech, “Echoes of the Brain within Default Mode, Association, and Heteromodal Cortices,” *Journal of Neuroscience*, vol. 33, no. 35, pp. 14031–14039, Aug. 2013.
- [25] P. Garcés, E. Pereda, J. A. Hernández-Tamames, F. Del-Pozo, F. Maestú, and J. Ángel Pineda-Pardo, “Multimodal description of whole brain connectivity: A comparison of resting state MEG, fMRI, and DWI: Multimodal Description of Whole Brain Connectivity,” *Human Brain Mapping*, vol. 37, no. 1, pp. 20–34, Jan. 2016.
- [26] C. Rosazza, A. Andronache, D. Sattin, M. G. Bruzzone, G. Marotta, A. Nigri, S. Ferraro, D. Rossi Sebastiano, L. Porcu, A. Bersano, R. Benti, M. Leonardi, L. D’Incerti, L. Minati, and on behalf of the Coma Research Center, Besta Institute, “Multimodal study of default-mode network integrity in disorders of consciousness: DMN Integrity in DOC,” *Annals of Neurology*, vol. 79, no. 5, pp. 841–853, May 2016.
- [27] D. A. Gusnard, M. E. Raichle, and M. E. Raichle, “Searching for a baseline: functional imaging and the resting human brain,” *Nat. Rev. Neurosci.*, vol. 2, no. 10, pp. 685–694, Oct. 2001.
- [28] M. E. Raichle, “The Brain’s Default Mode Network,” *Annu. Rev. Neurosci.*, May 2015.
- [29] M. E. Raichle and M. A. Mintun, “Brain work and brain imaging,” *Annu. Rev. Neurosci.*, vol. 29, pp. 449–476, 2006.
- [30] H. McGurk and J. Macdonald, “Hearing lips and seeing voices,” *Nature*, vol. 264, no. 5588, pp. 746–748, Dec. 1976.
- [31] N. P. Holmes and C. Spence, “Multisensory integration: space, time and superadditivity,” *Curr. Biol.*, vol. 15, no. 18, pp. R762-764, Sep. 2005.
- [32] C. Kayser, C. I. Petkov, M. Augath, and N. K. Logothetis, “Functional imaging reveals visual modulation of specific fields in auditory cortex,” *J. Neurosci.*, vol. 27, no. 8, pp. 1824–1835, Feb. 2007.

- [33] S. Kennett, M. Eimer, C. Spence, and J. Driver, “Tactile-visual links in exogenous spatial attention under different postures: convergent evidence from psychophysics and ERPs,” *J Cogn Neurosci*, vol. 13, no. 4, pp. 462–478, May 2001.
- [34] V. Romei, M. M. Murray, L. B. Merabet, and G. Thut, “Occipital transcranial magnetic stimulation has opposing effects on visual and auditory stimulus detection: implications for multisensory interactions,” *J. Neurosci.*, vol. 27, no. 43, pp. 11465–11472, Oct. 2007.
- [35] B. E. Stein, T. R. Stanford, M. T. Wallace, J. W. Vaughan, and W. Jiang, “Crossmodal spatial interactions in subcortical and cortical circuits,” in *Crossmodal space and Crossmodal attention*, C. Spence and J. Driver, Eds. Oxford ; New York: Oxford University Press, 2004.
- [36] M. S. Gazzaniga, R. B. Ivry, and G. R. Mangun, *Cognitive neuroscience: the biology of the mind*, Fourth edition. New York, N.Y: W. W. Norton & Company, Inc, 2013.
- [37] B. Odegaard and L. Shams, “The Brain’s Tendency to Bind Audiovisual Signals Is Stable but Not General,” *Psychol Sci*, Mar. 2016.
- [38] V. Alluri, P. Toivainen, I. P. Jääskeläinen, E. Glerean, M. Sams, and E. Brattico, “Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm,” *Neuroimage*, vol. 59, no. 4, pp. 3677–3689, Feb. 2012.
- [39] R. J. Zatorre and A. R. Halpern, “Mental concerts: musical imagery and auditory cortex,” *Neuron*, vol. 47, no. 1, pp. 9–12, Jul. 2005.
- [40] M. Lotze, “Kinesthetic imagery of musical performance,” *Front Hum Neurosci*, vol. 7, p. 280, 2013.
- [41] R. Kanai and G. Rees, “The structural basis of inter-individual differences in human behaviour and cognition,” *Nat Rev Neurosci*, vol. 12, no. 4, pp. 231–242, Apr. 2011.
- [42] W. J. Freeman, “The Physiology of Perception,” *Scientific American*, vol. 264, no. 2, pp. 78–85, 1991.
- [43] W. J. Freeman, *Societies of Brains: A Study in the Neuroscience of Love and Hate*. Hillsdale, N.J: Psychology Press, 1995.
- [44] S. Palgi, E. Klein, and S. G. Shamay-Tsoory, “Oxytocin improves compassion toward women among patients with PTSD,” *Psychoneuroendocrinology*, vol. 64, pp. 143–149, Feb. 2016.
- [45] D. A. Abrams, S. Ryali, T. Chen, P. Chordia, A. Khouzam, D. J. Levitin, and V. Menon, “Inter-subject synchronization of brain responses during natural music listening,” *Eur. J. Neurosci.*, vol. 37, no. 9, pp. 1458–1469, May 2013.
- [46] I. Cross, “The Origins of Music: Some Stipulations on Theory,” *Music Perception*, vol. 24, no. 1, pp. 79–82, Sep. 2006.
- [47] I. Cross, “The nature of music and its evolution,” in *The Oxford handbook of music psychology*, S. Hallam, I. Cross, and M. H. Thaut, Eds. Oxford: Oxford University Press, 2016, pp. 3–17.
- [48] G. Tomlinson, *A million years of music: the emergence of human modernity*, First edition. New York: Zone Books, 2015.
- [49] M. H. Thaut, “Chapter 13 - The discovery of human auditory–motor entrainment and its role in the development of neurologic music therapy,” in *Progress in Brain Research*, vol. 217, S. F. and F. B. Eckart Altenmüller, Ed. Elsevier, 2015, pp. 253–266.

- [50] M. H. Thaut, G. C. McIntosh, and V. Hoemberg, “Neurobiological foundations of neurologic music therapy: rhythmic entrainment and the motor system,” *Front Psychol*, vol. 5, Feb. 2015.
- [51] E. Altenmüller and G. Schlaug, “Chapter 12 - Apollo’s gift: new aspects of neurologic music therapy,” in *Progress in Brain Research*, vol. 217, S. F. and F. B. Eckart Altenmüller, Ed. Elsevier, 2015, pp. 237–252.