Investigating metaphors of musical involvement

Immersion, flow, interaction and incorporation

Oskari Koskela
University of Jyväskylä
Dept. of Music, Art and Culture Studies
P.O. Box 35, FI-40014, University of Jyväskylä, Finland
oskari.j.koskela@student.jyu.fi

Kai Tuuri
University of Jyväskylä
Dept. of Music, Art and Culture Studies
P.O. Box 35, FI-40014, University of Jyväskylä, Finland
kai.tuuri@jyu.fi

ABSTRACT

The concept of immersion, despite being relatively unknown within music research, presents a potentially productive way for understanding the well acknowledged phenomenon of "being drawn into music". This paper 1) discusses immersion as a metaphor for conceptualizing musical involvement by drawing on the research into video games and virtual reality and 2) aims to clarify the metaphor of immersion by utilizing the concept of image schema to analyze it in relation to alternative metaphors of flow, interaction and incorporation. The theoretical stance of the paper is based on the paradigm of enactive cognitive sciences, which stresses the bodily, constructive and interactive nature of experience. As a conclusion, the paper suggest several ways to consider the differences between the chosen metaphors based on their image schematic structures. In line with the enactive approach, it is suggested that the experience of immersion should be considered as a constructive activity of using music, thereby highlighting the view of experience as a skillful activity. All in all, the paper aims to offer one kind of approach for considering different experiences with media and to stress the role of metaphors in how we understand experiences.

CCS CONCEPTS

• **Applied computing** → **Sound and music computing**; *Arts and humanities*; • **Computing methodologies** → *Cognitive science*; *Theory of mind*;

KEYWORDS

music, immersion, experience, image schema, conceptual metaphor, enactive perception

ACM Reference Format:
Oskari Koskela and Kai Tuuri. 2018. Investigating metaphors of musical involvement: Immersion, flow, interaction and incorporation. In *Audio Mostly 2018: Sound in Immersion and Emotion (AM'18), September 12–14, 2018, Wrexham, United Kingdom.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3243274.3243293

AM'18, September 12–14, 2018, Wrexham, United Kingdom

ACM ISBN 978-1-4503-6609-0/18/09…$15.00
https://doi.org/10.1145/3243274.3243293

1 INTRODUCTION

Text

Music is first and foremost an experiential phenomenon. Thus, understanding the conceptualization of musical experience can be taken as the crucial task of music research. Since the advent of the so-called affective turn, musical experience has been considered mainly in terms of emotions. While the merits of this approach are undeniable, the focus on emotions has perhaps steered attention away from other kinds of views on musical experience. Moreover, as the concept of emotions has remained elusive, there has been some tendency to treat them as reified entities, thus neglecting the wider structure of the experience and adopting what John Sloboda [31] has called a "pharmaceutical model" of particular music causing particular effects on a passive listener.

The approach chosen here is to consider the various metaphors used in describing musical experiences. Instead of being mere arbitrary figures of speech, metaphors can be taken as the primary way we understand different things: We make sense of fleeting and abstract experiences by invoking our understanding of concrete interactions with the world. However, metaphors are not only ways of describing the experience, but they also direct our understanding of the experience by structuring it according to the logic of the metaphor [24]. As such, metaphors relate directly to our understanding of what the experience is like. This paper discusses four metaphors of musical experience, immersion, flow, interaction and incorporation, aiming to clarify what kind of experiences they describe.

Since immersion has received the least attention of these metaphors within music research, it will be the main concern of this paper. Despite being generally used to describe experiences of deep involvement, immersion is mostly associated with a sense of inhabiting a different world afforded by virtual reality and video games. In this regard, it has been questioned whether the metaphor of immersion should be extended to cover experience with different, namely non-ergodic,[1] kind of media [2]. To consider this question, the current paper draws from the enactive theory of cognition [34] to present a view of musical experience as an interactive achievement of constructing a world.

Ultimately, this paper aims to provide a means of understanding the conceptual particularities of the chosen metaphors, while it also proposes the use of image schemas as a tool for making explicit the conceptual ways and possibilities of describing human experience

[1]The term ergodic, as used in the context of interactive media, is meant to highlight the active role of the user in constructing the media. In his book on cybertext, Espen Aarseth defines ergodic literature as literature where "nontrivial effort is required to allow the reader to traverse text" [1, p 1]. Another somewhat common way of describing ergodic media is that with it, the actions of the user have an effect on the media and the media requires active input to come into being [2].

more generally. Moreover, the enactive approach is suggested as a potentially fruitful framework for approaching experience fundamentally as a constitution of embodied activity. We hope that the discussion will offer certain clarification and ideas about the ways of experiencing and describing experiences within music and media in general.

2 EXPERIENCE AS A PROCESS OF EMBODIED INTERACTION: THE ENACTIVE PARADIGM

Enactivism is a relatively recent research program which originally sought to reconcile insights of cognitive science with traditions focused on lived human experience, most notably phenomenology and eastern philosophy [34]. Since its original inception enactivism has diffused into several strands of research, enough to frame the approach in the present paper as an interpretation of the enactive paradigm rather than the paradigm itself, which nevertheless shares the idea of mind as emerging from the bodily action of an organism with the environment [13]. In many respects, enactivism can be taken as an attempt to integrate several philosophical and psychological ideas, such as the constructed nature of reality, coupling between action and perception and the essential role of emotion and experience in cognition, into one biologically grounded and phenomenologically plausible framework.

Taking its inspiration from the phenomenological tradition, enactivism holds that the world we live in is a world of meaning [34]. This is not to deny the reality of the physical world, but to stress the experiential aspect of being-in-the-world: Each organism aims to create and sustain itself which leads to the world appearing according to how it hinders or advances this basic drive. So we cannot experience, or even understand, the world outside this perspective, which turns objective physical reality into experienced, meaningful and valenced reality.

According to enactivism, meaning is primarily about interaction. Since the basis of life is in self-preservation, and self-preservation is tied to actions within the environment, the most basic way of how we understand the world is about what kind of possibilities for action it affords. This can be appreciated through the ecological psychology concept of affordance, which refers to opportunities of interaction in the environment and in its objects with respect to the sensory-motor capabilities of the organism [14]. Even though there is a certain disagreement over the exact definition of affordances the crucial idea in the context of this paper is that affordances posit meaning as a relational property of interaction between organism and environment. Likewise the action-oriented nature of affordances suggests that, prior to any conscious propositional conceptualization, our world is already made meaningful by our capabilities of interacting with it. Here one should consider action in itself as a form of understanding: Even though we tend to consider meaning in terms of propositions, our most primary way of understanding the world is by successful action within it.

As the world we live in is a meaningful world, and meaning is about action, it is justifiable to say that we create our experiential world by interacting with it. This is to highlight the active and constructive nature of experience. According to enactivism, there is no pre-given objective world which we could inhabit, but life is a process of "bringing forth" a world, which is also a process of generating a self [32]. Briefly stated, reality is a transactional construct, based on continuous interaction between organism and its environment.

The idea of constructed world, brought forth by individual beings interacting with an environment, raises the question of how we can still have a shared and stable sense of reality. A short answer would be that because we have certain kinds of bodily capabilities for action, we also have certain stable, or rather recurring, patterns of interaction with the environment. Through time, in both the evolutionary and developmental sense, these recurring patterns lead to structural congruities between ourselves and our environment, which are called structural couplings [26]. The idea of structural coupling relates not only to regularities in sensory-motor interactions with the environment, how perception is dependent on the actions of the perceiver, but also aims to illustrate the interdependency of living systems more generally. As such, the reality of human beings is in important ways also a social and cultural construction.[2]

Altogether, the enactive paradigm holds that that experience (and cognition) is, at its core, action grounded in the sensory-motor interaction with the environment. As this view posits no ontological gap between bodily action and mind it can be questioned how exactly the body gives rise to thought. One way of considering this is the theory of conceptual metaphor.

2.1 Continuity of Bodily Action and Thought: The Theory of Conceptual Metaphor

The basic idea behind the theory of conceptual metaphor is that our understanding is to a great extent structured by metaphors [24]. More precisely, we understand abstract things by utilizing our experiential understanding of concrete things: by metaphorically projecting the pattern of experience from concrete source domain to abstract target domain [20]. For example, the abstract concept of "idea" can be understood by our concrete experience of physical objects, hence the metaphor IDEAS ARE OBJECTS. This is illustrated in phrases such as "I gave you that idea" and "It's hard to get that idea across to him" [24, pp. 10–11].

Instead of being merely figures of speech, the theory of conceptual metaphor holds that metaphors structure our understanding by lending their logic from a concrete source to an abstract target. In the example cited above, the logic of physical objects leads to ideas being "given" and "carried". Therefore metaphors are not arbitrary but closely connected to the way we constitute our world. In a sense, things are what they are partly because of the metaphors we use to structure them.

Conceptual metaphors are based on image schemas. These refer to "recurring, dynamic patterns of our perceptual interactions and motor programs that gives coherence and structure to our experience" [20, p. xiv]. As Mark Johnson [20] emphasizes, an important aspect of image schemas is their nature as unified wholes, or gestalts, which makes them inherently meaningful. Briefly stated,

[2] While not venturing into the details here, it is worth noting that enactive approaches tend to stress the social nature of cognition [13], considering human consciousness as essentially intersubjective [32].

image schemas can be understood as describing the most elementary meaning structures of experience.

The theory of conceptual metaphor is deeply rooted in the larger context of embodied cognition. As Johnson [20, 21] has stressed, the idea is precisely to give an account of how our abstract thoughts are grounded and structured by concrete sensory-motor activity, thus evading the need to postulate different ontologies for bodily action and mind. Even though the theoretical agreement between the theory of conceptual metaphors and the enactive approach is somewhat debatable [13], we take them to be consistent enough for the aims of this paper. Therefore, following Johnson [21], image schemas can be understood as experiential structures arising from the structural couplings discussed above.

2.2 Music and Other Technologies

The enactive perspective stresses the active nature of experience, cognition and life itself. Hence, following Alva Noë [28], music and other media can be conceived of in terms of tools and technologies. While the idea of media as something that is actively used is hardly a novel one, a few of the implications of the enactive view on technological practices merit consideration.

As understood by Noë, technology doesn't only refer to objects but also to practices. The core meaning of technology can thus be framed, in line with the enactive approach, as "skillful activity" [28, p. 24]. Moreover, technology is not only something we work with but something that "works on us". In this sense, Noë defines technologies as "evolving patterns of organization" or "organized ways of doing things" [28, p. 19, 25]. This means that once adopted, technologies start to structure our actions and being. Matt Hayler [17] refers to similar idea by pointing out the domesticating nature of technology. Finally, an important issue concerning Noë's account on technology is that technologies are natural for human beings [28]. As such, our life is to a great degree governed by how we use technologies to extend our abilities and to give organization to our activities.

Perhaps the most crucial aspect in the enactive view on technology is that, through enactivism's non-dualistic stance, it aims to dispel the traditional distinction between the subject and object of technological interaction. The conception of technology as, at its core, skillful activity, suggests that there is no technology (as abstract entity) apart from someone using and experiencing it. Likewise, the idea of technology as organizing activity suggests that in use technologies become effectively part of the user.[3] While this characterization may risk confusing the concept of technology, it is taken as a fruitful premise for considering music, among other arts and media, as a form of technology [8, 28] or cognitive extension [23] more generally. In this sense, music can be considered as a way to extend our experiential capabilities by giving tools for structuring and creating experiences.

[3]The idea of interdependency between media and its user (that there is no game without a player nor player without a game) as such has been well acknowledged within research on video games and music, but the exact interpretation of this coupling is not unambiguous. However, there have been some empirical studies suggesting that tools may indeed become integrated as a part of the cognitive system [9].

3 IMMERSION AS A METAPHOR FOR EXPERIENCE OF INVOLVEMENT

As a general concept, immersion has been defined as "absorption in some condition, action, interest, etc." [29].[4] In line with this definition immersion has been used to describe experiences of deep mental involvement with different media. Nevertheless, in recent discussions, immersion has been especially linked to the relatively recent technologies of virtual reality and video games. Within this context, it has gained a more specific meaning as an experience of "being in a different reality" [27], which highlights the metaphor of transportation alongside that of absorption [2].

Besides this basic idea there has been no unanimous agreement on the exact meaning or definition of immersion. This stems partly from conceptual ambiguity; especially the concept of presence has been used as a synonym for immersion, or presented alongside as referring to the psychological aspect of "being there", whereas immersion is understood as referring to the technical side (see [2, 15, 16]). Another kind of confusion is related to the use of immersion as the aforementioned general concept as opposed to a more specific definition, which is linked to the question of whether immersion should be limited to descriptions of experiences with ergodic media, thought to afford qualitatively different kind of experiences [2]. Moreover, the talk of immersion as a desirable quality in itself, especially in video-games marketing, has added a sort of "idealized" aura to it, thereby blurring its meaning by assuming immersion to account for more (e.g., being the main reason for the pleasure of media) than should be reasonable [30].

3.1 Gordon Calleja and Four Challenges of the Concept of Immersion

In his book "In-Game", Gordon Calleja [2] aims to clarify the concept of immersion. He sets the stage by framing four challenges that hinder the clear understanding of the concept. These four challenges, re-arranged and modified to suit the needs of this paper, are here taken as a starting point for the discussion of immersion and its use as a description of musical experiences.

The first issue is what Calleja calls the challenge of technological determinism. Briefly stated, the idea of technological determinism is that immersion is dependent on the quality of the technology (e.g., high quality sound and visuals, bigger screen etc.), which tends to posit the creation of immersion as a technological challenge of creating "more immersive" devices. While Calleja doesn't deny the importance of the features of the medium, he considers immersion from an experience-oriented perspective, which stresses the active role of the player in constructing the game as experienced.

Perhaps the most important aspect of technological determinism is the idea that immersion is related to the ability of technology to create the impression of realism or fidelity to the real world. The idea of considering immersion as straightforwardly dependent on the realism of the virtual world has been criticized on several grounds (see [2, 15, 30]), one of which is the question of how realism should be understood: Should the games employing "un-realistic"

[4]Besides the definition quoted above, the entry in the Oxford English online dictionary also gives a physical definition of immersion as "dipping or plunging into water or other liquid" [29], which hints at the other definition of "being in something".

points of view, fantasy elements or cartoonish graphics be considered inherently non-immersive?

While realism in terms of visual and auditory fidelity to the real world is indeed one factor, it could be argued that more important aspects for immersion would be contextual realism and the realism of interaction [15]. This means basically that the virtual world functions according to a similar logic to that of the real world, or more precisely, has a similar overreaching inner consistency and predictability as the real world. This is in line with the player oriented perspective, since it stresses the expertise of the player to construct a logic in the virtual world (what can, and should be done, and what kind of effects the action has in the world), which can be considered a prerequisite for immersion.

The second issue is the idea of immersion as a monolithic concept. This was already accounted for by Ermi and Mäyrä in their famous SCI model, which distinguishes between sensory, challenge-based and imaginative immersion [11]. In similar fashion, Calleja's "player involvement model" presents six dimensions of involvement: Alongside spatial, ludic and narrative involvement, which bear a certain resemblance to the SCI model, Calleja presents dimensions of kinesthetic, shared and affective involvement. Besides extending the list of possible ways of being immersed, Calleja also posits each of these dimensions as having a macro phase, referring to experiences outside the game, along with an in-game micro phase. The point is that immersion is not a single form of experience caused by a single factor in an on/off-manner, but can be derived from different constituents and varies in intensity.

While Ermi and Mäyrä's and Calleja's models agree on the multi-dimensional nature of immersion, the latter can be seen to be more specific in regards to video games and a particular conception of immersion. This leads to the third issue, which Calleja sums up as a confusion between the metaphors of immersion as transportation and immersion as absorption. While both can be seen as aspects of immersion, the failure to distinguish the two or the usage of immersion in a more general sense risks confusing the concept and jettisoning its more specific meaning. Calleja's solution to this is replacing the metaphor of immersion with a dual metaphor of incorporation, which refers to both player being incorporated into the gameworld and the gameworld being incorporated as part of the player's consciousness.

Lastly, the final issue concerns the specificity of immersion to a certain, namely ergodic, kind of media. The basic line of reasoning Calleja presents is that since immersion is a concept taken to describe experiences with certain kind of media, it should not be extended to cover media that afford different kinds of affordances for engagement. Specifically, in order for the medium to afford immersion, it has to acknowledge the player's presence in the virtual world and thus grant agency to the player.

Given the subject-matter of this paper is musical immersion, the last issue merits further discussion. While there are reasonable grounds to suppose that different media and technologies lend themselves to different kinds of experiences, it is not that clear whether the experience of immersion should be considered specific to certain media. Moreover, since enactivism suggests that all experiences are interactive and have a world-constituting quality, it can be questioned whether, or in what sense, "ergodicity" is the distinguishing feature of video games in the case of immersion.

3.2 Musical Immersion

While immersion has not been widely discussed within music studies, the concept has nevertheless been used and the experience of "being there" in itself is well acknowledged [12, 22]. As discussed above, for a medium to be considered as affording immersion, it should provide an environment of a sort and afford agency (or possibilities for interaction). Since listening to music doesn't afford either world or agency in the same sense as video games or virtual realities, in what sense would it afford immersion? This question is considered here in relation to three listening situations.

Beginning with sound in video games, it can be taken as somewhat obvious that as video game worlds are also sonic worlds, the audio plays a part in affording the overall gameplay immersion [4, 10, 15]. Game audio can also be considered to have more specific roles in immersion, so that it should not just be overhastily grouped into other dimensions of immersion. Karen Collins [4], for example, discusses game audio as extending the game into the real world and facilitating identification with the game avatar in a manner that the screen-bound visual aspect is incapable of. Likewise, Mark Grimshaw [15, 16] has argued for the importance of audio in creating the sense of spatiality and therefore sense of inhabiting a space. Considering immersion specifically within music in video games, Isabella van Elferen [10] has suggested a dimensional model, similar to those of Calleja as well as of Ermi and Mäyrä discussed above. Her model scrutinizes how music in video games affords immersion through affect, literacy (e.g., cultural cues familiar from cinema) and interaction. Audio can thus be considered an integral part of the game world and, since the player responds to and generates sounds, it can be taken affording interaction in the same sense as other elements of the game.

To continue with another obvious notion, "everyday reality" is sonic in the same way as virtual reality. Music is no different from other sounds in that it can also be used to create spaces. One may, for example, consider how commercial spaces (such as shops and cafeterias) use music to demarcate their space and create a certain mood, how one may use portable music devices to shut out the surrounding audial environment or how one may modify the atmosphere of one's surroundings by using music (see [8]). This "background music" is also interactive in the sense that it can be affected by changing our physical position in relation to the sound source. Moreover, interactivity of a sort is evident in the situations, where one can select and initiate the music from the music player.

In the enactive view, the effort required for interaction with music is not trivial but constitutive. In order to constitute the non-organizational stimuli into meaningful music, one has to make-sense of it through musical affordances. Since affordances are relational (i.e., what an object affords depends on the capabilities and needs of the subject) this dynamic interaction gives rise and modifies the music as experienced.

Sound in video games and music in everyday life presents two cases where music is a constitutive part of the environment affording immersion. However, one may argue that in these cases immersion is not ultimately that much about music rather than about contextual environment and activity: While music has a role in shaping the environment and affording immersion, it is actually the activity in a video game world, or the everyday life chores, into

which one gets immersed. With respect to this, the third listening situation discussed here concerns the more traditional idea in music studies of "pure" or "absolute" music.

Briefly stated, the idea of absolute music is that music in itself ought to be distinguished from the extra-musical elements (such as lyrics, cultural connotations or emotional references). In this view, the essence of music is in its purely formal structure. The enjoyment of music, therefore, is the aesthetic enjoyment of appreciating the composition as it unfolds, which requires understanding of the theoretical and formal properties of music.

The formalist position presents music in itself as affording a world into which one can immerse oneself by utilizing music-theoretical knowledge. However, as this view also posits music as an autonomous entity with a fixed structure, it is at odds with enactive views of continuity between bodily action and abstract thought and co-constitution of the subject and object. The enactive alternative to a formalist conception of the musical world is offered by Joel Krueger [22]. His premise is the idea referred to above, namely that as sounds in general carry spatial information, music is likewise experienced as spatially structured. While this spatiality is the way we perceive music as music, as spatially organized instead of random noise, we may also take a more sensitive and understanding listening attitude towards music, which Krueger labels "deep listening". In deep listening, one takes an active role towards musical space, exploring its inner sonic structure, thereby inhabiting the musical space. Unlike formalist appreciation of form, this sonic exploration rests not upon music-theoretical knowledge but upon our "everyday" bodily skills of interacting with musical affordances.

4 METAPHORS FOR INVOLVEMENT: IMMERSION, FLOW, INTERACTION AND INCORPORATION

The four metaphors were chosen mainly because they seem to describe different crucial aspects of musical involvement, yet the metaphors are often used interchangeably with each other. Therefore the aim was to consider their differences in order to clarify the metaphors and to gain a better sense of what they entail.

The method of investigating the metaphors alongside image schemas can be described as one of cross-fertilization. Since the metaphors chosen can be used in different ways and as image schemas may be realized in different fashions, neither could be taken as exactly fixed. Therefore the aim was to consider the metaphors and image schemas in a mutually constraining manner, hoping to arrive at generally acceptable as well as illuminating conclusions. It is acknowledged that different schemas could have been used and their logic interpreted otherwise.

Besides describing the metaphor and image schema it suggests, the investigation considered three aspects related to the logic of image schema (see [3]). The one most strictly related to the logic of image schema was the bipolarity or axiology inherent in the schema. Another aspect of the logic of image schemas is that they can be realized either in a static fashion as a state, or in a dynamic fashion as a process. Finally, as the basic idea was to consider music as a technology, the question of how agency is organized in experience according to the metaphors was discussed. The summary is presented in Table 1.

4.1 Immersion

Immersion, as an experience of "being there", is most closely aligned with the CONTAINER image schema and its in-out orientation. The CONTAINER schema refers to the pervasive experiences of physical containment as in being contained within our surroundings or placing an object within another [20]. As such, the metaphor of immersion frames music as a place in which one is contained.

The very basic logic of CONTAINER dictates a boundary separating inside and outside [20]. Immersion is thus about closing in and blocking out the world around. Containment typically involves an idea of protection from or resistance to external forces, which in this case may be considered in "escapist" fashion as whatever distractions, worries and mundanities of the everyday world. Basically, when immersed one is situated in the same place as elements of music, instead of looking at them from the outside,[5] but also shut out from the "real" world.

Like most of the image schemas, CONTAINER can be interpreted in either a static or dynamic fashion. While immersion is deeply connected with dynamic notions of action (e.g. being constituted by interaction or being itself a form of action or a dynamic experience of transportation [27]), it is by itself perhaps more easily understood as a state. Therefore one may consider immersion as a static state that is constituted and kept up by action, and affording action and experiences special to that state (as being in a place affords different actions than looking at it from the outside).

Finally, the formulation of immersion as a state of containment also has implications for the crucial notion of agency. As mentioned before, the CONTAINER image schema has a quality of restricting the actions within a container (for example, taking a deliberately reflective stance towards one's experience of music would be "stepping out" of the music therefore breaking the immersion) which also limits the agency of the subject. In this sense, immersion is about discarding one's own agency and subjecting it to the rules of the place.

At this point it would be worthwhile briefly taking note of immersion's relation to presence which, as noted before, is often used in relation to immersion and sometimes interchangeably with it. The basic line of reasoning in the present paper is that the metaphor of immersion is used to describe the experiential or psychological quality of "being inside something", thereby downplaying the technical connection (e.g., immersive technology) usually attributed to immersion in distinction to presence [2]. In this sense, it is mostly a terminological question whether one wishes to denote the experience in question with a metaphor of immersion or with a metaphor of presence. Nevertheless, should these two metaphors be demarcated, one could begin by considering presence as experiential "closeness to something" or "sharing a space with something" (through image schemas of, for example, CONTACT or NEAR-FAR). With respect to immersion, presence would therefore not be about being inside a bounded space but about being in contact with something. In this sense, the metaphor of presence in not describing an

[5] This "crossing of the boundary" can be taken as referring to the sense of nonmediation crucial to presence (see [25]).

Table 1: Image-schematic structures of metaphors of involvement

	Immersion	Flow	Interaction	Incorporation
Description	Being inside something	Moving along with something	Doing something with something	Taking something as a part of oneself
Image schema	CONTAINER	PATH	LINK	MERGING
Bipolarity	In - Out	Towards - Away	Connected - Disconnected	Ready to hand - Present at hand
Dynamic / Static	State	Process	Process	State
Agency	Restricted	Directed	Shared	Reorganized

experience of transporting oneself to "being there" but an experience of bringing something to "being here".

4.2 Flow

The concept of flow, as originally coined by Mihaly Csikszentmihalyi [5], describes "the holistic sensation that people feel when they act with total involvement" [5, p. 36]. This basic idea of losing oneself in pleasurable action can be further defined with notions of flow as autotelic or optimal experience [6]. These refer to experience being in itself rewarding and desirable, and to the optimal balance between skills and the challenge of the task.

As can be seen from the description above, flow is very closely related to immersion, and both terms are often used in connection to each other [5, 11, 15]. Most notably, both include an idea of deep absorption, which in the case of flow is present in the characteristics of intense and focused concentration, merging of action and awareness, loss of reflective self-consciousness and distortion of the temporal experience [6]. In order to differentiate the metaphors, the similarities will be downplayed and differences emphasized.

In contrast to immersion, the metaphor of flow is more about ongoing activity and motion (as in "being carried on by the flow" [6, p. 40]). Therefore the basic image schema would be PATH or the more elaborated SOURCE-PATH-GOAL. Arising from the concrete connectings of start and end points, the PATH image schema refers to moving from one point to another [20]. However, besides the idea of movement, flow also contains an idea of optimal balance between skill and challenge, which refers to the BALANCE image schema. These image schemas frame music as a journey which one travels and as a challenging task one is carrying out.

Perhaps the most notable quality of the PATH schema, in the case of flow, is the sense of a clear goal [6]. The goal may be fixed beforehand or during the action, or it can be more or less abstract in nature, but it is nevertheless something that structures the path one is travelling: One knows what one is aiming for, and can therefore stay on the path or construct the path with this goal in mind. Nevertheless, by virtue of a clear goal, there is also a certain aspect of automation and (perhaps paradoxically) a lack of creativity: While PATH need not be fixed beforehand, it is still a route that defines the possible actions and in a sense carries the traveller along it. As opposed to CONTAINER, which restricts the actions to those possible in a certain space, PATH restricts the actions with respect to the route towards the GOAL.

With regards to the static/dynamic dichotomy, there is a similar tension within the concept of flow as with the concept of immersion. While again deeply interwoven with action, flow is clearly described as a state or a zone resulting from the balance of skills and challenges [5, 6]. Similar tension is also present in the two image schemas assigned to flow above: While PATH and BALANCE could both be interpreted as either dynamic or static, the former tends to refer to the process of travelling whereas the latter tends to refer to the state of equilibrium.[6] Admitting a certain arbitrariness, it could nevertheless be argued that flow gives more weight to experience's dynamic qualities of "action following action" [5, p. 36] and is hence more easily understood as a process.

4.3 Interaction

The metaphor of interaction is here considered after Tia DeNora's [8] conception of human-music interaction. According to this view, music is understood as a technology that can be implicated in the work of constructing subjectivity.

The metaphor of interaction is open enough to allow for application of several schemas. As the basic idea can however be described as "action with something", one natural choice would be the image schema of LINK, which refers to experiences of different kinds of couplings and connections between two things. As such, the metaphor of interaction also entails several image schemas related to forces between the linked things [20]. With respect to the framing of music, as Lakoff and Johnson have pointed out, the word "with" has the feature of indicating both instrument and accompaniment [24]. Music is therefore framed as something that is used for something and as someone who is accompanying the interaction.

To begin with the dynamic-static dichotomy, interaction, as understood within DeNora's [8] music "in action" approach, is clearly an active metaphor. Therefore the LINK image schema should be understood either as a process of linking or as itself a part of the larger dynamic structure. These two conceptions can be illustrated by considering interaction in terms of extension and exchange.

The basic idea in doing something with something is to extend one's capabilities, which refers to the very basic definition of technology [17]. In this sense, linking with music makes it possible to use music for something that cannot be achieved without it [8]. By this "making possible" aspect, LINK is linked to the image schema of ENABLEMENT. Thus the LINK image schema is a part of a wider structure of "connecting in order to be able", which as a whole is

[6] Although, as Johnson notes, the bodily basis for the experience of balance is precisely in the activity of balancing [20]

a dynamic process. This description applies as well to objects and persons: After all, the ability of humans to survive has been from the very beginning and continues to be in a very concrete way dependent on linking with other humans [20].

The second idea refers to interaction as a LINK of exchange. This means that interaction is a process of giving and taking: The one being interacted with is not merely used but let to affect the action, its outcome and the other parts of the interaction. In this sense, interaction is about dynamic linking, which includes several different force-schemas referring to exchange between participants (e.g. COUNTERFORCE when intentions of the participants clash in one way or another). As in the case of extension, this description applies to both interaction with someone and something: While the first case is quite evident, the use of technology involves similar negotiation where the technology, by affording certain actions rather than others, shapes the interaction as well as the user (see [8]).

Finally, the notion of active agency is crucial for interaction. When considering interaction as extending, it is supposed that the one using something for something is indeed working as an active agent. Similarly, interaction in the sense of exchange presupposes that all participants are contributing to the interaction. Moreover, as opposed to the previous metaphors of immersion and flow, interaction is necessarily open-ended action. While there may be a predetermined and definite goal for interaction, as a process it is not strictly governed by the goal, since the linking of actors (and actants) leaves room for change in the course of the interaction and in the participants.

The open-endedness of reciprocal involvement can be seen as more or less the point of interaction. This can be exhibited with another metaphor, play, which emphasizes the often "non-purposive" nature of constituting musical experiences in interaction with music. Indeed, the embodied involvement in listening can be seen as imaginary playacting (see [13, 19, 33]). A special case of listening interaction, musical imagining (i.e., listening to music "inside one's head"), holds certain performative characteristics that could be considered playful even in the ergodic sense. Hence, open-endedness of the process of imagining music seems genuinely reciprocal, not only relating to structuring one's experience by playing with the musical material, but also structurally modifying the musical material itself (see [18]).

4.4 Incorporation

Incorporation bears a close resemblance to interaction, so the relationship between the two can be seen as a continuity. This can be illustrated via growing expertise in using a tool: When one has reached a certain level of expertise, the tool ceases to be a separate part in interaction and becomes an incorporated part of the user [17]. Therefore instead of linking, incorporation refers to image schema of MERGING. Music can thus be framed as a part of one's body, as in when speaking of music as a prosthetic technology [8].

Even though incorporation, like interaction, is fundamentally about action (what we do), it describes not so much a process as an established state. When interacting, the technology one interacts with is experienced as an extension of one's capabilities. Incorporation, in contrast, changes the perspective of one's capabilities in the way that the experience is not in the same sense about extension: One does not feel one's arm as an extension but rather as a part of oneself (see [7]). Likewise, since the object has become a familiar part of oneself, there is no exchange between the two.

Incorporation can be considered through Heideggerian (here following Hayler [17]) notions of present-at-hand and ready-to-hand. When encountered as present-at-hand, objects are considered in "scientific" attitude as observable. As ready-to-hand, this conscious consideration ceases and objects are brought into one's experience of oneself, therefore switching the focus from the object to the task being done with it. In a sense, incorporation is not so much about doing something with something than doing something through or as something.

An important aspect of incorporation is its modification of agency. As incorporation is about MERGING something into oneself, it is also an act of changing oneself. In enactive parlance, incorporation refers to a change in body-schema, therefore a change in sense-making capabilities. As noted by Helena De Preester and Manos Tsakiris [7], incorporation refers not to an extension but to internal reorganization of body-model. The crucial notion is therefore an experience of completeness: Losing of body-part generally leads to a feeling of being "not-whole" in a way losing a tool seldomly does.

Finally, the idea of incorporation as reorganization posits it as a continuous way of experiencing. This can refer to how technologies may become reliable enough as to be necessary for certain temporary but recurrent activities, like when relying on music in organizing activity during exercise or managing one's emotions (see [8, 23]). In a more comprehensive sense, by reorganizing our sense-making abilities incorporation of technology can mean a change in how we constitute our world as meaningful. With respect to music, one may consider, for example, how Noë describes music as a technology for reorganizing our rhythmic, melodic and tonal ways of being and acting in everyday life [28].

5 CONCLUSIONS

While the main focus of this paper was on the conceptual use of metaphors for experiences of musical involvement, it also suggests, due to the ontological basis of image schemas, a certain view of musical experience. Different metaphors can thus be considered as different ways of constructing music in the act of listening: the activity of structuring experience according to the different patterns of sensory-motor interaction the music affords.

In line with the dimensional models of immersion [2, 10, 11], the exact nature of these experiences is left as an open question. This means that, in the case of immersion, music may work as a container for "purely" musical elements as well as for (for example) emotions, memories or bodily choreographies. The important point in this case is that the experience is structured according to the bodily logic of containment: shutting out the world around and bringing whatever is inside into the presence of oneself.

What this view also highlights is the role of expertise in experience. While a certain skillfulness is usually considered a prerequisite for all of the experiences discussed here.[7], it can also be taken as

[7] Expertise is most clearly linked to flow and incorporation, but also immersion is often presented as requiring a certain level of familiarity with the media (at least the mastery of the controls) In the case of interaction, one may consider for example DeNora's use of the concept "aesthetic reflexivity" [8].

referring to the entire activity of experiencing. To take up again the idea of "technological determinism" discussed in relation to immersion, there is a tendency to consider media as "causing" certain experiences rather than "affording" them. The notion of expertise emphasizes instead the relational nature of affordance, framing experiences as skills and dispositions, as much dependent on personal aptitudes and cultural practices as they are on properties of the media and medium.

As a last remark concerning the view of experience presented here, it should be pointed out that through image schemas, the investigation of ways of experiencing is ultimately limited to a description of their very basic structure. As such, it leaves out much of the felt dimension of the experience (see [21]). With respect to the idealized status of especially immersion and flow, it can be questioned whether this kind of approach can account for what really is essential in these experiences. Despite the possible limitations, the conceptual analysis of this paper has pointed out how the metaphoric and image schematic qualities, attributed to words we choose to use to describe our involvement with sounds and music, also constitute our basic understanding of the phenomenon in the first place.

To conclude, we propose that the enactive approach alongside the image schematic analysis presented here provides explicit and ontologically enlightened ways of understanding the psychology of musical involvement. Besides taking part in theoretical discussions concerning musical experience and its conceptualization, we hope the present paper offers insights to those working with technology mediated sonic experiences in general. Especially, the perspective of image schemas may provide interaction designers with novel tools for describing and communicating the intended interaction experience when designing sound in games and media, or musical interfaces for new digital music instruments.

ACKNOWLEDGMENTS

This work is partly funded by Finnish Cultural Foundation (Central Finland Regional Fund) and Academy of Finland (Centre of Excellence in Game Culture Studies, project n:o 312397).

REFERENCES

[1] Espen J. Aarseth. 1997. *Cybertext: Perspectives on Ergodic Literature.* The John Hopkins University Press, London.
[2] Gordon Calleja. 2011. *In-Game: From Immersion to Incorporation.* The MIT Press, London.
[3] Alan Cienki. 1997. Some Properties and Groupings of Image Schemas. In *Lexical and Syntactical Constructions and the Construction of Meaning*, Marjolijn Verspoor, Kee Dong Lee, and Eve Sweetser (Eds.). John Benjamins, Amsterdam, 3–15.
[4] Karen Collins. 2013. *Playing with Sound: A Theory of Interacting with Sound and Music in Video Games.* The MIT Press, London.
[5] Mihaly Csikszentmihalyi. 1975. *Beyond Boredom and Anxiety.* Jossey-Bass Publishers, London.
[6] Mihaly Csikszentmihalyi. 1990. *Flow: The Psychology of Optimal Experience.* HarperCollins, London.
[7] Helena De Preester and Manos Tsakiris. 2009. Body-extension versus body-incorporation: Is there a need for a body-model? *Phenomenology and the Cognitive Sciences* 8, 3 (Sep 2009), 307–319. https://doi.org/10.1007/s11097-009-9121-y
[8] Tia DeNora. 2004. *Music in Everyday Life.* Cambridge University Press, Cambridge.
[9] Dobromir G. Dotov, Lin Nie, and Anthony Chemero. 2010. A Demonstration of the Transition from Ready-to-Hand to Unready-to-Hand. *PloS One* 5, 3 (03 2010), 1–9. https://doi.org/10.1371/journal.pone.0009433
[10] Isabella van Elferen. 2016. Analysing Game Musical Immersion: The ALI Model. In *Ludomusicology: Approaches to Video Game Music*, Michiel Kamp, Tim Summers, and Mark Sweeney (Eds.). Equinox Publishing, Sheffield, 32–52.
[11] Laura Ermi and Franz Mäyrä. 2005. Fundamental Components of the Gameplay Experience: Analysing Immersion. In *Changing Views: Worlds in Play, Selected Papers of the 2005 Digital Games Research Association's Second International Conference.*
[12] Alf Gabrielsson. 2011. *Strong Experiences With Music: Music is Much More Than Just Music.* Oxford University Press, Oxford.
[13] Shaun Gallagher. 2017. *Enactivist Interventions: Rethinking the Mind.* Oxford University Press, Oxford.
[14] James J. Gibson. 1986. *The Ecological Approach to Visual Perception.* Lawrence Erlbaum Associates, Hilsdale.
[15] Mark Grimshaw. 2011. Sound and Player Immersion in Digital Games. In *The Oxford Handbook of Sound Studies*, Trevor Pinch and Karin Bijsterveld (Eds.). Oxford University Press, Oxford, 347–366.
[16] Mark Grimshaw. 2017. Presence through sound. In *Body, Sound and Space in Music and Beyond: Multimodal Explorations*, Clemens Wöllner (Ed.). Routledge, London, 279–298.
[17] Matt Hayler. 2015. *Challenging the Phenomena of Technology.* Palgrave Macmillan, Hampshire.
[18] Erkki Huovinen and Kai Tuuri. in press. Pleasant Musical Imagery: Eliciting Cherished Music in the Second Person. *Music Perception* (in press).
[19] Martine Huvenne. 2013. Sound in film as inner movement: Towards embodied listening strategies. In *Moving Imagination: Explorations of gesture and inner movement*, Helena De Preester (Ed.). John Benjamins Publishing, Amsterdam, 133–148.
[20] Mark Johnson. 1990. *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason.* The University of Chicago Press, London.
[21] Mark Johnson. 2017. *Embodied Mind, Meaning, and Reason: How Our Bodies Give Rise to Understanding.* The University of Chicago Press, London.
[22] Joel Krueger. 2009. Enacting Musical Experience. *Journal of Consciousness Studies* 16, 2-3 (2009), 98–123.
[23] Joel Krueger. 2014. Affordances and the musically extended mind. *Frontiers in Psychology* 4 (2014), 1003. https://doi.org/10.3389/fpsyg.2013.01003
[24] George Lakoff and Mark Johnson. 2003. *Metaphors We Live By.* The University of Chicago Press, London.
[25] Matthew Lombard and Theresa Ditton. 1997. At the Heart of It All: The Concept of Presence. *Journal of Computer-Mediated Communication* 3, 2 (1997). https://doi.org/10.1111/j.1083-6101.1997.tb00072.x
[26] Humberto R. Maturana and Francisco J. Varela. 1998. *The Tree of Knowledge: The Biological Roots of Human Understanding.* Shambhala Publications, London.
[27] Janet H. Murray. 2016. *Hamlet on the Holodeck: The Future of Narrative in Cyberspace.* The Free Press, New York.
[28] Alva Noë. 2015. *Strange Tools: Art and Human Nature.* Hill and Wang, New York.
[29] OED Online. 2018. "Immersion, n.". Retrieved March 14, 2018 from www.oed.com/view/Entry/91885
[30] Katie Salen and Eric Zimmerman. 2004. *Rules of Play: Game Design Fundamentals.* The MIT Press, London.
[31] John Sloboda. 2005. *Exploring the musical mind: Cognition, emotion, ability, function.* Oxford University Press, Oxford.
[32] Evan Thompson. 2010. *Mind in Life: Biology, Phenomenology, and the Sciences of Mind.* The Belknap Press of Harvard University Press, London.
[33] Kai Tuuri and Henna-Riikka Peltola. in press. Building worlds together with sound and music: Imagination as an active engagement between ourselves. In *The Oxford Handbook of Sound and Imagination*, Mark Grimshaw, Mads Walter-Hansen, and Martin Knakkegaard (Eds.). Oxford University Press.
[34] Francisco J. Varela, Evan Thompson, and Eleanor Rosch. 1993. *The Embodied Mind: Cognitive Science and Human Experience.* The MIT Press, London.