

Mental Representations in Musical Processing and their Role in Action-Perception Loops

REBECCA S. SCHAEFER [1,2]
University of California, Santa Barbara

ABSTRACT: I address the diverging usage of the term “imagery” by delineating different types of imagery, each of which is supported by multimodal mental representations that are informed and modulated by the body and its in- and outputs, and that in turn modulate and inform perception and action through predictive processing. These multimodal representations, viewed here as mental models, underlie our individual perceptual experience of music, which is constructed in the listener as it is perceived and interpreted. While tracking incoming auditory information, mental representations of music unfold on multiple levels as we listen, from regularities detected across notes to the structure of entire pieces of music, generating predictions for different musical aspects. These predictions lead to specific percepts and behavioral outputs, illustrating a tight coupling of cognition, perception and action. This coupling and the prominence of predictive mechanisms in music processing are described in the context of the broader role of predictive processing in cognitive function, which is well suited to account for the role of mental models in musical perception and action. As a proxy for mental representations, investigating the cerebral correlates of constructive imagination may offer an experimentally tractable approach to clarifying how mental models of music are implemented in the brain.

Submitted 2014 March 31; accepted 2014 November 28.

KEYWORDS: *music imagination, predictive processing, embodied music cognition*

IF one thing has become abundantly clear from decades of research in music cognition, it is that musical experiences are highly personal. Although it is clear that a range of common (or predictable) responses to music can be identified, there is no assurance that two people will feel the same way about a piece of music, or even necessarily have the same experience at repeated listening. The internalized, automatic processes that drive these experiences are the focus of the current paper. When perceiving music, we can generally say that a waveform of sound – with no inherent meaning – is transformed into music once it is perceived and interpreted through processes driven by the unique listening biography of the listener. This may in turn trigger various memory associations, or lead to differences in directing attention to specific musical aspects that are most salient to that person. The way we make sense of music thus depends on how we relate what we hear to previous experience – models built up from exposure to a lifetime of music, created by the unique auditory environment in which an individual has existed. This previously encountered environment provides the statistical basis for what we expect from auditory signals around us, by tracking incoming information and placing it into our own unique context of what we have previously experienced, based on the incidence of occurrences of those experiences. In this way, the unique implicit representations we have of what is common (and thus expected) drive our unique response to music. It also follows that having a greater overlap in music exposure in your listening biography with other people, as for instance through growing up in the same place, or being born in the same year, will lead to more similar responses to the same stimuli than when fewer biographical details are shared.

In the following, I will argue that the implicit expectations that influence perception and action are driven by mental models, and that research into music imagery offers an informative look at these mental representations of music, as well as the mechanisms of active, engaged perception. I here use the term “mental model” to refer to an internal representation of a percept or action, as it is built up through experience and statistical learning. Furthermore, I pose that these representations are informed and modulated by the body and its in- and outputs (namely perception and actions), in line with how the brain

functions with its many efferent and afferent connections, thus making a strong case that mental representations have embodied components. The recently asserted theoretical framework of predictive processing as a driving mechanism in the coupling of perception, cognition and action, which also fits well with how we think about the predictions that drive music processing, includes an account of imagery as highly related to active perception, further supporting these ideas. Here, I will first discuss some of the work looking at music imagination in relation to mental models and then turn to findings of the coupling of perception, cognition and action, closing with more general considerations and implications.

IMAGERY AND PERCEPTION

Imagery has been a topic of investigation for a considerable time, initially from a philosophical perspective, and more recently within the fields of experimental psychology and cognitive neuroscience. A detailed overview of experimental findings of music imagery is given by Hubbard (2010), who concludes that auditory imagery preserves many structural and temporal features of the imagined stimulus, involves semantically interpreted information and expectancies, has depictive and descriptive components and is related to musical abilities and experience. Neuroscientific research has found imagery in various modalities (i.e. visual, auditory, or movement imagery) to overlap with modality-specific perception or action in the brain, in terms of anatomical locations as well as in the temporal signatures of the electrical brain activity. Although the precise involvement of different neural regions, and specifically the primary modality-specific brain regions, is not fully established (cf. Daselaar, Porat, Huijbers & Pennartz, 2010), overlap in modality-specific brain activity has been reported for visual (cf. Kosslyn, Ganis & Thompson 2001; Cichy, Heinzle, & Haynes, 2012), movement (Pfurtscheller, Brunner, Schlögl & Da Silva, 2006; Oosterhof, Tipper, & Downing, 2012), and music imagination (Halpern & Zatorre 1999; Herholz, Halpern, & Zatorre, 2012; Herholz, Lappe, Knief, & Pantev, 2008; Schaefer, Desain & Farquhar, 2013; Schaefer, Desain & Suppes, 2009; Schaefer, Vlek & Desain, 2011a).

In thinking about what this shared activation represents, the first priority is to establish agreement on what we mean by imagery, as this diverges considerably in the literature. While most imagery research involves the deliberate imagination of sounds (cf. Halpern 2001 and others), visual images (cf. Cichy *et al.*, 2012; Reddy, Tsuchiya, & Serre, 2010) or movements (cf. Oosterhof *et al.*, 2012; Pfurtscheller *et al.*, 2006), phenomena such as spontaneously imagined music (also termed “earworms”) (Bailes, 2007; Williamson *et al.*, 2011), comparisons of perception with a representation held in working memory (cf. Navarro Cebrian & Janata, 2010a, 2010b), manipulating elements of a musical image in working memory (Zatorre, Halpern, & Bouffard, 2009), judging the orientation of an object or body part (cf. Parsons, 1987, and many others, but see also Viswanathan, Fritz, & Grafton, 2012), implicit action outcome (cf. Keller, Dalla Bella, & Koch, 2010; Keller & Koch, 2008) or pitch expectations (Janata & Paroo, 2006), mental simulation of future scenarios (cf. Schacter, Addis, & Buckner, 2007) and various other tasks have all been considered imagery, or at least claimed to be supported by imagery processes. In some cases, the difference in task leads authors to question whether imagery really occurred (see for instance Hubbard’s (2010) review in which this potential problem is mentioned repeatedly for music imagery), whereas in several cases the use of imagery is simply inferred from the idea that some form of modality-specific representation is being used to perform a task, such as is the case for the hand-turning paradigms in which no deliberate imagery of movement is performed (cf. Vingerhoets, de Lange, Vandemaële, Deblaere & Achten, 2002). Although these processes surely have something in common, there are also clear differences, at least phenomenologically. When comparing and interpreting experimental outcomes, it is crucial to take the differences between these cognitive tasks into account, with implications for overlap or similarities with perceptual processes.

Types of Imagery

To clarify the different outcomes of studies of imagination and better distinguish types of imagery, it is useful to go beyond definitions utilized in psychology, and consider how philosophers have characterized functional differences in imagery types. A useful source of such thought is provided by Strawson’s essay “Perception and Imagination” (1974). Basing his ideas on the uses of the word “imagination” by Hume, Kant and Wittgenstein, Strawson distinguishes four types of imagery, further elaborated on and specified for music imagery by Moore (2010), who also provides specific terms for each. The four types that are thus

described are sensory imagery, creative imagery, propositional imagery and constructive imagery (Moore, 2010). Here, the first is functionally closest to how psychologists tend to approach imagery, namely as a mental (re-)creation of a sensory experience, while the second refers to the most common non-academically used meaning of “imagination”, or being imaginative or creative. The third type relates to false beliefs, or propositional attitudes (imagining that something is the case), and the final imagery type, most interesting to the current discussion, is described as “the faculty that organizes percepts for coherent cognition” (Moore, 2010). Although the term “constructive imagery” is used differently in other sources, there referring to the active construction of mental images, more fitting in the category of sensory imagery (cf. Van Leeuwen, 2013), I will here use this term as Moore has characterized the type of imagery that Strawson delineates. This type of imagery is driven by implicit internal representations, and instrumental in interpreting that which is perceived, which is very much in line with the way in which implicit imagery processes involved in action outcome predictions or judging body part orientations might work. Although a thorough discussion of different possible types of imagery and their relation to perception is considerably larger than the scope of this paper, I argue here that specifically this notion of constructive imagery is useful in thinking about the commonalities between imagery and perception, and possibly accounts very well for certain instances where imagery is described but does not include effortful, deliberate (re-)creation of an internal perceptual experience.

Some authors have put forward the idea that music perception often (if not always) includes an imagery-like component (cf. Hargreaves, 2012) and this process of perceptual organization, based on person-specific knowledge and experience, also neatly accounts for the highly personal interpretation of a (musical) stimulus, referred to above. The delineation of imagery types also allows for the interesting possibility that this constructive process may *also* be present in deliberate, effortful sensory imagery, in some sense closing the gap with other uses of the term “constructive imagery”, but also offering an account for the commonalities and shared cerebral processing between perception and sensory imagery (cf. Schaefer, 2011). The main argument for this idea would be that the same perceptual coherence is present as we internally recreate a sensory experience as when the stimulus is actually heard, fitting with Hubbard’s (2010) conclusion of mostly intact structural and temporal features of auditory images. A distinction that is useful to draw here is that sensory imagery is a deliberate task that requires some effort or concentration [3], whereas constructive imagery happens automatically, as a part of making sense of incoming information. Through constructive imagery, the implicit influence of personal experiences, but also selective attention can be accounted for. A good example of this is what happens when, for instance in a crowded room, only fragments of a piece of music can be heard, leaving the listener to (re-)construct the full stimulus internally, and the missing auditory information is supplied by what we think is likely to fit with those parts of the stimulus we can perceive (similar to perceptual completion of a visual object that is partly occluded) or, in the case of a well-known stimulus, the representation we already have. In this situation, we actively involve our previous experience to reconstruct what the stimulus may have been, with an obvious dependence on musical training or listening biography. This process of interpreting incoming information based on previous experience is also consistent with a statistical learning-based account of music perception mentioned above, which is reviewed and described in more detail by Perruchet and Pacton (2006).

We generally consider attentive music listening to be an active process. As a part of this process, we automatically track incoming information on multiple structural levels, such as harmonic progressions, tempo changes, relative dynamics and other aspects of the musical stimulus. On each structural level, our previous experience constructs our expectations, whether they are in time (like the particular time point where a beat is expected to happen) or in pitch (like a chord or note that may logically follow), providing structural layers that may be independent, but also may be interrelated. If we consider how imagery may be part of a perceptual experience (cf. Hargreaves, 2012), this tracking behavior is an obvious candidate, as our expectations direct our attention and ultimately our interpretation of the music. Listener-specific and situation-specific factors may also mediate the allocation and direction of attention, leading to different ways in which we can engage with a musical stimulus. Listener-specific factors include cultural background, music training, specific experiences, but also the listener’s mood, while situation-specific factors would be the listening location and social situation or activity (such as a concert, an aerobics class or listening to the radio while driving). Indeed, the production rules for emotional responses to music as formulated by Scherer and Zentner (2001) state as much, distinguishing music-structural features, performance features (together creating the stimulus we perceive), listener features and contextual features as combining to produce a specific response to music. Arguably, the constructive imagery (or perceptual

organization) that takes place while listening is captured in the listener features through one's unique, personal interpretation, influencing a range of responses, including aesthetic appreciation and emotion, but also behavior. An example of the latter is that when moving to music, or even just tapping a finger to a metronome, mental models of the rhythm or periodicity of the stimulus allow us to align our movements with the sound, as suggested by the body of experimental work on finger tapping, showing anticipatory processes to be active through slightly early taps (cf. Repp, 2005; Repp & Su, 2013). Here, movements are driven by predictions based on perceptual input in combination with a mental or cognitively created model of how this percept is likely to continue, while recalibrating the motor output based on action outcomes and perceptual inputs (i.e. being early or late with a tap). This process is an excellent example of how mental models, in this case a temporal structure, informed by perceptual input (a metronome) and action feedback from tapping, allow us to accomplish a task, such as entraining movements to sound. These tasks can vary from incredibly easy and automatic to highly complex and dependent on long-term training. An example of the latter is joint music making, which arguably needs a very tight coupling of perception, cognition and action as musicians adjust their movements to feedback from their own as well as other musicians' instruments, while tracking multi-layered music and executing very fine motor control, often preattentively, indicating intricately linked processes.

COUPLING OF PERCEPTION, COGNITION AND ACTION

The idea that perception, cognition and action are intricately coupled, and as such should not be seen (or investigated) as independent functional systems, is well supported by decades of research in music cognition, even more so than in general psychology. Bidirectional modulating influences have been demonstrated experimentally between perception and cognition, between perception and action and between cognition and action. Of these, the coupling between perception and cognition is perhaps the most extensively researched and needs the least description, with as obvious examples the cognitive models that have been described for tonal or temporal structures (cf. Deutsch & Feroe, 1981; Krumhansl, 1979; Palmer & Krumhansl, 1990) but of course also including effects of attention on perception, the effects of perceptual conditions on memory and so on.

Action and perception have also been shown to bidirectionally influence each other; we know that movements are regularized when carried out to music (Leman *et al.*, 2013; Thaut, Kenyon, Hurt, McIntosh, & Hoemberg, 2002), that exercising to music decreases perceived exertion (Bood, Nijssen, van der Kamp, & Roerdink, 2013; Karageorghis & Priest, 2012) and that action is preattentively modulated by what we perceive (Furuya & Soechting, 2010). Currently, the use of auditory cues in rehabilitation of disordered movement is an active area of research (cf. Schaefer, 2014a) that holds considerable promise for therapies based on the coupling between perception and action. Going in the other direction, movement is reported to enhance temporal acuity in auditory perception (Manning & Schutz, 2013; Su & Pöppel, 2012), and to modulate the perception of rhythmically ambiguous temporal patterns (Phillips-Silver & Trainor, 2005, 2008). Previous experience appears to be crucial to this coupling, such that in highly trained musicians, the association between perception and action is sufficiently strong that just listening to music activates specific motor regions in the brain that are related to how the music would be performed (cf. Bangert *et al.*, 2006; Haueisen & Knosche, 2001).

Finally, findings are now also emerging of cognition and action directly impacting each other. Examples of cognition impacting movement are that mental imagery helps music performance and practice (Holmes, 2005; Parncutt, 2007), both auditory and movement imagination (or mental training) aid motor learning (Brown & Palmer, 2013), and that "cognitive" states can be translated into actions, such as emotional expression being recognizable in dance (Burger, Saarikallio, Luck, Thompson, & Toiviainen, 2013). The influence of cognition on action has also been demonstrated at the level of movement-related brain activations; moving to imagined music was found to lead to different activation patterns within the motor network of the brain than when identical movements were carried out without music imagery (Schaefer, Morcom, Roberts & Overy, 2014). Probably the least investigated effect direction is that of movement impacting cognition, but various findings have been reported, specifically in investigations addressing specific aspects of the theoretical framework of embodied cognition (for a description of different aspects, see Wilson, 2002). For instance, the way we move influences emotional appraisal of music (Maes & Leman, 2013), and outside the realm of music research the effects of exercise on cognition are gaining considerable interest, including general effects of physical fitness (Angevaren, Aufdemkampe, Verhaar, Aleman, & Vanhees, 2008; Eggermont, Milberg, Lipsitz, Scherder, & Leveille, 2009), as well as

findings that exercising while learning new words is more effective than learning in a passive condition (Schmidt-Kassow, Kulka, Gunter, Rothermich, & Kotz, 2010).

On closer inspection of these findings we may find that, although in these studies the impact of one functional domain (as in cognition, action and perception) on another was investigated, a role of the third is often inherent to the mechanism under investigation, and processes that have been considered imagery by some authors (although probably not the effortful, sensory kind) are arguably at play in cases focusing on action-perception coupling. For instance, in the case of rhythmic movement impacting later perception and preference, where babies bounced to differing interpretations of an ambiguous rhythm then showed a preference for that interpretation (Phillips-Silver & Trainor, 2005), it is likely that the movement created an internal accenting structure that led to a stable mental model of one interpretation of the ambiguous rhythm, something that would not have happened without some form of a mental representation of that accent structure, leading to a specific percept. A spontaneously emerging subjective accenting structure, taking the shape of an imagined pattern in time, has been shown to be present in perception-related brain signatures even when actual accents were never presented (Brochard, Abecasis, Potter, Ragot & Drake, 2003), but can also be added to ambiguous stimuli in a deliberate manner (Schaefer *et al.*, 2011a). Mental representations of time structures have been argued to be inherent to both actual and imagined actions, similarly to perception and sensory imagery of music (Schaefer, 2014b), possibly even sharing neural resources. Likewise, in the anticipatory processing inherent to tapping along with a rhythm as mentioned earlier, the mental models of the rhythm appear to be directly informed by afferent motor feedback from the finger, as is suggested by the finding in a small sample of deafferented patients, who do not receive haptic or proprioceptive feedback from their movements and did not show the commonly seen slightly early taps, but instead were tapping right on time (Stenneken, Prinz, Cole, Paillard, & Aschersleben, 2006). These results imply that the motor system is crucially involved in setting up mental representations of temporal structure, highlighting the multimodal nature of the representations of rhythm that unfold over time.

Predictive Processing

The idea that predictive processes are a driving force in perception, cognition and action, which has in recent years gained traction in more general cognitive science and neuroscience (cf. Bar, 2009; Bubic, von Cramon, & Schubotz, 2010) is not a new idea within the field of music cognition. In the latter, the importance of predictive mechanisms in music listening, especially in terms of affective processing, has been stressed in a number of influential theories of music perception (cf. Meyer, 1956; Narmour, 1990; Huron, 2006). The impact of statistical learning on perceptual mechanisms has been experimentally evaluated for a range of perceptual phenomena, such as perceptual segmentation (Pearce, Ruiz, Kapasi, Wiggins & Bhattacharya, 2010), learning a new tonal system (Loui, Wessel & Hudson Kam, 2010) or even the longer-term effects of one's mother tongue, the rhythmic properties of which are related to composition and perception of musical rhythms (Iversen, Patel, & Ohgushi, 2008; Patel & Daniele, 2003). Our knowledge of statistical learning is sufficiently advanced that it also offers a way of investigating specific perceptual disorders such as congenital amusia (Omigie, Pearce, & Stewart, 2012). Additionally, various aspects of music perception can be accounted for using Bayesian models (e.g. Temperley, 2007), even bringing together apparently diverging datasets from rhythm perception and production tasks (Sadakata, Desain & Honing, 2006), showing that rhythm perception data can be predicted accurately from rhythm production data using Bayesian modeling. Although these examples are by no means an exhaustive review of the studies relevant to the coupling of perception, cognition and action in musical experiences or predictive mechanisms in music processing, they support the long-standing argument that conceptually, predictions and predictive processing are of substantial importance in understanding various aspects of musical behavior. Additionally, the findings illustrate how exceptionally suitable music is as a topic of research when investigating these mechanisms, and how musical paradigms represent a range of methodological opportunities for investigations into predictive processing.

Based on predictive coding theory initially developed in the context of visual perception (cf. Rao & Ballard, 1999), a more general theoretical framework for the mechanism of predictive processing in perception, cognition and action is described in Clark (2013). Here, Clark asserts that "Brains [...] are essentially prediction machines. They [...] support perception and action by constantly attempting to match incoming sensory inputs with top-down expectations or predictions. [...] Such accounts offer a unifying model of perception and action [and] illuminate the functional role of attention [...]." (p. 1) Thus, the

predictive coding account of cognition is that prior knowledge is used to guide sensory interpretations and action decisions, fitting exquisitely with the previously discussed coupling between cognition, perception and action. The mismatch between the predictions based on prior knowledge and the inputs from the senses and the body is termed the prediction error, which in turn informs new mental models. Clark here distinguishes surprisal, the subconscious violation of implicit predictions based on statistically generated models of the world, from surprise, or the conscious violation of explicit predictions about the world. When considering this framework specifically for music processing, this distinction can neatly capture the different kinds of prediction violations that have been described as crucial for music processing. While some musical phenomena are clearly and overtly surprising, such as a sudden change in time signature or an unlikely chord change (with the level of conscious surprise likely depending on the listener's musical expertise), the more subtle prediction violations, such as rhythmic syncopation in music composition and micro-timing deviations or pitch bending in expressive performance, are not consciously processed in the listener. This is also a good description of how, after many repeated listening (presumably removing any remaining conscious surprises for the listener, although perhaps adding some heightened anticipation for specific musical events), the affective response to music can still be based on the deviations from what is predictable in the context of the statistical properties of our auditory environment. Moreover, it can also account for the preferences we can develop for a specific performance of a musical piece, often leading to repeated listening without losing its attraction even after hearing it many times, and even leading to an aversive response when hearing a different performance. Indeed, the high incidence of repetition of musical patterns within a single piece or musical section is a characteristic element of music and has been put forward as being crucial to eliciting emotions (Margulis, 2013), suggesting that the exposure needed to build expectations (which in turn elicit responses) is already packaged within a single piece of music at a range of structural levels.

A number of papers have already discussed the more specific implications of a predictive coding account of perception and cognition for music processing. Data supporting these ideas have been presented for pitch perception (Kumar *et al.*, 2011), processing rhythmic incongruities (Vuust, Ostergaard, Johanne, Bailey, & Roepstorff, 2009) but also emotional responses to music (Gebauer, Kringelbach, & Vuust, 2012; Vuust & Frith 2008). The strength of the predictive processing framework as presented by Clark (2013) is that it supports the existence of a multitude of levels or dimensions on which hierarchical predictions are formed simultaneously. In this way, the implementation of the predictive mechanisms are tailored to that specific musical aspect, whether the predictions are about musical style, rhythmic periodicity or tempo, specific scale notes or micro-timing. For each of these levels, the impact of the prediction error may be different, with in some cases a certain amount of prediction error leading to a more positive aesthetic response than precisely accurate predictions, such as for instance in naturally timed versus mechanical percussion (Schaefer, Overy & Nelson, 2013). This complexity of real-world music is identified as a problem for more traditional computational models of prediction (Rohrmeier & Koelsch, 2012). For more basic auditory processing, not restricted to, but conceptually including musical materials, Winkler *et al.* (2009) put forward an account of auditory perception based on predictive processing, which was further elaborated to a predictive coding view in a later paper (Winkler, Denham, Mill, Bohm, & Bendixen, 2012) and supported by a range of experimental findings (Bendixen, SanMiguel, & Schröger, 2012; Chennu *et al.*, 2013; Winkler & Czigler, 2012). Indeed, clinical findings of auditory hallucinations or tinnitus, often appearing in the absence of perceptual stimulation, have been interpreted to be related to the predictive processing in the auditory cortex, creating the impression of external stimulation without sound (Kumar *et al.*, 2014; SanMiguel, Widmann, Bendixen, Trujillo-Barreto, & Schröger, 2013; Vanneste, Song, & De Ridder, 2013). Specifically the temporal structure in sound has been shown to be highly conducive to predictive mechanisms (Holm & Madison, 2013; Schwartz, Tavano, Schröger, & Kotz, 2012; Tavano, Widmann, Bendixen, Trujillo-Barreto, & Schröger, 2014), supporting the type of higher-order processing that is necessary for musical rhythm perception with its multiple hierarchical levels of structure.

The use of prior knowledge, collected from an individual's perceptual senses and action outcomes, in interpreting sensory input, also fits with the concept of using mental models or constructive imagery in active perception as described in the previous section. In a different paper, Clark (2012) states that "Prediction-driven [...] [perceptual] systems, simply as part and parcel of their ability to perceive, must develop an ability to self-generate perception-like states." He goes on to say that "...converging paradigms [...] suggest that perception is inextricably tied up not just with action but also with something functionally akin to imagination. They suggest, moreover, that prediction-driven processing [...] plays a major role in the development and deployment of perceptual understanding, and that self-generated action enables and

enhances such learning in a variety of important ways.” (pp. 760-1) This description of the process by which perceptual input is interpreted or organized, including the involvement of the motor system, implies that something like imagery, and in this case fitting the description of constructive imagery in the shape of mental models of that which is perceived, would be a necessary component of successful perception. The implication is that listening with some previous experience of that sound, be it a rhythm, a specific instrument, a specific genre and so on, the model we have of that periodicity, timbre or style will be included in our listening experience, and this representation may have salient aspects in a range of modalities. Based on experimental data of event-related brain potentials, Winkler and Czigler (2012) come to a similar conclusion and argue that perceptual object representations are closely related to the generative models assumed by predictive coding theories. In this way, the theoretical framework of predictive processing also potentially ties together the neuroscientific findings of overlap in the brain between active or attentive perception and active sensory imagery, although there are some indications that this overlap may also be mediated by the complexity of the stimulus (Schaefer *et al.*, 2013).

DISCUSSION

In the above, I have aimed to bring together ideas and findings from experimental psychology, cognitive neuroscience and philosophy to clarify the role of imagery in perception and action. The main argument is built on the assertion that what is referred to as imagery in the literature actually covers a range of phenomena, sometimes leading to conflicting findings and confusion in their interpretation, although in all cases including some sort of internal representation. These phenomena of imagery may occur with or without awareness, and with varying degrees of deliberateness and effort. Clarification of terminology and delineation of different types of imagery may help here, and subtypes of imagery are described.

The most relevant types for the current discussion are sensory imagery and constructive imagery, of which I here suggest that the latter may be included in the former, as a vividly imagined stimulus comprises structural and temporal features that need to be perceptually organized. That is, if indeed the structural and temporal features of an imagined stimulus are preserved (cf. Hubbard, 2010), the relations between pitches, durations, and even sections of a musical piece are retained in the image, including the unconscious predictions created for the different hierarchical levels of the stimulus, which can all be considered a part of perceptual organization. The concept of constructive imagery thus not only offers an account of the overlap between sensory imagery and active perception that is reported in the neuroimaging literature, but also fits well with a theoretical framework of predictive processing as a unifying mechanism in perception, action and cognition. The generative models that are assumed in this framework are informed by previous perception as well as action, leading to representations that impact our interpretation of what we perceive, which also include action-related aspects. The tight coupling of perception, action and cognition implies that the representations we have are multimodal, in that one set or level of predictions may be how we cognitively process an incoming stimulus (such as judging the tempo of a rhythm), whereas a set of predictions concern the behavioral response (such as tapping your foot) is generated concurrently, but not independently. Although this account is not the only one in which cognition, perception and action are thought of as tightly coupled (see for instance Hurley, 2008), the role of prediction error, which is so salient in music processing, emphasizes the importance of internal models in this coupling. Moreover, the multi-leveled structure of music may be exquisitely suited to further investigate more detailed questions about how these models and predictions may be generated, and whether there are specific types of models and predictions that are more strongly interrelated (such as perhaps that expectations about where in time harmonic changes happen are related to expectations of where in time a downbeat will occur), while other types may exist quite independently (such as the expectation of where a downbeat at a certain time point happening independently of which instrumental timbres are expected).

With the current contribution I aim to address an aspect of the personalized experience of music, based on internal mental models that incorporate characteristics related to both perception and action, and are generated from previous experience. By considering these internal mental models, which organize perception and direct our attention to specific stimulus features, as a type of imagery, as some authors already do, tension between research findings utilizing very different experimental tasks can be resolved. Although fully explaining individual differences in musical experiences is beyond the aim of the current perspective, the core idea that an individual combination of previous experience contributes to basic mechanisms of perception and action through construction of mental models that drive predictions that are crucial to perceptual processes is indeed relevant to these differences. Secondly, I argue for the relevance of

researching music imagery with the goal of understanding how these mental models drive our musical experiences, while making clear distinctions in what aspect of imagery is researched, and what can be concluded from the results. For instance, by clarifying the impact of previous exposure on implicit, and memory functions on explicit, deliberate imagery tasks [4], gaining a better understanding of individual differences in sensory imagery ability, and the role of imagery in expert skills, a lot may be learned about the role of automatic as well as deliberate mental simulation in cognition, perception and action.

In the context of the embodied music cognition, this account may still sound quite cognition-centered, however from a neurologist's perspective, the concept of the "non-neural body" or, similarly, the disembodied brain makes very little sense, given the rich bidirectional connections within the central nervous system as well as between the central and peripheral nervous system. Our mental representations, traditionally designated as "cognitive", clearly contain aspects of perception and action, most obviously present in the neural signatures of mental simulation, and likely in many other cognitive operations, especially in the case of music processing, which inherently engages multiple senses and bodily responses, from tears and shivers (cf. Sloboda, 1991) to induced movement (Janata *et al.*, 2012). The predictive processing ideas, while possibly uniting apparent contradictions within the concepts of embodied cognitive science, concerning what the specific computational role of the body is in cognition (Kiverstein, 2012), implicate something like multimodal imagination to be at play continuously during music processing (and perception more generally), impacting the interpretation of perceptual signals as well as the performance of movements. This closely resembles how imagery has been considered from the perspective of music perception research as well as performance science. Imagery skills tend to be thought of as a mark of musicianship (and thus appear in several historical tests of musical aptitude, see for instance Gordon, 1965; Seashore, 1915), although this more often relates to skills in deliberate sensory imagery rather than interpretative perceptual skills. Of course the two are likely related (to each other, and to musical training, see Pecenka & Keller, 2009), but are arguably phenomenologically different processes. The constructive imagery that is part of tracking and predicting one's environment is dependent on successful model construction through experience, whereas deliberate imagery, while making use of models or representations, presumably also requires attention and skill. This argues for experimental investigations that isolate task or modality effects, i.e. effects related to moving, perceiving, or imagining in a specific modality, from stimulus effects, based on the content of the action, percept or image (see for instance Daselaar *et al.*, 2010; Schaefer, Vlek & Desain, 2011b).

The use of deliberate imagery in music has been described in the context of a number of goals, mostly focusing on mental rehearsal, but also the use of metaphors in acquiring technique or attaining expressive performance, and even imagery related to attenuating performance anxiety and stress (Gregg, Clark & Hall, 2008). The fact that these goals are all related to motor output and include sensory representations further underscores the aforementioned point on the multimodal nature of these representations. These deliberate types of imagery, and specifically mental rehearsal, are likely very closely related to musical expertise, which is a reason for many studies of sensory imagery to use trained musicians as participants, and may also be true for different domains of expertise (see Schaefer, 2014b for more discussion). Mental rehearsal is reported as an effective aid in learning music, athletics, dance and used in movement rehabilitation and brain-computer interfaces (Gerven *et al.*, 2009; Owen *et al.*, 2006), but is also dependent on sensory imagery skills (Clark & Williamon, 2011), which has been identified as a problem for its application in rehabilitation (Malouin & Richards, 2010). Somewhat conflicting results can be found in the literature for other modalities of imagination, such as for instance athletes (Ozel, Larue, & Molinaro, 2004) and dancers (Malouin *et al.*, 2009) reporting more vivid deliberate movement imagery than non-experts, while dancers counter-intuitively show slower performance in mental rotation (Jola & Mast, 2005), the latter finding interpreted as due to a difference in strategy. This interpretation is supported by findings illustrating that for some aspects of spatial abilities, such as mental rotation, a range of strategies may be used, that may also impact task ability (Logie, Pernet, Buonocore, & Della Sala, 2011). Again, the difference between active sensory or motor imagery and constructive processes such as mental rotations and manipulations appear to dissociate (as already reported by Farah, Weisberg, Monheit, & Peronnet, 1989), supporting the use of conceptual subtypes such as sensory and constructive imagination.

The role of mental representations in action-perception loops, such as the coordination of behavioral output to an external stimulus, is thus mainly one involving the generation and evaluation of predictions. The constructive imagery that is a part of active perception is still a somewhat elusive phenomenon that in some cases happens completely subconsciously, but in other cases emerges spontaneously into our awareness. Examples are the spontaneous but conscious imagery experiences

involved in auditory anticipation (Leaver *et al.*, 2009) or gap filling (Kraemer, Macrae, Green, & Kelley, 2005), indicating that the distinctions between imagery types are not always clear. Hargreaves (2012) makes a case for viewing imagery as a central aspect of music perception, formulating a three-way reciprocal feedback model of music processing in which the agent, be it a listener, performer, composer or improviser interacts with the musical material on one hand and the situation or context on the other (which are also bidirectionally connected), with imagination forming the core of the model, impacting both perception and production and being credited with something like musical executive function, crucial to musical creativity. For piano performance, deliberate music imagery has recently been described as a great aid in practice and memorization, but also offering performers a way to reduce cognitive load while performing, once a detailed representation of a piece, including auditory, visual and bodily aspects, is present (Davidson-Kelly, 2014).

While the terminology might still benefit from further specification, this way of representing the role of imagery in music processing aligns closely with the view described here, with an explicit focus on the personalized nature of musical experiences. Researching this, and other types of imagery, will likely offer more clarification of imagery subtypes, the extent to which the subtypes may overlap or co-occur, and the nature of the representations that facilitate difference types of imagery. Given the practical uses of mental practice, a better understanding of the skill set that is involved in this type of imagination would have direct practical implications. Further investigation of the individual differences in constructive imagery, potentially related to other musical sub-skills (cf. Pecenka & Keller, 2011), appears integral to a better understanding of musical experiences. Finally, given the great potential for music to facilitate shared experience (cf. Overy, 2012) the concept of shared mental models, be it through a commonality in previous exposure or the current listening situation, may offer a theoretical framework for potential investigations of rich interpersonal experiences in musical settings.

CONCLUSION

To summarize, the divergent use of the term ‘imagery’ in scientific literature groups tasks together that can not reasonably be considered to be functionally the same, as the level to which they are effortful or automatic, as well as the extent to which they are consciously experienced, vary greatly. By specifying our terms more clearly, more progress can be made towards an understanding of the mechanism of imagery phenomena, and their role in other cognitive, perceptual and bodily processes. Importantly, the representations that underlie imagery phenomena should be recognized to be multimodal, treating the brain and the body as one system. A theoretical account that could drive future research is that of predictive processing, which has already been established as highly suited to musical behavior. The concept that mental models underlie perception but also various types of imagery provides a potential explanation for the shared neural activation between deliberate as well as automatic imagery tasks with actual perception or action.

NOTES

[1] All correspondence can be addressed to rebecca.schaefer@sagecenter.ucsb.edu. This work is supported by the European Commission under the Marie Curie Intra-European Fellowship Programme (grant FP7-2010-PEOPLE-IEF 276529), and the SAGE Center for the Study of the Mind at the University of California, Santa Barbara.

[2] This paper is based on a talk presented at the Embodied Music Cognition (EMuCog) Conference, Edinburgh, UK, in July 2013.

[3] Spontaneous sensory musical imagery, also termed an ‘earworm’ or ‘involuntary musical imagery’ is arguably not effortful (as this is a spontaneous phenomenon), however in terms of characteristics and experience still more akin to sensory imagery than constructive imagery, which happens subconsciously.

[4] Although slightly outside the scope of the current paper, it is important to comment on the close link between imagery and memory functions. Not only is imagery considered an important component of vivid memory, it is arguably used during working memory rehearsal, and imagery has long been considered a

mechanism of improving associative memory. Conversely, vivid sensory imagination of an existing piece is highly dependent on long-term memory, with better knowledge of a stimulus (or action) supporting the ease of imagery. In this context it is useful to bear in mind that memory functions and imagery functions are likely intertwined and cannot be viewed separately. Arguably, the (multimodal) representations that are encoded in memory have much in common (or overlap) with the representations discussed here, however this deserves its own discussion elsewhere.

REFERENCES

Angevaren, M., Aufdemkampe, G., Verhaar, H. J. J., Aleman, A., & Vanhees, L. (2008). Physical activity and enhanced fitness to improve cognitive function in older people without known cognitive impairment. *Cochrane Database of Systematic Reviews*, 3, CD005381.

Bailes, F. (2007). The prevalence and nature of imagined music in the everyday lives of music students. *Psychology of Music*, 35, 555–570.

Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., Heinze, H.-J. & Altenmüller, E. (2006). Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction. *NeuroImage*, 30, 917 – 926.

Bar, M. (2009). Predictions: a universal principle in the operation of the human brain. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 364, 1181–1182.

Bendixen, A., SanMiguel, I., & Schröger, E. (2012). Early electrophysiological indicators for predictive processing in audition: A review. *International Journal of Psychophysiology*, 83(2), 120–31.

Bood, R. J., Nijssen, M., van der Kamp, J., & Roerdink, M. (2013). The power of auditory-motor synchronization in sports: Enhancing running performance by coupling cadence with the right beats. *PLoS One*, 8(8), e70758.

Brochard, R., Abecasis, D., Potter, D., Ragot, R., & Drake, C. (2003). The “ticktock” of our internal clock: Direct brain evidence of subjective accents in isochronous sequences. *Psychological Science* 14(4), 362–366.

Brown, R. M., & Palmer, C. (2013). Auditory and motor imagery modulate learning in music performance. *Frontiers in Human Neuroscience*, 7, 320.

Bubic, A., von Cramon, D. Y., & Schubotz, R. I. (2010). Prediction, cognition and the brain. *Frontiers in Human Neuroscience*, 4, 25.

Burger, B., Saarikallio, S., Luck, G., Thompson, M. R., & Toiviainen, P. (2013). Relationships between perceived emotions in music and music-induced movement. *Music Perception*, 30(5), 517–533.

Chennu, S., Noreika, V., Gueorguiev, D., Blenkmann, A., Kochen, S., Ibáñez, A., Owen, A. & Bekinschtein, T. A. (2013). Expectation and attention in hierarchical auditory prediction. *The Journal of Neuroscience*, 33(27), 11194–205.

Cichy, R. M., Heinze, J., & Haynes, J. (2012). Imagery and perception share cortical representations of content and location. *Cerebral Cortex*, 22(2), 372-380.

Clark, A. (2012). Dreaming the whole cat: Generative models, predictive processing, and the enactivist conception of perceptual experience. *Mind*, 121(438), 753-771.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181-204.

- Clark, T., & Williamon, A. (2011). Imagining the music: Methods for assessing musical imagery ability. *Psychology of Music, 40*(4), 471–493.
- Daselaar, S. M., Porat, Y., Huijbers, W., & Pennartz, C. M. A. (2010). Modality-specific and modality-independent components of the human imagery system. *NeuroImage, 52*, 677–685.
- Davidson-Kelly, K., (2014). *Mental Imagery Rehearsal Strategies for Expert Pianists* (Unpublished Doctoral Dissertation). University of Edinburgh, Edinburgh, UK.
- Deutsch, D., & Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological Review, 88*(6), 503–522.
- Eggermont, L. H. P., Milberg, W. P., Lipsitz, L. A., Scherder, E. J. A., & Leveille, S. G. (2009). Physical activity and executive function in aging: The MOBILIZE Boston Study. *Journal of the American Geriatrics Society, 57*(10), 1750–6.
- Farah, M. J., Weisberg, L. L., Monheit, M., & Peronnet, F. (1989). Brain activity underlying mental imagery: Event-related potentials during mental image generation. *Journal of Cognitive Neuroscience, 1*(4), 302–316.
- Furuya, S., & Soechting, J. F. (2010). Role of auditory feedback in the control of successive keystrokes during piano playing. *Experimental Brain Research, 204*(2), 223–37.
- Gebauer, L., Kringsbach, M. L., & Vuust, P. (2012). Ever-changing cycles of musical pleasure: The role of dopamine and anticipation. *Psychomusicology: Music, Mind & Brain, 22*(2), 152–167.
- Gerven, M. V., Farquhar, J., Schaefer, R. S., Vlek, R. J., Geuze, J., Nijholt, A., Ramsey, N., Haselager, P., Vuurpijl, L., Gielen, S. & Desain, P. (2009). The brain-computer interface cycle. *Journal of Neural Engineering, 6*(4), 41001.
- Gordon, E. (1965). *Musical Aptitude Profile*. Boston: Houghton Mifflin
- Gregg, M. J., Clark, T. W. & Hall, C. R. (2008). Seeing the sound: An exploration of the use of mental imagery by classical musicians. *Musicae Scientiae, 12*(2), 231-247.
- Halpern, A. R. (2001). Cerebral substrates of music imagery. *Annals of the New York Academy of Sciences, 930*, 179–192.
- Halpern, A. R. & Zatorre, R. J. (1999). When that tune runs through your head: A PET investigation of auditory imagery of familiar melodies. *Cerebral Cortex, 9*, 697–704.
- Hargreaves, D. J. (2012). Musical imagination: Perception and production, beauty and creativity. *Psychology of Music, 40*(5), 539–557.
- Haueisen, J., & Knosche, T. R. (2001). Involuntary motor activity in pianists evoked by music perception. *Journal of Cognitive Neuroscience, 13*, 786–792.
- Herholz, S. C., Halpern, A. R., & Zatorre, R. J. (2012). Neuronal correlates of perception, imagery, and memory for familiar tunes. *Journal of Cognitive Neuroscience, 24*(6), 1382–1397.
- Herholz, S. C., Lappe, C., Knief, A., & Pantev, C. (2008). Neural basis of music imagery and the effect of musical expertise. *European Journal of Neuroscience, 28*, 2352–2360.

Holm, L., & Madison, G. (2013). Whenever next: Hierarchical timing of perception and action. *Behavioral and Brain Sciences*, 36(3), 37–38.

Holmes, P. (2005). Imagination in practice: a study of the integrated roles of interpretation, imagery and technique in the learning and memorisation processes of two experienced solo performers. *British Journal of Music Education*, 22(3), 217–235.

Hubbard, T.L. (2010). Auditory imagery: Empirical findings. *Psychological Bulletin* 136(2), 302-329.

Hurley, S. (2008). The shared circuits model (SCM): How control, mirroring, and simulation can enable imitation, deliberation, and mindreading. *Behavioral and Brain Sciences*, 31, 1-58

Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT Press.

Iversen, J. R., Patel, A. D., & Ohgushi, K. (2008). Perception of rhythmic grouping depends on auditory experience. *Journal of the Acoustical Society of America*, 124(4), 2263–2271.

Janata, P., & Paroo, K. (2006). Acuity of auditory images in pitch and time. *Perception & Psychophysics*, 68, 829–844.

Janata, P., Tomic, S. T. & Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology: General*, 141(1), 54-75

Jola, C., & Mast, F. W. (2005). Mental object rotation and egocentric body transformation: Two dissociable processes? *Spatial Cognition and Computation*, 5(2&3), 217–237.

Karageorghis, C. I., Priest, D. L. (2012). Music in the exercise domain: A review and synthesis (Part I). *International Review of Sport and Exercise Psychology*, 5, 44-66

Keller, P. E., & Koch, I. (2008). Action planning in sequential skills: relations to music performance. *Quarterly Journal of Experimental Psychology*, 61(2), 275–91.

Keller, P. E., Dalla Bella, S., & Koch, I. (2010). Auditory imagery shapes movement timing and kinematics: evidence from a musical task. *Journal of Experimental Psychology. Human Perception and Performance*, 36(2), 508–13.

Kiverstein, J. (2012). The meaning of embodiment. *Topics in Cognitive Science*, 4, 740–758.

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience*, 2, 635–642.

Kraemer, D. J. M., Macrae, C. N., Green, A. E., & Kelley, W. M. (2005). Sound of silence activates auditory cortex. *Nature*, 434, 158.

Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, 11, 346–374.

Kumar, S., Sedley, W., Barnes, G. R., Teki, S., Friston, K. J., & Griffiths, T. D. (2014). A brain basis for musical hallucinations. *Cortex*, 52, 86-97.

Kumar, S., Sedley, W., Nourski, K. V, Kawasaki, H., Oya, H., Patterson, R. D., Howard III, M. A., Friston, K. J. & Griffiths, T. D. (2011). Predictive coding and pitch processing in the auditory cortex. *Journal of Cognitive Neuroscience*, 23(10), 3084–94.

- Leaver, A. M., Van Lare, J., Zielinski, B., Halpern, A. R., & Rauschecker, J. P. (2009). Brain activation during anticipation of sound sequences. *Journal of Neuroscience*, *29*, 2477–2485.
- Leman, M., Moelants, D., Varewyck, M., Styns, F., van Noorden, L., & Martens, J.-P. (2013). Activating and relaxing music entrains the speed of beat synchronized walking. *PLoS One*, *8*(7), e67932.
- Logie, R. H., Pernet, C. R., Buonocore, A., & Della Sala, S. (2011). Low and high imagers activate networks differentially in mental rotation. *Neuropsychologia*, *49*(11), 3071–3077.
- Loui, P., Wessel, D. & Hudson Kam, C. L. (2010). Humans rapidly learn grammatical structure in a new musical scale. *Music Perception*, *27*, 377–88.
- Maes, P.-J., & Leman, M. (2013). The influence of body movements on children’s perception of music with an ambiguous expressive character. *PLoS One*, *8*(1), e54682.
- Malouin, F., & Richards, C. L. (2010). Mental practice for relearning locomotor skills. *Physical Therapy*, *90*(2), 240–251.
- Malouin, F., Richards, C. L., Durand, A., Descent, M., Poiré, D., Frémont, P., Pelet, S., Gresset, J. & Doyon, J. (2009). Effects of practice, visual loss, limb amputation, and disuse on motor imagery vividness. *Neurorehabilitation and Neural Repair*, *23*(5), 449–63.
- Manning, F., & Schutz, M. (2013). “Moving to the beat” improves timing perception. *Psychonomic Bulletin & Review*, *20*(6), 1133–9.
- Margulis, E. H. (2013). Repetition and emotive communication in music versus speech. *Frontiers in Psychology*, *4*, 167.
- Meyer, L. B. (1956) *Emotion and meaning in music*. Chicago: University of Chicago Press.
- Moore, M. E. (2010). *Imagination and the mind’s ear*. (Unpublished doctoral dissertation). Temple University, Philadelphia, PA.
- Narmour, E. (1990) *The analysis and cognition of basic melodic structures: The implication-realization model*. Chicago: University of Chicago Press.
- Navarro Cebrian, A., & Janata, P. (2010a). Electrophysiological correlates of accurate mental image formation in auditory perception and imagery tasks. *Brain Research*, *1342*, 39–54.
- Navarro Cebrian, A., & Janata, P. (2010b). Influences of multiple memory systems on auditory mental. *Journal of the Acoustical Society of America*, *127*(5), 3189–3202.
- Omigie, D., Pearce, M. T., & Stewart, L. (2012). Tracking of pitch probabilities in congenital amusia. *Neuropsychologia*, *50*(7), 1483–1493.
- Oosterhof, N. N., Tipper, S. P., & Downing, P. E. (2012). Visuo-motor imagery of specific manual actions: a multi-variate pattern analysis fMRI study. *NeuroImage*, *63*(1), 262–71.
- Overy, K. (2012). Making music in a group: synchronization and shared experience. *Annals of the New York Academy of Sciences*, *1252*, 65–8.
- Owen, A. M., Coleman, M. R., Boly, M., Davis, M. H., Laureys, S. & Pickard, J. D. (2006). Detecting awareness in the vegetative state. *Science*, *313*, 1402.

- Ozel, S., Larue, J., & Molinaro, C. (2004). Relation between sport and spatial imagery: Comparison of three groups of participants. *The Journal of Psychology*, *138*(1), 49–63.
- Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology - Human Perception and Performance*, *16*(4), 728–741.
- Parncutt, R. (2007). Can researchers help artists? Music performance research for music students. *Music Performance Research*, *1*(1), 13–50.
- Parsons, L. M. (1987). Imagined spatial transformations of one's hands and feet. *Cognitive Psychology*, *19*, 178–241.
- Patel, A. D., & Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition* *87*, B35-B45.
- Pearce, M. T., Ruiz, M. H., Kapasi, S., Wiggins, G. A. & Bhattacharya, J. (2010). Unsupervised statistical learning underpins computational, behavioral, and neural manifestations of musical expectation. *NeuroImage*, *50*, 302–313.
- Pecenka, N. & Keller, P. E. (2011) The role of temporal prediction abilities in interpersonal sensorimotor synchronization. *Experimental Brain Research*, *211*, 505–15.
- Pecenka, N. & Keller, P. E. (2009). Auditory pitch imagery and its relationship to musical synchronization. *Annals of the New York Academy of Sciences*, *1169*, 282–286.
- Perruchet, P. & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*(5), 233–8.
- Pfurtscheller, G., Brunner, C., Schlögl, A. & Lopes da Silva, F. H. L. (2006). Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. *NeuroImage*, *31*(1), 153–159.
- Phillips-Silver, J. & Trainor, L. J. (2005). Feeling the beat: movement influences infant rhythm perception. *Science*, *308*, 1430.
- Phillips-Silver, J. & Trainor, L. J. (2008). Vestibular influence on auditory metrical interpretation. *Brain and Cognition*, *67*(1), 94–102.
- Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience* *2*(1), 79-87
- Reddy, L., Tsuchiya, N., & Serre, T. (2010). Reading the mind's eye: Decoding category information during mental imagery. *NeuroImage*, *50*(2), 818–825.
- Repp, B. H. (2005). Sensorimotor synchronization: a review of the tapping literature. *Psychonomic Bulletin & Review*, *12*, 969–92.
- Repp, B. H. & Su, Y.-H. (2013). Sensorimotor synchronization: a review of recent research (2006-2012). *Psychonomic Bulletin & Review*, *20*, 403–52.
- Rohrmeier, M. A. & Koelsch, S. (2012). Predictive information processing in music cognition. A critical review. *International Journal of Psychophysiology*, *83*(2), 164–75.
- Sadakata, M., Desain, P. & Honing, H. (2006) The Bayesian way to relate rhythm perception and production. *Music Perception*, *23*, 269–88.

SanMiguel, I., Widmann, A., Bendixen, A., Trujillo-Barreto, N., & Schröger, E. (2013). Hearing silences: Human auditory processing relies on preactivation of sound-specific brain activity patterns. *The Journal of Neuroscience*, 33(20), 8633–9.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: The prospective brain. *Nature Reviews Neuroscience*, 8(9), 657–61.

Schaefer, R. S. (2014a). Auditory rhythmic cueing in movement rehabilitation: Findings and possible mechanisms. *Philosophical Transactions of the Royal Society B* 369: 20130402.

Schaefer, R. S. (2014b). Images of time: Temporal aspects of auditory and movement imagination. *Frontiers in Psychology: Perception Science* 5:877.

Schaefer, R. S. (2011). *Measuring the mind's ear: EEG of music imagery*. Radboud University Nijmegen, Nijmegen, The Netherlands.

Schaefer, R. S., Desain, P. & Farquhar, J. (2013). Shared processing of perception and imagery of music in decomposed EEG. *NeuroImage*, 70, 317-326.

Schaefer, R. S., Desain, P. & Suppes, P. (2009). Structural decomposition of EEG signatures of melodic processing. *Biological Psychology*, 82, 253-259.

Schaefer, R. S., Farquhar, J. & Desain, P. (2011). Name that tune: Decoding music from the listening brain. *NeuroImage*, 56, 843-849.

Schaefer, R. S., Morcom, A.M., Roberts, N. & Overy, K. (2014). Moving to music: Effects of heard and imagined musical cues on movement-related brain activity. *Frontiers in Human Neuroscience*, 8: 774.

Schaefer, R. S., Overy, K. & Nelson, P. (2013). Affect and non-uniform characteristics of predictive processing in musical behaviour. *Behavioral and Brain Sciences* 36(3), 226-227.

Schaefer, R. S., Vlek, R.J. & Desain, P. (2011a). Decomposing rhythm processing: Electroencephalography of perceived and self-imposed rhythmic patterns. *Psychological Research*, 75(2), 95-106.

Schaefer, R. S., Vlek, R.J. & Desain, P. (2011b). Music perception and imagery in EEG: Alpha band effects of task and stimulus. *International Journal for Psychophysiology*, 82(3), 254-259.

Scherer, K. R. & Zentner, M. R. (2001). Emotional effects of music: Production rules. In Juslin, P.N. & Sloboda, J.A. (Eds.) *Music and emotion: Theory and research*. Oxford: Oxford University Press.

Schmidt-Kassow, M., Kulka, A., Gunter, T. C., Rothermich, K. & Kotz, S. A. (2010). Exercising during learning improves vocabulary acquisition: Behavioral and ERP evidence. *Neuroscience Letters*, 482(1), 40–4.

Schwartz, M., Tavano, A., Schröger, E. & Kotz, S. A. (2012). Temporal aspects of prediction in audition: Cortical and subcortical neural mechanisms. *International Journal of Psychophysiology*, 83(2), 200–7.

Seashore, C. E. (1915). The measurement of musical talent. *The Musical Quarterly*, 1(1), 129-148.

Sloboda, J. A. (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music*, 19, 110-120

Stenneken, P., Prinz, W., Cole, J., Paillard, J. & Aschersleben, G. (2006). The effect of sensory feedback on the timing of movements: Evidence from deafferented patients. *Brain Research*, 1084(1), 123–31.

- Strawson, P. F. (1974). *Imagination and Perception*. In: *Freedom and Resentment*. London: Methuen.
- Su, Y.-H. & Pöppel, E. (2012). Body movement enhances the extraction of temporal structures in auditory sequences. *Psychological Research*, 76(3), 373–82.
- Tavano, A., Widmann, A., Bendixen, A., Trujillo-Barreto, N. & Schröger, E. (2014). Temporal regularity facilitates higher-order sensory predictions in fast auditory sequences. *The European Journal of Neuroscience*, 39(2), 308-318.
- Temperley, D. (2007). *Music and probability*. Cambridge, MA: MIT Press.
- Thaut, M. H., Kenyon, G. P., Hurt, C. P., McIntosh, G. C. & Hoemberg, V. (2002). Kinematic optimization of spatiotemporal patterns in paretic arm training with stroke patients. *Neuropsychologia*, 40(7), 1073–81.
- Van Leeuwen, N. (2013). The meanings of “Imagine” part I: Constructive imagination. *Philosophy Compass*, 8(3), 220-230.
- Vanneste, S., Song, J.-J. & De Ridder, D. (2013). Tinnitus and musical hallucinosis: The same but more. *NeuroImage*, 82, 373–83.
- Vingerhoets, G., de Lange, F. P., Vandemaele, P., Deblaere, K. & Achten, E. (2002). Motor imagery in mental rotation: An fMRI study. *NeuroImage*, 17, 1623-1633.
- Viswanathan, S., Fritz, C. & Grafton, S. T. (2012). Telling the right hand from the left hand: Multisensory integration, not motor imagery, solves the problem. *Psychological Science*, 23(6), 598–607.
- Vuust P. & Frith, C. D. (2008). Anticipation is the key to understanding music and the effects of music on emotion. *Behavioral and Brain Sciences*, 31, 599–600.
- Vuust, P., Ostergaard, L., Johanne, K., Bailey, C., & Roepstorff, A. (2009). Predictive coding of music – Brain responses to rhythmic incongruity. *Cortex*, 45(1), 80–92.
- Williamson, V. J., Jilka, S. R., Fry, J., Finkel, S., Mullensiefen, D., & Stewart, L. (2011). How do “earworms” start? Classifying the everyday circumstances of Involuntary Musical Imagery. *Psychology of Music*, 40(3), 259-284.
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4), 625-636.
- Winkler, I., & Czigler, I. (2012). Evidence from auditory and visual event-related potential (ERP) studies of deviance detection (MMN and vMMN) linking predictive coding theories and perceptual object representations. *International Journal of Psychophysiology*, 83(2), 132–43.
- Winkler, I., Denham, S. L., & Nelken, I. (2009). Modeling the auditory scene: Predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences*, 13(12), 532–540.
- Winkler, I., Denham, S., Mill, R., Bohm, T. M., & Bendixen, A. (2012). Multistability in auditory stream segregation: A predictive coding view. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 367(1591), 1001–12.
- Zatorre, R. J., Halpern, A. R., & Bouffard, M. (2009). Mental reversal of imagined melodies: A role for the posterior parietal cortex. *Journal of Cognitive Neuroscience*, 22(4), 775–789.