A Review of Music-Evoked Visual Mental Imagery: Conceptual Issues, Relation to Emotion, and Functional Outcome

Liila Taruffi[1] and Mats B. Küssner[2]

[1]Freie Universität Berlin, Berlin, Germany

[2]Humboldt-Universität zu Berlin, Berlin, Germany

Abstract

Visual mental imagery has been characterized as an important aspect of our mental life, which consists of "seeing" in the absence of a sensory stimulus. However, the mechanisms underlying how visual mental images unfold during music listening have remained largely neglected. Here, we review the existing literature on the relation between music-evoked emotions and images and we draw attention on how visual mental imagery has been previously conceptualized in the music domain. We also propose to adopt a conceptual framework from research on spontaneous cognition, which will promote a more nuanced and comprehensive understanding of the different types of music-evoked visual mental imagery. Finally, we highlight how music's capability to trigger images can be harnessed in daily life as well as therapeutic practices to foster the benefits and minimize the costs of visual mental imagery.

Keywords: visual mental imagery, music-evoked emotions, internally-oriented cognition

Introduction

What is visual mental imagery?

In contrast to vision, which occurs due to retinal stimulation, visual mental imagery arises when perceptual information is accessed not from an immediate visual stimulus but from memory, thereby allowing the individual "to see with the mind's eye" (Kosslyn, Ganis, & Thompson, 2001). However, visual mental imagery does not always necessarily overlap with episodic memory (the recall of previously perceived objects or events), and can also draw from the modification of previous perceptual information in novel ways. In cognitive psychology and neuroscience, visual mental imagery is traditionally understood as a specific type under the general domain of imagery abilities, which encompasses auditory, gustatory, olfactory, haptic as well as motor senses (Halpern, 2001; Kosslyn et al., 2001; Pearson, 2007; Plailly, Delon-Martin, & Royet, 2012; Schaefer, 2014). In this sense, it is postulated that multimodal mental imagery may even be the rule rather than the exception. As Nanay (2018) has argued, multimodal imagery is a pivotal element of our everyday perception. In this review article, we focus on visual mental imagery because the experiences of mental imagery triggered by music listening are either predominantly visual or a combination of visual and other mental imagery modalities (Küssner & Eerola, same special issue). It should be noted, however, that other types of mental imagery, such as motor imagery, play an important role in various musical contexts. Musical sounds usually afford strong images of motion in the listener—a process that is relevant also for music-evoked emotions (Overy & Molnar-Szakacs, 2009). Several authors have argued that we perceive (or imagine) motion in music through the use of conceptual metaphors, which, by definition, are grounded in our bodily experiences (Cox, 2016; Eitan & Granot, 2006; Godøy, 2003; Johnson & Larson, 2003). Experientially, it can be hard to separate motor from visual imagery. For instance,

professional musicians—as well as performers from other fields—regularly engage in kinaesthetic and visual motor imagery when learning a new piece (Bernardi, De Buglio, Trimarchi, Chielli, & Bricolo, 2013; Di Rienzo et al., 2016; Lotze & Halsband, 2006; Schuster et al., 2011); thus, they may feel and see their bodily movements on stage and use this imagery as an additional dimension of aptitude in performance preparations. Different imagery abilities across modalities are highly correlated (for example, auditory pitch imagery contributes to sensorimotor synchronization in ensemble musicians; Pecenka & Keller, 2009) and support a wide range of functions including memory, emotion and mood, movement or action, and different types of learning (D'Argembeau & Van der Linden, 2006; Holmes & Mathews, 2005; Jeannerod & Decety, 1995; Tartaglia, Bamert, Mast, & Herzog, 2009). In particular, the ability to visually represent the world in our minds has immense value in its applications across other aspects of life, allowing us, for example, to navigate and predict the complex social environments in which we live, to travel backwards in time and remember our past experiences, or to picture ourselves in the future (D'Argembeau & Van der Linden, 2006). Moreover, visual mental imagery assists in human beings' urge to create, which is often channeled through art (many writers imagine fictional characters and/or scenarios to facilitate creative writing). Further applications extend to educational settings, where it was shown that guided imagery practice stimulates creative writing in the classroom (Jampole, Mathews, & Konopak, 1994). Besides its role in the arts and creativity (Pearson, 2007), visual mental imagery is also used in both clinical and informal practices to promote mental health and well being (e.g., Holmes, Arntz, & Smucker, 2007).

In philosophy, understanding imagery has historically been a core issue (e.g., Kant, 1965; Wittgenstein, 1953), and has been especially relevant in conceptions of the mind (Tye, 2000). Despite its established inextricable relevance, 20th-century philosophers often overlooked visual mental imagery. For example, in 1913 Behaviorist John B. Watson denied

that mental images even existed, mainly due to their covert nature, which evaded empirical assessment. However, during the late 1970s and early 1980s a lively debate arose amongst philosophers and cognitive psychologists surrounding the nature of visual mental imagery (Sterelny, 1986). The two main theoretical positions fell into "Pictorialism" and "Descriptionalism", respectively. Advocates of "Pictorialism" claimed that mental images have a picture-like structure, and thereby serve to depict, not to describe objects (e.g., Kosslyn, 1980), whereas advocates of "Descriptionalism" held that representations underlying the experience of mental imagery are the same type as those used in language (e.g., Pylyshyn, 1973). In recent decades neuroimaging literature has provided evidence of a large overlap between visual perception and visual imagery. For example, a number of studies suggest that the early visual cortex crucially supports the construction of visual mental images (for a review see Kosslyn & Thompson, 2003), and this correspondence between neural resources involved in imagery and perception largely applies to the other sensory modalities as well (Kosslyn et al., 2001). Nevertheless, neuroimaging studies have revealed differences between the neural correlates of imagery and perception, with frontal areas involved in memory processes being more engaged during imagery than during perception and primary sensory areas showing stronger activity during perception than during imagery, likely due to their respective bases in recall and recombination of memories and sensory processes (Albers, Kok, Toni, Dijkerman, & de Lange, 2013; Ishai, Ungerleider, & Haxby, 2000; Schaefer, Desain, & Farquhar, 2013).

Visual mental imagery and music

While visual mental imagery has been studied extensively across several disciplines including philosophy, cognitive neuroscience, and visual arts (e.g., Kosslyn, 1980; Nanay, 2015), visual mental imagery in music is still largely underexplored and has considerable

potential to advance our understanding of the role and function of visual mental imagery in humans (see, for example, the KOSMOS Workshop on "Mind Wandering and Visual Mental Imagery in Music" that took place during May 2018 at Humboldt University in Berlin). In relation to music, visual mental imagery refers to the mechanism whereby music stimulates internal images in the listener consisting of pictorial representations (e.g., natural landscape, colors), embodied image-schemata (e.g., picturing a melodic movement as an ascending or descending image), or complex visual narratives (e.g., similar to that of a movie). In music psychology research, visual mental imagery has been discussed predominantly as a trigger for affective responses following the theoretical framework proposed by Juslin and Västfjäll (2008). In this landmark paper, the authors proposed six mechanisms—*brain stem reflex*, *evaluative conditioning*, *emotional contagion*, *visual imagery*, *episodic memory*, and *musical expectancy*—through which music listening evokes emotions and moods. Additional mechanisms, such as *rhythmic entrainment* and *aesthetic judgment*, were included more recently (Juslin, 2013). Several of these mechanisms have received much attention by music psychologists and thus warranted further theoretical discussion and empirical testing. Examples include *episodic memory*, where emotion induction is caused by music evoking the memory of a particular event in the listener's life, or *musical expectancy*, where emotion induction is caused by a specific feature of the music that violates, delays, or confirms the listener's expectations about the unfolding of the music (Cuddy, Sikka, Silveira, Bai, & Vanstone, 2017; Gurr, Foxhall, Shinoda, & Baird, 2014; Huron, 2006; Janata, Tomic, & Rakowski, 2007; Janata, 2009; Lehne, Rohrmeier, & Koelsch, 2014; Meyer, 1956). However, other mechanisms including *visual imagery* have been largely neglected by music psychology research despite their crucial role in music-evoked emotions (Day & Thompson, same special issue; Vuoskoski & Eerola, 2015).

Despite the sparseness of research conducted on the topic, evidence relating to vivid visual mental imagery evoked by music has been provided both anecdotally and by findings in music psychology. Anecdotal evidence is long-standing, dating back to Albert Einstein in 1929 mentioning that he "liv[ed] his daydreams in music" (Saturday Evening Post, 26 October 1929, p. 17). Likewise, in a recent study, Küssner and Eerola (same special issue) reported that 77% of music listeners have experienced visual mental imagery during music listening. Similar estimates were provided by Vuoskoski and Eerola (2015), who found that around 80% of their participants (48 out of 60 individuals) reported visual mental imagery after listening to evocative, instrumental film music combined with narrative descriptions regarding the context of the music pieces. However, these estimates refer to a retrospective survey method (Küssner & Eerola, same special issue) and a single music trial (Vuoskoski & Eerola, 2015), leading to obvious limitations in the generalization of these findings (for more details see the *Outstanding Issues and Future Directions* section). Although Vuoskoski and Eerola (2015) limited the effects of demand characteristics by avoiding specific instructions related to imagery, their findings still require further corroboration. While investigating the form of spontaneous thoughts (or *mind-wandering*; see next section) evoked by music, Taruffi and colleagues (2017) observed that spontaneous thoughts occur more significantly in the form of images compared with words during both sad and happy music, with large effect sizes for both sad and happy music ($d = 0.67$ and $d = 0.83$, respectively). The above findings demonstrate the prevalence and importance of visual mental imagery during music listening. Given the ubiquity of music (in an experience sampling study, 64% of 141 participants reported to listen to music "several times a day"; Juslin & Laukka, 2004) and the pervasiveness of visual mental processes during music listening (Küssner & Eerola, same special issue; Taruffi, Pehrs, Skouras, & Koelsch, 2017; Vuoskoski & Eerola, 2015), it is surprising that there has been no attempt to systematically characterize visual mental images

evoked by music. Several outstanding issues remain unexplained and include but are not limited to the following: What types of images are evoked by music (e.g., colors, abstract shapes, etc.)? Is visual mental imagery shaped by the cultural background of the listener? Are the music's formal properties associated with specific features of the images? Does emotion function as a mediating mechanism between music and images? The precise nature of visual imagery processes during music listening episodes, their structure, phenomenology, function, and underlying brain mechanisms still remain to be determined.

Conceptual and Definitional Challenges of Music-Evoked Visual Mental Imagery

Visual mental imagery as a modality of mind-wandering, musical daydreams, and autobiographical memories

Visual mental imagery while listening to music may fall under a broader family of mental experiences; however, consensus on its appropriate terminology has yet to be determined. Examples of previously articulated terms encompass: *mind-wandering* or *daydreaming*, *musical daydreams,* and *music-evoked autobiographical memories*. The variety of terms mirrors a fuzzy conceptual framework that can not clearly account for this type of mental experiences, and this conceptual ambiguity may be partly inherited from challenges in the conceptualization of internally-oriented cognitive phenomena such as mind-wandering (Christoff, Irving, Fox, Spreng, & Andrews-Hanna, 2016; Seli et al., 2018).

Visual mental images that are spontaneously evoked and naturally flow over time have been characterized by some as being part of a larger family of cognitive experiences traditionally known as *mind-wandering* (Martarelli, Mayer, & Mast, 2016; Taruffi et al., 2017). In particular, visual mental images are one possible modality (along with words) in which mind-wandering can occur (Delamillieure et al., 2010; Taruffi et al., 2017). Mind-

wandering is a ubiquitous and very frequent mental phenomenon (up to 50% of our waking time; Killingsworth & Gilbert, 2010), and is characterized by the spontaneous emergence of internally-oriented images and thoughts that are largely unrelated to the present, external sensory environment. Although mind-wandering is traditionally understood as a spontaneous mental experience, it can also occur intentionally (Seli, Risko, & Smilek, 2016). Martarelli and colleagues (2016) have used *daydreaming* as a term similar to mind-wandering, emphasizing both the independence from the ongoing perceptual experience (e.g., current task) and the internal focus of attention. Even more recently, Ruth Herbert defines yet another fraternal term *musical daydreams* as "lived experiences of music marked by a fluctuating distributed attentional focus" (2018). Musical daydreams are characterized by a systemic interaction between perceiver, affordances of music and environment, and involve simultaneously distributed internally- and externally-focused attention (differently from mind-wandering or daydreaming, in which attention is decoupled from the sensory environment and is directed inwards; Schooler et al., 2011). According to this definition, visual mental images are merely one of the various components of musical daydreams, which are intrinsically multi-modal phenomena and also feature verbal and non-verbal thought as well as interaction with surroundings (Herbert, 2011). Another common term used in music research is *music-evoked autobiographical memories* (or MEAMs) (Janata et al., 2007; Janata, 2009). While previous definitions have captured the inherently introspective nature of visual mental images, their task-unrelatedness, and/or their multimodal and flexible characteristics, MEAMs pinpoint the role of episodic memory in shaping these mental experiences. Furthermore, terms such as *involuntary imagery* or *involuntary autobiographical memories* have been occasionally used by researchers (e.g., El Haj, Fasotti, & Allain, 2012). However, a disadvantage of extending such terms to visual mental imagery research in general is that they exclude many other experiences in which we purposely

imagine something, for example, to entertain ourselves in a boring situation. It is important to clarify at this point that visual mental images can be spontaneous/involuntary (e.g., while attending a classical concert, a person suddenly thinks of a natural landscape, then imagines the ideal location for the next summer holidays) or deliberate/goal-directed (e.g., during a guided meditation session with music, a person pictures different beloved people as instructed by the therapist). From this brief analysis of the terminology, it is apparent that visual mental images experienced during music listening encompass a broad, heterogeneous, and somewhat fluid spectrum of mental phenomena. While previous definitions differ from one another along quite critical points of departure, they also complement one another and consistently highlight key dimensions of the phenomenon itself, such as intentionality (spontaneous vs. deliberate visual mental imagery) and task-relatedness.

Organizing the broad range of phenomena underlying music-evoked visual mental imagery

Christoff and colleagues (2016) have recently introduced a theoretical framework for understanding various typologies of spontaneous thoughts, including mind-wandering. They define spontaneous cognition as "a mental state, or a sequence of mental states, that arises relatively freely due to an absence of strong constraints on the contents of each state and on the transitions from one mental state to another" (Christoff et al., 2016, p. 719). Importantly, the authors propose two mechanisms that restrict and govern the transition between an individual's various mental states when engaging in the aforementioned cognitive phenomena: *deliberate* and *automatic* constraints. Deliberate constraints are implemented through cognitive control (in a top-down manner; e.g., "we can deliberately maintain our attention on a dry and boring lecture, bringing our thoughts back to the lecture whenever they begin to stray", Christoff et al., 2016, p. 719), whereas automatic constraints allow attention

to be held on a restricted set of information and typically arise due to sensory or affective salience (in a bottom-up manner; e.g., "despite our efforts, we may find ourselves unable to disengage our attention from a fly buzzing in a quiet library or from a preoccupying emotional concern", Christoff et al., 2016, p. 719). According to the authors, mind-wandering, dreaming, and creative thinking all belong to the overarching "cognitive family" of spontaneous thought, but differ in their level of deliberate constraints as well as in their form or modality (i.e. visual mental images or inner language). For example, mind-wandering can be defined as a special case of spontaneous thought that tends to be more deliberately constrained than dreaming, but less deliberately constrained than creative thinking. In addition, mind-wandering can be clearly distinguished from rumination and other types of cognition that are marked by a high degree of automatic constraints, such as obsessive thoughts (Christoff et al., 2016, Fig. 1, p. 719). Neuroimaging literature aligns with this perspective on spontaneous cognition and particularly points to the role of the Salience Network (SN) in automatic constraints (Seeley et al., 2007) and the Frontoparietal Control Network (FPCN) in deliberate constraints (Vincent, Kahn, Snyder, Raichle, & Buckner, 2008). Both SN and FPCN dynamically interact with the Default Mode Network (DMN, the main neural network contributing to internally-oriented cognition; Andrews-Hanna, Reidler, Sepulcre, Poulin, & Buckner, 2010) in regulating the temporal flow of spontaneous thoughts (Christoff et al., 2016).

We believe that this theoretical framework can be successfully applied to the family of visual mental images evoked by music, as it is understood as a modality in which spontaneous cognition can unfold. Further, this framework may parsimoniously account for the distinctions that have been highlighted above (i.e. between the dimensions of intentionality and task-relatedness), encouraging a more nuanced and precise understanding of the several typologies of visual mental images. Within this framework, music-evoked

visual mental imagery shifts in both deliberate and automatic constraints, from weak to strong (Figure 1). For instance, the affective tone conveyed by the music represents a primary source of automatic constraints (see also next section, *Relation to Emotion*), while the context in which visual mental images arise has a strong impact on the level of applied deliberate constraints (e.g., experiencing visual mental imagery while in a music therapy session, where the patient is encouraged to focus on the images evoked by the music, vs. while listening to music in a club, where the listener's imagery is not explicitly directed by external input). When deliberate constraints are weaker, visual mental imagery tends to coincide with spontaneous cognition; inversely, when they are stronger, an opposite shift towards goal-directed cognition occurs. Furthermore, the variability in levels of automatic constraints is helpful in distinguishing between mind-wandering, as evoked in a non-musical context such as driving, and music-evoked visual mental imagery, with the latter being potentially more strongly modulated by the music's features. Regardless of the context in which visual mental imagery is evoked, be it musical or non-musical, personality traits (e.g., neuroticism) and affective disorders (e.g., anxiety and depression) should be marked by higher levels of automatic constraints (Christoff et al., 2016), in turn impacting the phenomenological content of the evoked images (e.g., where an individual with the above trait and disorders experience more negatively hued images; Perkins, Arnone, Smallwood, & Mobbs, 2015; see also the section *Functional Outcome in Daily Life: Benefits & Costs*).

[Place Figure 1 about here]

Relation to Emotion

Images and emotions evoked by music are closely intertwined

Because visual mental imagery was initially proposed as a mechanism of emotion evocation via music (Juslin & Västfjäll, 2008), most of the empirical research conducted so far has focused on the phenomenon's relation to emotion. Recent behavioral and neuroimaging research has provided strong evidence for a link between emotion, mind-wandering, and visual mental imagery (Taruffi et al., 2017). In Taruffi et al.'s study (2017), a group of 216 participants listened to various instrumental music excerpts conveying and evoking emotions of sadness and happiness and reported their level and form (*visual imagery* or *inner language*) of mind-wandering. A separate group of 24 participants underwent fMRI while listening to 4 min uninterrupted instrumental music conveying and evoking emotions of sadness and happiness. Results showed that sad music (compared with happy music) was associated with higher levels of mind-wandering and with stronger activity within the main nodes of the brain's DMN, which is the principal neural contributor to mind-wandering (e.g., Mason et al., 2007). Regarding the form of mental experiences, visual mental imagery was by far the dominant modality for both sad and happy music conditions (compared with inner language), pointing to a strong link between visual mental imagery and music processing. Overall, these findings demonstrate how mind-wandering levels (and likewise DMN activity) can be modulated as function of the emotions evoked by the music, and that music-evoked mind-wandering is predominantly a visual phenomenon, regardless of the emotional experience. However, let it be noted that the authors did not assess the role of musical training, and additionally that both sad and happy music pieces were unfamiliar to participants; therefore, future research should include empirical analysis of whether musical training and/or familiarity, among other factors, can contribute to the modality through which mind-wandering occurs. The findings of Taruffi et al. (2017) are consistent with general mind-wandering literature, which has underscored how affective processes play a pivotal role in shaping the quantity and quality of mind-wandering episodes (Smallwood & Schooler,

2015). Although there is robust evidence of a relation between mind-wandering and negative affect (with higher rates of mind-wandering leading to enhanced negative mood; Killingsworth & Gilbert, 2010; Smallwood, O'Connor, Sudbery, & Obonsawin, 2007), it has also been shown that this association is strongly mediated by the content of thoughts and/or images, with past-related thoughts/images being linked to higher levels of unhappiness (Ruby, Smallwood, Engen, & Singer, 2013; Smallwood & O'Connor, 2011; for more details see the subsection *The phenomenological content of music-evoked visual mental imagery*).

Martarelli and colleagues (2016) manipulated the valence of music to investigate the mediating role of daydreams on relaxation and music liking. They asked participants to listen to one excerpt of sad music ("The Schindler's List Theme" of John Williams) and happy music ("Tritsch-Tratsch-Polka" of Johann Strauss) and to answer a few questions about the valence and vividness of their daydreams as well as the liking and relaxation during the music experience. They found that daydreams arising in the course of the musical experience mediated the effect of the type of music (sad vs. happy) on relaxation and liking, suggesting that daydreams play a functional role when listening to music. Specifically, sad music correlated with increased relaxation, while happy music promoted more positive daydreams, which in turn facilitated relaxation and were associated with participants' increased liking of the music. However, it is important to notice that participants' reports regarding daydreams, relaxation, and liking took place simultaneously. Therefore conclusions about causal direction are not warranted and it is possible that greater relaxation induced by sad music might have influenced the valence of daydreams rather than vice versa.

In a survey study, Küssner and Eerola (same special issue) explored the prevalence, nature, and functions of visual mental imagery in music and how it is associated with domain-specific skills by assessing individual differences in musical sophistication (using the Goldsmiths Musical Sophistication Index, or Gold-MSI; Müllensiefen, Gingras, Musil, &

Stewart, 2014). Interestingly, they observed that one of the main functions of visual mental imagery during music listening is the modulation of emotional arousal. Utilizing a music-related visual imagery questionnaire with 24 items, two components labeled *Vivid Visual Imagery* and *Soothing Visual Imagery* were identified, suggesting that both musically trained and untrained listeners use visual mental imagery to calm or relax them, while musically trained individuals additionally use imagery to feel more energetic and excited. This finding provides further evidence for the importance of visual mental imagery in music-evoked emotions.

Empirical support that visual mental imagery is involved in the evocation of emotion via music was also demonstrated by means of experience sampling, thereby extending the ecological validity of previous studies (Juslin, Liljeström, Västfjäll, Barradas, & Silva, 2008). In this study, 32 college students carried a palmtop that emitted a sound signal seven times per day at random intervals for two weeks. When signaled, participants were required to complete a questionnaire, asking whether they had experienced a music-evoked emotion and if so, what its cause had been. This approach allowed the researchers to explore emotions in music as they naturally occurred in everyday life. Results showed that visual imagery was rated as the fourth most commonly reported cause of emotion (7%), following emotional contagion (32%), brain stem response (25%), and episodic memory (14%). Further, visual imagery was more frequently reported than evaluative conditioning (6%), lyrics (4%), musical expectancy (4%), and cognitive appraisal (2%). These findings were corroborated by Taruffi and Koelsch (2014) for the specific case of music-evoked sadness. Specifically, they found that episodic memory was the most important mechanism for eliciting sadness, followed by emotional contagion and visual imagery. Therefore, both studies (Juslin et al., 2008; Taruffi & Koelsch, 2014) point to the involvement of visual mental imagery in emotion evocation, although this does not seem to be the primary mechanism through which emotions

are evoked by music. However, both studies are limited by the use of self-reports; these are in fact not the ideal means to assess the mechanisms of emotion evocation, which describe processes that may occur partly unconsciously, thereby proving more difficult for participants to directly report. With regard to the specific case of fear, an fMRI study showed that listening to fear-evoking music is associated with an increased functional connectivity between the superficial amygdala and the visual cortex, suggesting that visual mental imagery may play a crucial role in the aesthetic appeal of fear-evoking music (Koelsch et al., 2013).

To sum up, the empirical studies outlined above indicate that visual mental imagery and emotional processes during music listening are closely intertwined. In particular, the emotional tone of the music is linked to the amount and vividness of these mental experiences (Martarelli et al., 2016; Taruffi el al., 2017). Moreover, visual mental imagery appears to be a common phenomenon during music listening, often leading to evocation of emotion (Juslin et al., 2008; Küssner & Eerola, same special issue). Because individual differences are an inherent part of emotional processes evoked via music (e.g., Vuoskoski, Thompson, McIlwain, & Eerola, 2012), it will be essential for future research to investigate how listener attributes, such as personality traits, contribute to modulating the link between visual mental imagery and specific emotions induced by music. In the next subsection, we will discuss a recent neuroimaging study suggesting that individual differences in trait empathy are in fact associated with higher levels of visual mental imagery during sad music (Taruffi, Pehrs, Skouras, & Koelsch, 2018).

Visual mental imagery, sadness, and empathy

Empathy, or the capacity to experience and understand someone else's feelings as one's own (Zaki & Ochsner, 2012), often underlies music listening experiences. For example,

we may empathize with emotions that are presumed to be experienced or expressed by the composer and/or the performer (e.g., Scherer & Zentner, 2001) or we may internally "mirror" the perceived emotional expression of music itself (e.g., Davies, 2011). A third way through which listeners can empathize with the music consists of experiencing music as a narrative regarding an indefinite "musical persona" and empathizing with it and/or the imagined events (Levinson, 1997). This is similar to what readers do with a novel's fictional character, but with the difference that music listeners can imagine their own narrative unfolding on the basis of the musical events. As Levinson explained: "When we identify with music that we are perceiving or perhaps better, with the person whom we imagine owns the emotions or the emotional gestures we hear in the music – we share in and adopt those emotions as our own, for the course of the audition" (Levinson, 1997, p. 228). According to Levinson, such empathic processes that occur during music listening are largely based on the ability to imagine the music as a person, therefore linking empathy to visual mental imagery skills.

Interestingly, preliminary evidence from neuroscience suggests that individual differences in trait empathy are indeed associated with higher levels of visual mental imagery during sad music as reflected by stronger activity of the primary visual cortex (Taruffi et al., 2018). In this study, 24 participants underwent fMRI while listening to 4 min blocks of instrumental sad- and happy-evoking music (with their eyes closed) and completed a self-report measure of empathy (Interpersonal Reactivity Index, or IRI; Davis, 1980). Eigenvector centrality mapping and functional connectivity analyses were used to identify a functional network of brain regions underlying empathic responses to sad (compared with happy) music. Within this brain network, the primary visual cortex exhibited the highest centrality values and was by far the largest region engaged during sad music. These findings are important because they underscore that, in empathic individuals, specific emotional experiences (in this case, sadness) are associated with enhanced activity in brain structures involved in visual

mental imagery, corroborating the link between visual mental imagery and emotion as highlighted by previous behavioral studies (e.g., Martarelli et al., 2016; Taruffi et al., 2017). In other words, when listening to instrumental sad music, individuals with a predisposition to empathize engage in highly dense fantasies, pointing to strong involvement of visual mental imagery in empathic responses to sad music.

While the fMRI experiment by Taruffi et al. (2018) underscores the importance of the individual differences approach to studying the relation between visual mental imagery and emotion, another outstanding issue focuses on the causal relationship between visual mental imagery and felt emotions. Although Juslin and Västfjäll (2008) discussed visual mental imagery as being a precursor of emotion, there is no robust evidence on this yet: Does visual mental imagery precede emotion or vice versa? Or do emotions and images co-occur during music listening?

What is the causal link between visual mental imagery and affective processes?

Day and Thompson (same special issue) provided evidence that felt emotions actually precede visual mental imagery (in contrast to Juslin & Västfjäll's theory on visual imagery as a mechanism of emotion evocation). In a music listening experiment, they asked participants to indicate by key pressing when they (*i*) recognized an emotion in music, (*ii*) experienced an affective response to the music, and (*iii*) experienced a visual image related to music. Response-times were measured for 49 participants as they listened to thirty 20-second music excerpts of classical or pop music. Their results showed that the time taken to *perceive* an emotion was significantly shorter than the time taken to *feel* an emotion, which was, in turn, significantly shorter than the time taken to experience a visual image. While these findings suggest that emotional states in response to music typically precede the formation of visual imagery, this could be due to the quality of the emotional responses involved in participants'

task, presented as low-level affects described in terms of pleasantness, energy, and tension. More complex, mixed "aesthetic" emotions, such as nostalgia (Zentner, Grandjean, & Scherer, 2008), may in fact require more time to develop or may lead to other causal relations with evoked images.

This hypothesis was recently tested by Vroegh (2018). In this study, data from an online survey ($N = 602$) were analyzed using structural equation modeling. Participants were asked to listen to music excerpts varying in valence and arousal and to provide information regarding their attentional focus, frequency and vividness of imagery, and quality of felt emotions. The best-fitting model accounting for the data was characterized by a unidirectional effect from emotion to imagery when explicitly positive emotions were concerned; however, in the case of mixed emotions, the results pointed to a reversed effect, from imagery to felt emotions. Importantly, these findings (Day & Thompson, same special issue; Vroegh, 2018) demonstrate the complexity of the relation between felt emotions and visual mental imagery and point to the crucial role of the quality of the affective experience (Vroegh, 2018). In this sense, future research will benefit from employing objective indices of emotion (i.e. neurophysiological correlates) and methods to assess the temporal dynamics (such as electroencephalography, or EEG) of cognitive and emotional processes during music listening.

Functional Outcome in Daily Life: Benefits & Costs

This section considers evidence of the benefits and costs of music-evoked visual mental imagery, emphasizing that the context in which the images unfold is crucial in determining the impact of the overall experience. Similarly, the content of these images, which appears to be modulated by the music's features, is likely to be fundamental in

determining the impact of visual imagery processes on mood, well being, and mental health. In this sense, the capacity of music to regulate visual mental imagery has been informally harnessed mostly in therapeutic contexts, but also in music performance, for example, in accomplishing desired interpretations or sound qualities (Black, 2017). On the other hand, the role of visual imagery in music performance is ambivalent, since it can also lead to undesired distraction. Finally, individual differences in visual imagery abilities can impact other cognitive-related phenomena such as episodic memory and future thinking. We explore possible ideas on how music could be harnessed to optimize such mental processes in daily life.

Use in therapy

Visual mental images have a pivotal role in self-understanding and psychotherapy. The archetypal psychology of Carl Jung (1969) was the first to recognize the importance of the myriad of fantasies and archetypes that shape and are shaped by our psychological lives. In this framework, spontaneous images are considered to be symbols or archetypes, which can open the door of the unconscious mind and help to endorse authentic psychological change. According to Jung, archetypes embody emotions and are collective and transpersonal, yet when activated, they are experienced by each individual in a unique personal way (Jung, 1969). Because detrimental mental imagery acting unconsciously can lead to emotional and physical distress, it is essential to bring these images to consciousness, allowing the healing process to unfold (Jung, 1969).

Visual mental imagery plays a pivotal role in a wide range of mental and neurological disorders (from post-traumatic stress disorder [PTSD], depression, anxiety disorders to bipolar disorder and schizophrenia) and their treatments (for reviews see Holmes & Mathews,

2010; Pearson, Naselaris, Holmes, & Kosslyn, 2015). For instance, suicide-related imagery is an important component in the phenomenology of depression, and imagining suicidal acts can increase the risk of suicide (Crane, Shah, Barnhofer, & Holmes, 2012). Further, one of the most distinctive features of PTSD is the involuntary intrusion of vivid images in which the traumatic scenes are re-experienced as though they were occurring in the present (i.e. flashbacks; Brewin, 2014). These frequent involuntary images do not simply co-exist with PTSD but represent a cognitive mechanism driving the maintenance of the ongoing clinical disorder (Brewin, 2014). Psychological therapies have traditionally focused on verbal communication rather than on exploring the patient's visual imagery. However, applications of imagery in therapy are mounting and represent still an underexplored field, with special relevance for those clinical conditions mentioned above where intrusive visual mental imagery is central. Examples of currently employed imagery-focused techniques are cognitive behavioral therapy (involving repeated imaginary exposure to a feared object in patients with anxiety disorders; Foa, Steketee, Turner, & Fischer, 1980), imagery re-scripting (featuring the switch of negative to positive imagery; Holmes et al., 2007), systematic desensitization (aiming to remove the fear response of a phobia, and to substitute a relaxation response to the conditional stimulus gradually using counter conditioning; Wolpe, 1958), and eye movement desensitization and reprocessing (triggering lateral eye movements during the recall of emotional memories to diminish their vividness; van den Hout et al., 2012). Noteworthy imagery-focused cognitive behavioral therapy has the strongest success rate (up to 75%) on treating PTSD and social phobia (Pearson et al., 2015).

Visual mental imagery is also actively used in combination with music in one of the most effective types of music therapy known as the "Bonny Method of Guided Imagery and Music" (GIM; Bonny, 2002), which employs Western classical music to stimulate imagery in order to access and work through emotional processes. In a typical GIM session, the client

reclines with his or her eyes closed while listening to music "to work on significant life issues, for instance, disturbing old memories, losses, traumata, bothering health conditions, and relationship issues. While being guided, strong emotions are released and the client finds helpful resolutions" (Kestele, 2018). The music is chosen by the therapist to match the client's imagery or to facilitate working on a specific personal issue. GIM is a well-established method, encompassing different programs featuring carefully selected music pieces. The "Nurturing" program, for example, features music evoking attachment-related feelings such as tenderness. GIM has been linked to positive health outcome (e.g., on mood and quality of life in cancer patients; Burns, 2001) and has been shown to reduce cortisol levels (McKinney et al., 1997a) as well as β-endorphins (McKinney et al., 1997b), which are two biomarkers of psychological stress. However, both studies (McKinney et al., 1997a, 1997b) require replication of findings in order to isolate the effects of the music *per se*, given that appropriate control conditions (i.e. music listening and imagery alone) were not provided (Chanda & Levitin, 2013). In fact, the relation between the therapist and the client is likely to contribute significantly to the effectiveness of the therapy. Along these lines, a recent pioneering experiment employed a dual-EEG approach in the context of a "real-world" GIM session to explore the neural underpinnings of social interactive behavior between the therapist and the client (Fachner, 2018). Synchronously with EEG, the session was video-taped and subsequently analyzed by two independent raters (as well as the therapist and client) to identify three important moments in the session. The results revealed that emotional processing (reflected by Frontal Alpha Asymmetries; FAA) was associated with the emotional impact of the emerging imagery. In particular, a positive shift in the therapist's FAA indicated a strong emotional response, which paralleled the client's FAA peak during the emergence of crucial personal imagery. Importantly, this study is the first to show how FAA dynamics can be used to describe shared emotional processes evolving in time in a

"real-world" GIM session, providing evidence for the beneficial emotional effects of imagery in a social interaction context.

Episodic memory and future thinking

As mentioned in the *Introduction*, a mental image is generally created from previously stored information held in long-term memory. This close relation between visual mental imagery and memory is supported by lesion studies showing that damage to areas known to support visual imagery can result in an impairment of memory (e.g., Greenberg & Rubin, 2003). A previous study examined whether individual differences in visual imagery abilities (as measured by the Vividness of Visual Imagery Questionnaire, or VVIQ; Marks, 1973) affect memory for past events and/or future thinking (D'Argembeau & Van der Linden, 2006). The VVIQ comprises 16 items referring to different situations that the participant is asked to visualize; subsequently, the participant rates image vividness on a 5-point scale. For both past and future events, the VVIQ significantly predicted the amount of visual and other sensory details, indicating that participants with more vivid visual imagery created representations of past and future events that contained more visual and other sensory details (D'Argembeau & Van der Linden, 2006). Similarly, mind-wandering research has demonstrated that individuals with high working memory capacity are more likely to engage in prospective mind-wandering and that prospective mind-wandering frequently involves autobiographical planning (Baird, Smallwood, & Schooler, 2011). Together this evidence suggests that visual imagery abilities can mediate the benefits that memory and future thinking can bring to daily life, by enabling cognitive operations that help individuals to better navigate the world and by creating and sustaining social systems (Baumeister & Masicampo, 2010). Along these lines, the capability of music to stimulate vivid visual images

could be used to improve "mental time travel" skills, including both remembering the past and projecting into the future, even though research has shown that music-related visual imagery and the vividness of visual imagery in general do not necessarily overlap (Küssner & Eerola, same special issue).

Distracted listening

Despite the situations highlighted above in which visual mental images can play a vital role in healthy cognition and emotion, visual imagery can also have negative consequences for listeners and musicians, especially when it acts as a distraction undermining music listening and performance. Distractions in the form of visual mental images can be both internally and externally originated, and create a temporal switch of individuals' attention away from the music (Gritten, 2016). This type of disruptive visual mental imagery seems to occur predominantly in the form of spontaneous imagery and overlaps with *task-unrelated thought*, a term used to refer to mind-wandering that is unrelated to one's ongoing task (Smallwood & Schooler, 2006). Task-unrelated thoughts are associated with costly outcomes such as impaired reading comprehension and test-taking, attenuated processing of the environment, automobile accidents, disruptions to learning, and affective dysfunction (for a review see Smallwood & Schooler, 2015). These detrimental outcomes of visual imagery and mind-wandering raise the question of what strategies may be used to minimize the aforementioned negative consequences. For the specific case of music, Gritten (2018) has suggested to incorporate distractions into the listening or performance as an opportunity to enhance aesthetics and creativity. In essence, Gritten argues for distracted attention to be embraced as a temporal "switch", which adds new value to music's dynamic unfolding over time by unlocking attentional resources in the listener. Moreover, *meta-awareness*, the ability

to take explicit note of the current contents of consciousness (Schooler et al., 2011), appears to be the most effective element that can counterbalance the negative effects of task-unrelated thoughts and images. As Smallwood and Schooler observed: "regularly checking in on the contents of one's mind may help to curtail episodes of mind-wandering" (2015, p. 23). In this regard, meta-awareness of mind-wandering was found to mediate the relationship between attention deficit hyperactivity disorder symptomatology and detrimental mind-wandering, suggesting that some of the negative consequences of mind-wandering can be ameliorated by strategies that facilitate meta-awareness (Franklin et al., 2017).

The phenomenological content of music-evoked visual mental imagery

The phenomenological content of visual mental imagery evoked by music is an important factor underlying its costs and benefits in daily life. Mind-wandering research suggests that individuals who are able to regulate the content of their spontaneous thoughts and images to positive or productive topics exhibit improved emotional health and well being, whereas negative or unproductive thought and image content is indicative of poor emotional health and well being (Smallwood & Andrews-Hanna, 2013). For example, images focused on the future (Ruby et al., 2013) or rated as interesting (Franklin et al., 2013) correlate with subsequent positive mood, and future thinking reduces cortisol levels following social stress (Engert, Smallwood, & Singer, 2014). Furthermore, *rumination*, a style of thinking characterized by repetitive thoughts focused on negative or self-related content, is a core feature of depression (Nolen-Hoeksema, 2000). In a similar way, the content of visual images evoked during music listening may contribute to congruent effects on subsequent mood and overall well being in the long-term.

Despite the fact that visual mental images are heterogeneous mental representations, a couple of studies exploring the content of these experiences have provided preliminary evidence that they can be organized along different phenomenological dimensions. Taruffi and colleagues (2017) showed that images during sad-evoking music are focused on emotions (both positive and negative) and nature, while happy-evoking music is linked to dancing imagery. The issue of content in music-related visual imagery also clearly concerns questions of valence and vividness. Work by Martarelli et al. (2016) has explored vividness and valence of daydreams experienced while listening to sad and happy music. Interestingly, the authors observed that sad music was associated with less vivid and more negative daydreams, whereas happy music was linked to more vivid and positive daydreams. These results are in line with the findings of Taruffi et al. (2017), in that the valence of images is significantly more positive during happy music. However, Taruffi et al. (2017) also found that images evoked during sad music featured both positive and negative emotions. Similarly, emotions evoked during sad music were rated overall pleasurable and positive, and did not significantly differ in valence from that of happy music (Taruffi el al., 2017). In this regard, it is important to note that sadness in music is often described by listeners as a melancholic yet pleasurable (Taruffi & Koelsch, 2014); therefore, its relation to mental images may be more complex and nuanced as compared to being a mood-congruent and straightforward relationship (i.e. sad music being linked to negatively hued images). Despite these inconsistencies (which should be addressed by future research), both studies (Martarelli et al., 2016; Taruffi el al., 2017) reveal that the content of visual mental images is highly sensitive to the emotional tone of the music, and that it can be productively examined and distilled into several major dimensions (for an example of content analysis and its dimensions, see Fig. 2 in Taruffi et al., 2017).

Furthermore, future studies should investigate which structural or acoustic aspects of the music are linked to dimensions of image content; this could be formulated as an issue of

abstraction vs. concreteness in imagery. For example, is it the more abstract melodic features (e.g., a series of pitches) or the particular expressive features (e.g., timbral, articulatory properties) of a sound fragment that are the most potent triggers of imagery? This potential line of research may contribute to building a taxonomy of music-evoked visual mental imagery including different levels from low to high degrees of resolution (i.e. specificity/generality). Most importantly, a better understanding of the content of music-evoked visual mental imagery may explain a considerable amount of variance in constructs relating to emotional health and well being, and could therefore be a valuable means in assisting music therapy practices. A crucial question to address will be identifying the directional nature of the relationship between imagery content and psychological well being. Does unhappiness or tendency to depression have a causal influence on the valence and personal significance of images, or does engagement with such images lead to poor emotional well being?

Pitfalls, Outstanding Issues, and Future Directions

The covert nature of visual mental imagery, or any other form of mental imagery, for that matter, leads to difficulties in its assessment because visual mental images are subjective and internally-oriented mental states. Therefore, it is crucial to recognize that our current understanding of visual mental imagery closely relies on how the researcher faces challenges in inducing, measuring, and controling such internal states of imagery. As highlighted by this review, investigations of music-evoked visual mental imagery have predominantly employed retrospective survey methods (i.e. asking participants about their previous experience of visual imagery during music listening; Küssner & Eerola, same special issue), probe-caught thought sampling (i.e. intermittently probing individuals regarding their current mental state while listening to music; Taruffi et al., 2017), and music listening behavioral experiments

(i.e. assessing the inner experience of participants as they complete a music listening task in a controlled experimental setting; Day & Thompson, same special issue; Martarelli et al., 2016; Vuoskoski & Eerola, 2015). Retrospective questionnaires can be inaccurate in their measurement of visual imagery, due to their vulnerability to memory biases. Experience sampling approaches may guarantee a higher degree of ecological validity as compared to other assessments (both surveys and music listening experiments), since they capture real-time, *in vivo* participants' mental states (Csikszentmihalyi & LeFevre, 1989). Despite these differences, all studies reviewed in this article ultimately assessed visual mental imagery evoked by music via self-reports, leading to obvious limitations such as memory biases and demand characteristics (Stone et al., 1998). In this sense, it is important for future research to corroborate self-reports of visual mental imagery with more objective external measurements (e.g., neurophysiological indices) to ensure that the results are not simply a consequence of self-reports' shortcomings. The same argument also applies to the assessment of music-evoked emotions, which, in the literature reviewed in this article, were mainly measured via self-reports.

Another issue is the primacy of visual perception, which seems to extend to imagery as well, but must not obscure our view on important non-visual forms of imagery. As discussed in the *Introduction*, experiences of music-evoked imagery are likely multimodal, as is our perception of the world. A stronger focus on other forms of imagery such as that which is motor, tactile, olfactory, and gustatory, is thus necessary to capture the full breadth of music-evoked imagery. Metaphorical descriptions of music (e.g., "silky smooth" sounds, "rough" textures or "bittersweet" harmonies)—which are grounded in our experiences of the world and also a part of our mental life—are pervasive in everyday language and a common way to make sense of musical episodes. It will thus be our task in future studies to ascertain how non-visual imagery might fit into the framework of music-evoked imagery.

Another major shortcoming of the above reviewed literature is that generalizations were often drawn without assessing individuals' own disposition to imagery. Because people vary widely in their imagery abilities (general imagery as well as across the various modalities) and in their reported frequency, content, and vividness of visual mental images, this variability must be taken into consideration before making general conclusions about the nature of visual mental imagery. For instance, the neuroimaging study by Taruffi et al. (2018) underlines that differences in empathic abilities correspond to varied levels of activity in the primary visual cortex during listening to sad and happy music. Future research should also consider the relation between individual listeners and the music, and therefore assess variables such as familiarity with the music piece and/or genre (e.g., does listening to self-chosen music vs. experimenter-selected music contribute to potential differences in the frequency and/or content of evoked visual mental imagery?).

Despite being a prominent topic of research in cognitive psychology, neuroscience and philosophy, visual mental imagery in a musical context has yet to be fully explored; this is especially important given its influence on a wide range of phenomena including emotion, episodic memory, prospective thinking, health and well being, creativity, and artistic expressions (D'Argembeau & Van der Linden, 2006; Holmes et al., 2007; Nanay, 2015; Pearson, 2007). Importantly, visual mental imagery has implications across a wide range of mental disorders, including those pertaining to intrusive images in PTSD, hallucinatory imagery in schizophrenia, and intrusive memories of negative past events in depression (Pearson, Deeprose, Wallace-Hadrill, Heyes, & Holmes, 2013). It is therefore crucial to investigate the modulatory impact of music on visual mental images, possibly leading to innovative opportunities for treatment and intervention in healthy and clinical individuals (Taruffi & Koelsch, 2017). The capability of music in evoking visual mental imagery for therapeutic purposes was already proven to be successful by the GIM method (Burns, 2001).

At present, however, a systematic phenomenology of music-evoked visual imagery has yet to be provided. Additionally, the process by which musical and acoustical features map onto specific types or characteristics of imagery remains unknown. Shedding light on the qualitative content of music-evoked images and on factors (i.e. music, listener, and context) that play a role in shaping imagery will be fundamental for creating *ad hoc* music interventions. With regard to emotion, although the literature reviewed in this article highlights that emotional experiences and visual mental images are closely intertwined and that quality of the emotion is crucial in determining the causal link between emotions and images, the nature of this relation is still blurry. We do not know, for example, whether different emotions mediate specific features of visual images (e.g., color, temporal orientation, social characteristics, concreteness vs. abstraction, etc.) or, in other words, whether there exists a higher-level modal correspondence between emotion and visual mental imagery.

Another outstanding issue to be addressed by future research includes the intentional or spontaneous nature of visual mental imagery processes during music. For instance, it remains to be specified which type of visual mental imagery—deliberate or spontaneous—is more typically experienced during music listening. This is especially important in describing the functional outcome of these internally-oriented mental experiences in daily life. Furthermore, visual mental images do not only reflect perceptual components of physical and natural environments, but also cultural and social dynamics (Duncum, 2001). Therefore, an investigation of how music-evoked visual mental imagery varies across cultures is important (Juslin, Barradas, Ovsiannikow, Limmo, & Thompson, 2016), yet is lacking to date, and may lead to a better understanding of how diverse cultural backgrounds may influence internally-oriented cognitive phenomena. In fact, music is a strongly culturally connoted art form; cross-cultural research has offered evidence of both culture-sensitive and culture-invariant

mechanisms of music perception (Balkwill, Thompson, & Matsunaga, 2004; McDermott, Schultz, Undurraga, & Godoy, 2016), leading to open questions such as the following: What cross-cultural aspects are operative in visual mental imagery processes during music listening? Are differences in music's tonal system (i.e., culturally familiar vs. unfamiliar music) reflected by variations of visual images' features? Finally, examining the relation between music and images has the potential to shed light on the nature of cross-modal associations between visual and auditory modes of perception—a topic that has been investigated for more than a hundred years (Pierce, 1901), but that it is still poorly understood—thereby contributing to the explanation of complex perceptual phenomena, such as synaesthesia.

References

Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & de Lange, F. P. (2013). Shared representations for working memory and mental imagery in early visual cortex. *Current Biology*, *23*, 1427-1431.

Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R., & Buckner, R. L. (2010). Functional-anatomic fractionation of the brain's default network. *Neuron*, *65*, 550-562.

Baird, B., Smallwood, J., & Schooler, J. W. (2011). Back to the future: Autobiographical planning and the functionality of mind-wandering. *Consciousness and Cognition*, *20*, 1604-1611.

Balkwill, L. L., Thompson, W. F., & Matsunaga, R. (2004). Recognition of emotion in Japanese, Western, and Hindustani music by Japanese listeners. *Japanese Psychological Research*, *46***,** 337-349.

Baumeister, R. F., & Masicampo, E. J. (2010). Conscious thought is for facilitating social and cultural interactions: How mental simulations serve the animal–culture interface. *Psychological Review*, *117*, 945-971.

Bernardi, N., De Buglio, M., Trimarchi, P., Chielli, A., & Bricolo, E. (2013). Mental practice promotes motor anticipation: Evidence from skilled music performance. *Frontiers in Human Neuroscience, 7,* 451.

Black, M. (2017). *Don't sing it with the face of a dead fish! Can imagery stimulate a vocal response in choral rehearsals?* Paper presented at the KOSMOS Dialogue "Music, Emotion, and Visual Imagery", Berlin, Germany, June 1-3, 2017.

Bonny, H. L. (2002). *Music & consciousness: The evolution of guided imagery and music*. Gilsum: Barcelona Publishers.

Brewin, C. R. (2014). Episodic memory, perceptual memory, and their interaction: Foundations for a theory of posttraumatic stress disorder. *Psychological Bulletin*, *140*, 69-97.

Burns, D. S. (2001). The effect of the bonny method of guided imagery and music on the mood and life quality of cancer patients. *Journal of Music Therapy*, *38*, 51-6.

Chanda, M. L., & Levitin, D. J. (2013). The neurochemistry of music. *Trends in Cognitive Sciences*, *17*, 179-193.

Christoff, K., Irving, Z. C., Fox, K. C., Spreng, R. N., & Andrews-Hanna, J. R. (2016). Mind-wandering as spontaneous thought: A dynamic framework. *Nature Reviews Neuroscience*, *17***,** 718-731.

Cox, A. (2016). *Music and embodied cognition: Listening, moving, feeling, and thinking*. Bloomington: Indiana University Press.

Crane, C., Shah, D., Barnhofer, T., & Holmes, E. A. (2012). Suicidal imagery in a previously depressed community sample. *Clinical Psychology & Psychotherapy*, *19*, 57-69.

Csikszentmihalyi, M., & LeFevre, J. (1989). Optimal experience in work and leisure. *Journal of Personality and Social Psychology*, *56*, 815–822.

Cuddy, L. L., Sikka, R., Silveira, K., Bai, S., & Vanstone, A. (2017). Music-evoked autobiographical memories (MEAMs) in Alzheimer disease: Evidence for a positivity effect. *Cogent Psychology*, *4*, 1277578.

D'Argembeau, A., & Van der Linden, M. (2006). Individual differences in the phenomenology of mental time travel: The effect of vivid visual imagery and emotion regulation strategies. *Consciousness and Cognition*, *15*, 342-350.

Davies, S. (2011). Infectious music: Music-listener emotional contagion. In A. Coplan, P. Goldie (Eds.), *Empathy. Philosophical and psychological perspectives* (pp. 134-148). Oxford, UK: Oxford University Press.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, *10*, 85.

Day, R. A., & Thompson, W. F. (same special issue). Measuring the onset of experiences of emotion and imagery in response to music. *Psychomusicology: Music, Mind, and Brain*.

Delamillieure, P., Doucet, G., Mazoyer, B., Turbelin, M. R., Delcroix, N., Mellet, E., ... & Joliot, M. (2010). The resting state questionnaire: An introspective questionnaire for evaluation of inner experience during the conscious resting state. *Brain Research Bulletin, 81***,** 565-573.

Di Rienzo, F., Debarnot, U., Daligault, S., Saruco, E., Delpuech, C., Doyon, J., . . . Guillot, A. (2016). Online and offline performance gains following motor imagery practice: A comprehensive review of behavioral and neuroimaging studies. *Frontiers in Human Neuroscience*, *10*, 315.

Duncum, P. (2001). Visual culture: Developments, definitions, and directions for art education. *Studies in Art Education*, *42***,** 101-112.

Eitan, Z., & Granot, R. Y. (2006). How music moves: Musical parameters and listeners' images of motion. *Music Perception, 23*, 221-248.

El Haj, M., Fasotti, L., & Allain, P. (2012). The involuntary nature of music-evoked autobiographical memories in Alzheimer's disease. *Consciousness and Cognition*, *21*, 238-246.

Engert, V., Smallwood, J., & Singer, T. (2014). Mind your thoughts: Associations between self-generated thoughts and stress-induced and baseline levels of cortisol and alpha-amylase. *Biological Psychology*, *103*, 283-291.

Fachner, J. (2018). *Sharing altered states and the emanation of imagery in GIM music therapy settings.* Paper presented at the KOSMOS Workshop "Mind Wandering and Visual Mental Imagery in Music", Berlin, Germany, May 16-19, 2018.

Foa, E. B., Steketee, G., Turner, R. M., & Fischer, S. C. (1980). Effects of imaginal exposure to feared disasters in obsessive-compulsive checkers. *Behaviour Research and Therapy*, *18*, 449-455.

Franklin, M. S., Mrazek, M. D., Anderson, C. L., Johnston, C., Smallwood, J., Kingstone, A., & Schooler, J. W. (2017). Tracking distraction: The relationship between mind-wandering, meta-awareness, and ADHD symptomatology. *Journal of Attention Disorders*, *21*, 475-486.

Franklin, M. S., Mrazek, M. D., Anderson, C. L., Smallwood, J., Kingstone, A., & Schooler, J. W. (2013). The silver lining of a mind in the clouds: Interesting musings are associated with positive mood while mind-wandering. *Frontiers in Psychology*, *4*, 583.

Godøy, R. I. (2003). Motor-mimetic music cognition. *Leonardo, 36*, 317-319.

Greenberg, D. L., & Rubin, D. C. (2003). The neuropsychology of autobiographical memory. *Cortex*, *39*, 687-728.

Gritten, A. (2016). Distraction in polyphonic gesture. In E. King, A. Gritten (Eds.), *New perspectives on music and gesture* (pp. 125-148). London, UK: Routledge.

Gritten, A. (2018). *Distracted listening: Problems, potentials, and pragmatics.* Paper presented at the KOSMOS Workshop "Mind Wandering and Visual Mental Imagery in Music", Berlin, Germany, May 16-19, 2018.

Gurr, B., Foxhall, M., Shinoda, J., & Baird, A. (2014). Rebuilding identity after brain injury: Standard cognitive and music-evoked autobiographical memory training. *International Journal of Therapy and Rehabilitation*, *21*, 289-296.

Halpern, A. R. (2001). Cerebral substrates of musical imagery. *Annals of the New York Academy of Sciences*, *930*, 179-192.

Herbert, R. (2011). *Everyday music listening: Absorption, dissociation and trancing.* Aldershot, UK: Ashgate.

Herbert, R. (2018). *Everyday musical daydreams and kinds of consciousness.* Paper presented at the KOSMOS Workshop "Mind Wandering and Visual Mental Imagery in Music", Berlin, Germany, May 16-19, 2018.

Holmes, E. A., Arntz, A., & Smucker, M. R. (2007). Imagery rescripting in cognitive behaviour therapy: Images, treatment techniques and outcomes. *Journal of Behavior Therapy and Experimental Psychiatry*, *38*, 297-305.

Holmes, E. A., & Mathews, A. (2005). Mental imagery and emotion: A special relationship *Emotion*, *5*, 489-497.

Holmes, E. A., & Mathews, A. (2010). Mental imagery in emotion and emotional disorders. *Clinical Psychology Review*, *30*, 349-362.

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

Ishai, A., Ungerleider, L. G., & Haxby, J. V. (2000). Distributed neural systems for the generation of visual images. *Neuron*, *28*, 979-990.

Jampole, E. S., Mathews, F. N., & Konopak, B. C. (1994). Academically gifted students' use of imagery for creative writing. *The Journal of Creative Behavior*, *28*, 1-15.

Janata, P., Tomic, S. T., & Rakowski, S. K. (2007). Characterisation of music-evoked autobiographical memories. *Memory*, *15*, 845-860.

Janata, P. (2009). The neural architecture of music-evoked autobiographical memories. *Cerebral Cortex*, *19*, 2579-2594.

Jeannerod, M., & Decety, J. (1995). Mental motor imagery: A window into the representational stages of action. *Current Opinion in Neurobiology*, *5*, 727-732.

Johnson, M., & Larson, S. (2003). "Something in the way she moves"- metaphors of musical motion. *Metaphor and Symbol, 18*, 63-84.

Jung, C. G. (1969). *The archetypes of the collective unconscious: Vol. 9. The collective works of C. G. Jung.* Princeton, NJ: Princeton University Press.

Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews, 10*, 235-266.

Juslin, P. N., Barradas, G. T., Ovsiannikow, M., Limmo, J., & Thompson, W. F. (2016). Prevalence of emotions, mechanisms, and motives in music listening: A comparison of individualist and collectivist cultures. *Psychomusicology: Music, Mind, and Brain*, *26*, 293-326.

Juslin, P. N., & Laukka, P. (2004). Expression, perception, and induction of musical emotions: A review and a questionnaire study of everyday listening. *Journal of New Music Research*, *33*, 217-238.

Juslin, P. N., Liljeström, S., Västfjäll, D., Barradas, G., & Silva, A. (2008). An experience sampling study of emotional reactions to music: Listener, music, and situation. *Emotion*, *8*, 668-683.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*, 559-575.

Kant, I. (1965). *Critique of pure reason.* New York: St. Martin's Press.

Kestele, G. (2018). http://www.gim-trainings.com/about.html (last access 15 June 2018).

Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science, 330*, 932.

Koelsch, S., Skouras, S., Fritz, T., Herrera, P., Bonhage, C., Küssner, M. B., & Jacobs, A. M. (2013). The roles of superficial amygdala and auditory cortex in music-evoked fear and joy. *Neuroimage*, *81*, 49-60.

Kosslyn, S. M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.

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

Kosslyn, S. M., & Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery? *Psychological Bulletin*, *129*, 723-746.

Küssner, M. B., & Eerola, T. (same special issue). The content and functions of vivid and soothing visual imagery during music listening: Findings from a survey study. *Psychomusicology: Music, Mind, and Brain.*

Lehne, M., Rohrmeier, M., & Koelsch, S. (2013). Tension-related activity in the orbitofrontal cortex and amygdala: An fMRI study with music. *Social Cognitive and Affective Neuroscience*, *9*, 1515-1523.

Levinson, J. (1997). Music and negative emotion. In J. Robinson (Ed.), *Music and meaning* (pp. 215–241). Ithaca, NY: Cornell University Press.

Lotze, M., & Halsband, U. (2006). Motor imagery. *Journal of Physiology-Paris, 99*, 386-395.

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, *64*, 17–24.

Martarelli, C. S., Mayer, B., & Mast, F. W. (2016). Daydreams and trait affect: The role of the listener's state of mind in the emotional response to music. *Consciousness and Cognition*, *46*, 27-35.

Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., & Macrae, C. N. (2007). Wandering minds: The default network and stimulus-independent thought. *Science*, *315*, 393-395.

McDermott, J. H., Schultz, A. F., Undurraga, E. A., & Godoy, R. A. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature*, *535***,** 547-550.

McKinney, C. H., Antoni, M. H., Kumar, M., Tims, F. C., & McCabe, P. M. (1997a). Effects of guided imagery and music (GIM) therapy on mood and cortisol in healthy adults. *Health Psychology*, *16*, 390–400.

McKinney, C. H., Tims, F. C., Kumar, A. M., & Kumar, M. (1997b). The effect of selected classical music and spontaneous imagery on plasma β-endorphin. *Journal of Behavioral Medicine*, *20*, 85–99.

Meyer, L.B. (1956). *Emotion and Meaning in Music.* Chicago: Chicago University Press.

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: An index for assessing musical sophistication in the general population. *PLoS One*, *9*, e89642.

Nanay, B. (2015). Perceptual content and the content of mental imagery. *Philosophical Studies*, *172***,** 1723-1736.

Nanay, B. (2018). Multimodal mental imagery. *Cortex, 105*, 125-134.

Nolen-Hoeksema, S. (2000). The role of rumination in depressive disorders and mixed anxiety/depressive symptoms. *Journal of Abnormal Psychology*, *109*, 504-511.

Overy, K., & Molnar-Szakacs, I. (2009). Being together in time: Musical experience and the mirror neuron system. *Music Perception, 26*, 489-504.

Pearson, D. C. (2007). Mental imagery and creative thought. In I. Roth (Ed.), *Imaginative minds* (pp. 187–212). Oxford, UK: Oxford University Press.

Pearson, D. G., Deeprose, C., Wallace-Hadrill, S. M., Heyes, S. B., & Holmes, E. A. (2013). Assessing mental imagery in clinical psychology: A review of imagery measures and a guiding framework. *Clinical Psychology Review*, *33*, 1-23.

Pearson, J., Naselaris, T., Holmes, E. A., & Kosslyn, S. M. (2015). Mental imagery: Functional mechanisms and clinical applications. *Trends in Cognitive Sciences*, *19*, 590-602.

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.

Perkins, A. M., Arnone, D., Smallwood, J., & Mobbs, D. (2015). Thinking too much: Self-generated thought as the engine of neuroticism. *Trends in Cognitive Sciences*, *19*, 492-498.

Pierce, A. H. (1901). *Studies in auditory and visual space perception*. New York: Longmans.

Plailly, J., Delon-Martin, C., & Royet, J. P. (2012). Experience induces functional reorganization in brain regions involved in odor imagery in perfumers. *Human Brain Mapping*, *33*, 224-234.

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, *80*, 1-24.

Ruby, F. J., Smallwood, J., Engen, H., & Singer, T. (2013). How self-generated thought shapes mood—the relation between mind-wandering and mood depends on the socio-temporal content of thoughts. *PLoS One*, *8*, e77554.

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

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

Scherer, K. R., & Zentner, M. R. (2001). Emotional effects of music: Production rules. In P. N. Juslin, J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 361–392). New York: Oxford University Press.

Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D., & Sayette, M. A. (2011). Meta-awareness, perceptual decoupling and the wandering mind. *Trends in Cognitive Sciences*, *15*, 319-326.

Schuster, C., Hilfiker, R., Amft, O., Scheidhauer, A., Andrews, B., Butler, J., . . . Ettlin, T. (2011). Best practice for motor imagery: A systematic literature review on motor imagery training elements in five different disciplines. *BMC Medicine*, *9*, 75.

Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., ... & Greicius, M. D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience*, *27*, 2349-2356.

Seli, P., Kane, M. J., Smallwood, J., Schacter, D. L., Maillet, D., Schooler, J. W., & Smilek, D. (2018). Mind-wandering as a natural kind: A family-resemblances view. *Trends in Cognitive Sciences*, *22*, 479-490.

Seli, P., Risko, E. F., & Smilek, D. (2016). On the necessity of distinguishing between unintentional and intentional mind wandering. *Psychological Science*, *27*, 685–691.

Smallwood, J., & Andrews-Hanna, J. (2013). Not all minds that wander are lost: The importance of a balanced perspective on the mind-wandering state. *Frontiers in Psychology*, *4*, 441.

Smallwood, J., & O'Connor, R. C. (2011). Imprisoned by the past: Unhappy moods lead to a retrospective bias to mind wandering. *Cognition & Emotion*, *25*, 1481-1490.

Smallwood, J., O'Connor, R. C., Sudbery, M. V., & Obonsawin, M. (2007). Mind-wandering and dysphoria. *Cognition & Emotion*, *21*, 816-842.

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, *132*, 946-958.

Smallwood, J., & Schooler, J. W. (2015). The science of mind wandering: Empirically navigating the stream of consciousness. *Annual Review of Psychology*, *66*, 487-518.

Sterelny, K. (1986). The imagery debate. *Philosophy of Science*, *53*, 560-583.

Stone, A. A., Schwartz, J. E., Neale, J. M., Shiffman, S., Marco, C. A., Hickcox, M., ... & Cruise, L. J. (1998). A comparison of coping assessed by ecological momentary assessment and retrospective recall. *Journal of Personality and Social Psychology*, *74*, 1670-1680.

Tartaglia, E. M., Bamert, L., Mast, F. W., & Herzog, M. H. (2009). Human perceptual learning by mental imagery. *Current Biology*, *19*, 2081-2085.

Taruffi, L., & Koelsch, S. (2014). The paradox of music-evoked sadness: An online survey. *PLoS One*, *9*, e110490.

Taruffi, L., & Koelsch, S. (2017). Implications of the Vienna Integrated Model of Art Perception for art-based interventions in clinical populations: Comment on" Move me, astonish me... delight my eyes and brain: The Vienna Integrated Model of top-down and bottom-up processes in Art Perception (VIMAP) and corresponding affective, evaluative, and neurophysiological correlates" by Matthew Pelowski et al. *Physics of Life Reviews*, *21*, 145-147.

Taruffi, L., Pehrs, C., Skouras, S., & Koelsch, S. (2017). Effects of sad and happy music on mind-wandering and the default mode network. *Scientific Reports*, *7***,** 14396.

Taruffi, L., Pehrs, C., Skouras, S., & Koelsch, S. (2018). *Sad Music, Empathy, and Visual Mental Imagery: An fMRI Study.* Paper presented at the International Conference on Music

Perception and Cognition (ICMPC) and Conference of the European Society for the Cognitive Sciences of Music (ESCOM), Graz, Austria, July 23-28, 2018.

Tye, M. (2000). *The imagery debate*. Cambridge, MA: Mit Press.

van den Hout, M. A., Rijkeboer, M. M., Engelhard, I. M., Klugkist, I., Hornsveld, H., Toffolo, M. J., & Cath, D. C. (2012). Tones inferior to eye movements in the EMDR treatment of PTSD. *Behaviour Research and Therapy*, *50*, 275-279.

Vincent, J. L., Kahn, I., Snyder, A. Z., Raichle, M. E., & Buckner, R. L. (2008). Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *Journal of Neurophysiology*, *100*, 3328-3342.

Vroegh, T. (2018). *Investigating the directional link between music-induced visual imagery and two qualitatively different types of emotional responses.* Poster presented at the KOSMOS Workshop "Mind Wandering and Visual Mental Imagery in Music", Berlin, Germany, May 16-19, 2018.

Vuoskoski, J. K., & Eerola, T. (2015). Extramusical information contributes to emotions induced by music. *Psychology of Music*, *43*, 262-274.

Vuoskoski, J. K., Thompson, W. F., McIlwain, D., & Eerola, T. (2012). Who enjoys listening to sad music and why? *Music Perception*, *29*, 311-317.

Wittgenstein, L. (1953). *Philosophical investigations.* New York: Macmillan.

Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition.* Stanford: Stanford University Press.

Zaki, J., & Ochsner, K. N. (2012). The neuroscience of empathy: Progress, pitfalls and promise. *Nature Neuroscience*, *15*, 675-680.

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion, 8*, 494–521.

Author Note

Some ideas presented in this article were discussed at the KOSMOS Workshop "Mind Wandering and Visual Mental Imagery in Music" at Humboldt University in Berlin (May 16-19, 2018). The authors thank Ilana Harris for proofreading the manuscript. *Correspondence concerning this article should be addressed to* Liila Taruffi, Department of Educational Sciences & Psychology, Freie Universität Berlin, Habelschwerdter Alle 45, 14195 Berlin, Germany. E-mail: liilataruffi@zedat.fu-berlin.de

Figure Caption

Figure 1. Conceptual organization of music-evoked visual mental imagery. Music-evoked visual mental images encompass a wide range of mental phenomena varying in their degree of automatic and deliberate constraints. Music-evoked mind-wandering or daydreaming (Martarelli et al., 2016; Taruffi et al., 2017) in the form of images overlap with spontaneous cognition and tend to have less deliberate constraints compared with musical daydreams (Herbert, 2011; 2018) and guided imagery (which often occurs in the context of music therapy practices such as GIM). MEAMs vary in their level of deliberate constraints. Images with higher level of automatic constraints are characterized as intrusive cognition/imagery and are typically found in a broad range of mental disorders (schizophrenia, post-traumatic stress disorder, and depression). The conceptual framework for music-evoked visual mental imagery presented here is conceived as a continuous space in which mental experiences shift in their level of mental constraints. Spontaneous cognition includes mind-wandering/daydreaming, creativity, and dreaming as described by Christoff et al. (2016). Spontaneous cognition can occur in the form of visual mental imagery or inner language (Delamillieure et al., 2010) and both modalities vary along this continuum. GIM = guided imagery and music; MEAMs = music-evoked autobiographical memories.