Int J. Body, Nature, Cult. Vol. 2, No. 2, 2023, pp. 151-166

https://doi.org/10.23124/JBNC.2023.2.2.151

A Simulation Theory of Musical Expressivity: An Expanded Version

Hye-yoon Chung*

Musicology Department, Korea National University of Arts, Seoul, Republic of Korea

Abstract *In the present paper, I critically examine Cochrane's "simulation theory of musical expressivity" and propose an expanded version as an enhanced and more comprehensive theory. Cochrane's theory, which is committed to the "low-level" simulation that concerns primitive feelings immediately aroused in listeners, quite successfully elucidates the mechanism underlying our recognition of musical expressivity at local and cognitively low levels. However, it fails to give an account of the mechanism underlying our recognition of musical expressivity for cognitively complex emotions. In addition, Cochrane's theory generally overlooks the dynamic interaction between music and listeners and the relational and emergent nature of musical expressivity. I argue that these problems can be solved by introducing the "high-level" simulation that concerns cognitively complex mental states. Ultimately, I argue that 'the expanded version of a simulation theory of musical expressivity' successfully overcomes not only the problems in Cochrane's theory but also those in other theories on musical expressivity providing a better account of the mechanism to which those theories fail to clarify or to give due attention.*

Keywords musical expressivity, musical persona, simulation, simulation theory of mindreading

1. Introduction

Recently, Cochrane proposed the "simulation theory of musical expressivity" (STME) in an attempt to elucidate the mechanism through which music is

* Author to whom correspondence should be addressed: Hye-yoon Chung
Email: hychung@mju.ac.kr

recognized as expressive of emotions(Cochrane, 'A Simulation Theory of Musical Expressivity').[1] His proposal is based on the view that the mechanism that makes it possible that "listeners to music can often make confident judgments regarding its expressive content" (191) is comparable to the mechanism through which we recognize the emotional states of others. What initially motivates this view is "the problem of how (purely instrumental) music could express emotions when it neither presents the situations that cause emotions nor has a body" (problem *X*) (193).[2] The strategy Cochrane adopts to solve this problem is to "compare the way we recognize emotions in music with the way we recognize emotions in other people."(193)[3] What justifies this comparison is the fact that both when we recognize emotions in others and when we recognize emotions in music, we usually do so only with perceptual information. For the comparison, Cochrane introduces a theory of mindreading that has been developed to explain how it is possible to know the mental states of others and proposes the STME endorsing a simulation theory of mindreading (STM). While insightful and seminal, Cochrane's theory has some crucial problems. In this paper, I will critically examine Cochrane's STME and propose an enhanced version of STME by clarifying and expanding it. Ultimately, I will argue that the STME I propose is enhanced not only compared to Cochrane's STME but also compared to other theories on musical expressivity.

2. Critical Reconstruction of Cochrane's STME

Presenting the arguments for his theory, Cochrane does not frame them in the format of STM. Without an account of the general way STM operates and of the specific way it works being applied to music, it is difficult to find coherence among his arguments and to understand exactly what each of his arguments is for. Accordingly, I will begin by restructuring Cochrane's STME framing his arguments in terms of STM.

While Cochrane calls his theory STME, he does not give any explanation of what simulation is. Supposedly, it is because today the term "simulation" is widely used. However, a minimal account is needed since the term is not used univocally. Goldman's definition is useful here since the simulation as adopted in Cochrane's STME is a part of the mindreading process. Goldman's definition demonstrates well what simulation is in this sense (Goldman, *Simulating Minds* 37-38). Goldman's definition of generic simulation is as follows.

> Process P simulates process P' = $_{df}$.
>
> (1) P duplicates, replicates, or resembles P' in some significant respects (significant relative to the purposes or function of the task), and
>
> (2) in its (significant) duplication of P', P fulfills one of its purposes or functions.

As a species of generic simulation, mental simulation is defined as follows.

> Mental Simulation: Process P is a mental simulation of target process P' = $_{df}$.
>
> Both P and P' are mental processes (though P' might be merely hypothetical), and P and P' exemplify the relationship of generic simulation as previously defined.

Now, what is important to our concern is that mental simulation is used to understand the mental states of others. This usage is established today as STM, a strong candidate for mindreading. Mindreading is to attribute a mental state to self or others (Gallese and Sinigaglia 512; Goldman, 'Mirroring, Simulating, and Mindreading', 235; Goldman, 'Mirroring, Mindreading, and Simulation' 312). STM is a theory that argues that it is a simulation that makes mindreading possible. What distinguishes STM from a competing theory such as theory theory of mindreading is its claim that mindreading must use one's own mental resources in an imitative way.[4] What is remarkable is that simulations involved in STM are not homogenous. Here, again, Goldman's distinction between low-level and high-level simulation is useful (Goldman, 'Two Routes to Empathy: Insights from Cognitive Neuroscience', 31-44; Goldman, 'Mirroring, Simulating, and Mindreading'; Goldman, *Simulating Minds*).[5] Low-level simulation occurs effortlessly through a "mirroring response" or immediate resonance. It is completely spontaneous and confined to observational modes such as visual or auditory sensation. In contrast, high-level simulation is an effortful, reconstructive process based on a conscious imagining of the target's perspective and situations. Notably, the imagination involved in high-level simulation is an "enactment imagination" in that it does not merely suppose the truth of a certain proposition but conjures up being in a certain mental state.[6] The characteristics of mindreading are determined by the level of simulation involved. While low-level mindreading is "comparatively simple, primitive, automatic, and largely below the level of consciousness" (Goldman, *Simulating Minds* 113), high-level

mindreading is cognitively more complex and potentially under intentional control. Face-based emotion recognition is representative of the former, while the recognition of intention, desire, judgment, or belief of others is typical of the latter.

Now, I will reconstruct Cochrane's STME in light of STM. In an attempt to give an account of the causal mechanism through which musical expressivity is recognized, Cochrane models his STME on STM for the recognition of *emotions* in others. Mindreading for emotions is usually thought to be low-level mindreading. However, it is not true that mindreading for emotions is necessarily categorized as low-level mindreading since there are cognitively complex emotions such as jealousy or hope that can be hardly recognized through purely perceptual information. It is usually from the knowledge beyond the perceptual level such as the knowledge of someone's personality and the situation the person is in that we know that the person feels jealous or hopeful. The fact that Cochrane proposes STME as a solution to the problem *X* might imply that his STME necessarily precludes cognitively complex emotions.[7] Cochrane's STME is exclusively committed to primitive feelings. Following this, I will reconstruct Cochrane's STME in the frame of STM involving low-level simulation. However, it should be noted that Cochrane's STME is not the only way to solve the problem *X*. As a matter of fact, it could be solved by showing that it is not a problem at all. I will elaborate on this later.

According to Cochrane, what makes simulation possible is the belief that music is a product of some intentional activity or particular mental state. This belief makes us treat music as we do other people and this naturally generates simulation for music. In this view, Cochrane introduces "persona", an imaginative being as a subject of occurrent emotions experienced in music.[8] Then, in Cochrane's STME the mental states of the target would be the occurrent emotions that the persona is imagined to be undergoing, and listeners would recognize musical expressivity by undergoing the mental processes themselves similar to those imagined to be occurring in the persona and attributing them to the persona. Cochrane assigns significant parts of his discussion to an account of the mechanism underlying musical arousal and now it becomes clear why he does this. For simulation to occur, listeners themselves should undergo the mental processes of a persona.

With this clarification, it is notable that Cochrane's STME is different from the normal form of STM in an important aspect. The resemblance between the mental

processes of the simulator and those of the target is absent in Cochane's STME since there exists only the former without the latter. The latter is completely dependent on and therefore reduced to the former. Cochranes's statement that "being […] aroused by an emotion that *mirrors* the one expressed by the music will allow us to perceive the music as possessing that emotional quality" (196) is circular, since in his theory the emotion expressed by the music is identified with the one aroused in listeners. Moreover, it gives the impression that musical expressivity should be something already given in music waiting to be recognized by listeners.

Now, it might seem that Cochrane's theory is a version of *the* arousal theory[9] since it identifies the emotions expressed by music with those aroused in listeners. However, Cochrane himself draws a clear line between his theory and *the* arousal theory claiming that "the simulation process is initiated by replicating certain attributes of the target and is geared throughout towards the goal of attributing properties to that target by a […] process of projection." ('Music, Emotions, and the Influence of Cognitive Sciences' 981). The problem is that the example Cochrane himself provides, "the visceral sickly feeling that is so well captured by the quiet unsynchronized glissandi of a violin section" ('A Simulation Theory of Musical Expressivity' 201) fails to satisfy the condition for simulation. According to Cochrane, the sickly feeling should be immediately attributed to music and music should be recognized to express it. However, in this case, we are not likely to say that music expresses the sickly feeling, just saying that music arouses it instead. It is the same that we would not say that the 2nd movement of Haydn's *Surprise* Symphony expresses surprise, but only say that it arouses surprise. In these cases, the attribution of mental states to the target does not occur. Cochrane claims that when feelings are aroused in us by the emotional impression of facial expressions or gestures of others, we experience this arousal immediately as belonging not to us, but to the other. He argues that this kind of projection occurs also for music leading to "a direct impression of the emotion *in* the music rather than a self-conscious impression of personal arousal." (195-196) However, it is not always true that the feeling aroused in listeners is recognized as "the emotion *in* the music" as the examples above demonstrate. Cochrane claims that conscious attention to music will suffice for projective perception (205). However, some feelings are not attributed to music, though aroused through the attention to music. For instance, although the feeling of relief aroused when a prolonged suspension is finally resolved occurs through the attention to music, no one will say that the

music expresses the feeling of relief, but will say instead that the music arouses it. In sum, Cochrane's STME cannot distinguish the feelings aroused by and attributed to music from those aroused by but not attributed to music. I do not think this is a fatal problem. I believe that Cochrane's STME would be strengthened by admitting that not all the feelings aroused in listeners can be encompassed in low-level simulation. In many cases, e.g., the feeling of solidity that a sequence of octaves arouses or the feeling of warmth that the timber of a horn arouses, is indeed well captured by Cochrane's STME. For these cases at least, Cochrane's STME is successfully distinguished from *the* arousal theory.[10]

In Cochrane's STME, we can find another argument that supports the distinction. Cochrane argues that music arouses emotions in listeners through the resemblance between musical structure and the outward display of emotions such as vocal utterances, expressive physical gestures, and dynamic qualities of emotions. This implies that in Cochrane's STME, feelings aroused in listeners are supposed to be connected to music reliably. In the case of *the* arousal theory, the connection between aroused feelings and music is not warranted.[11] The resemblance between musical structure and the appearances characteristic of emotions is exactly what *appearance emotionalism* appeals to in its attempt to overcome *the* arousal theory. What distinguishes Cochrane's STME from appearance emotionalism is its claim that such outward "resemblance must connect to feelings." (200)[12] I argue that it should be highly appreciated that while Cochrane's STME emphasizes the intimate connection between music and emotion, it still preserves our ingrained insight that music arouses feelings in listeners.[13]

It should also be highly appreciated that Cochrane's STME gives an excellent account of *how* the resemblance mentioned above is achieved across different modalities. This account constitutes a part of his answer to the problem *X*. Cochrane explains that the multimodality of our perception translates the auditory information music provides into the experience of movement which would lead to bodily changes and emotional feelings in listeners (199-200). He supports this idea with evidence showing that various forms of sensory information are transformed into a sense of movement. In low-level simulation, a simulation process is automatically prompted by perceptual information of others and this typically occurs through a mirroring process.[14] Cochrane's STME appeals to the mirroring process, too. In Cochrane's STME, it is mirroring for *actions* that enable listeners to simulate sensory-motor information perceived in music. In this

way, the purely perceptual information music provides affords the experience of emotions. Many philosophers have asserted the idea that listeners experience movement in music and that this might be the fundamental ground of musical expressivity though usually without any clear evidence.[15] Cochrane's STME is very significant in that it provides a plausible account of the mechanism with concrete evidence.

So far, I have critically examined Cochrane's STME. Besides the points mentioned above, there are several other points to be considered. First, while Cochrane's STME provides a very persuasive account of the musical expressivity for primitive feelings, it leaves a broad area of musical expressivity for cognitively complex emotions unexplained. Second, it is limited in that it adopts only the mirroring process for *actions* as a mechanism for musical simulation without considering the mirroring process for *emotions* and *sensations*.[16] Third, it presents listeners as passive beings who only respond to music in a primitive way ignoring the active role that listeners play. I believe most of the problems in Cochrane's STME could be successfully solved with the introduction of high-level simulation. In the next section, I will demonstrate how this could be done.

3. An Expanded Version of the STME

The apparent problem that emerges when we attempt to introduce high-level simulation in music is how music can provide the information required for cognitively complex mental states. The answer I propose is that as the perceptual information presented by music is accumulated through the unfolding of music, it can be transformed into the cognitive resources necessary for cognitively higher mental states. As Karl and Robinson's analysis of Shostakovich's 10th Symphony (Karl and Robinson, 'Shostakovich's Tenth Symphony and the Musical Expression of Cognitively Complex Emotions' 401-415) and Levinson's analysis of Mendelssohn's *Hebrides* Overture (Levinson, 'Hope in *The Hebrides*' 336-375) show, music can provide cognitive resources even without titles or programs. It is possible through listeners' imaginings and musical narratives constructed on these.[17] While listening, listeners not only respond to music in a primitive way but also interact dynamically with the music by constantly recognizing and structuralizing it. Listeners interpret this experience continuously ascribing meaning to it and through this interpretation musical narrative is constructed as a way of understanding. During this process, cognitive resources needed for

cognitively complex emotions such as intention, desire, and belief are formed, accumulated, and reinterpreted.

It seems obvious that simulation is involved in this process since listeners' mental processes are necessarily used in the formation of cognitive resources. As for the attribution, some might be doubtful, since cognitive resources emerge only from listeners' experience. Here again, as in Cochrane's STME, the mental states of the target are dependent on those of the stimulator. However, in contrast to Cochrane's STME, the mental states of the target are not reduced to those of the simulator. In Cochrane's STME, music has only a causal role and does not contribute to the contents of the mental states of the target. This makes it hard to distance music from listeners as *the other*. As mentioned above, a host of primitive feelings immediately aroused in listeners fail to be attributed to music and remain just as one's own feelings.[18] In high-level STME, the mental states of the target emerge as a result of the dynamic interaction between music and listeners and thereby become *about* music. In this process, the contribution music makes goes beyond the causal level. Additionally, listeners' efforts to understand and interpret their interaction with music as meaningful to themselves are likely to put some distance between music and themselves naturally, whereby music becomes *the other*.[19] Through this process, the mental processes emerging from the interaction between music and listeners are separated from listeners and attributed to music thereby completing the STME.

Certainly, in high-level STME, the resemblance between the mental processes of the simulator and those of the target is not established in a strict sense, since the latter is not completely independent of the former.[20] Some might claim that this forces us to deny STM for music. In my view, however, it would be more productive to accept high-level STME as a slightly deviated version of STM rather than to refute it.[21] Once we accept high-level STME as a version of STM, it would be highly appropriate precisely because of its deviation. In high-level STME, musical expressivity is understood as emerging and being attributed to music through the dynamic interaction between music and listeners, not as something already given in the structure of music independently of listeners. Since the 20th century, many discourses in the philosophy of music have been committed to discovering the *locus* of musical expressivity, while its relational and emergent nature has been ignored.[22] In this respect, the STME as I propose is significant.

The introduction of high-level STME does not mean discarding low-level STME. The STME I propose, not replacing but expanding Cochrane's STME, encompasses low-level STME. Moreover, low-level STME and high-level STME are not mutually exclusive. On the contrary, low-level STME plays a pivotal role in the formation of the cognitive resources that make high-level STME possible. The local expressivity of music can be appropriately accounted for in terms of low-level STME and the musical expressivity at this primitive, micro-level is utilized in the process of the formation of the cognitive resources needed for high-level STME. Listeners construct a musical narrative based on the musical expressivity recognized through low-level STME along with the structural information music provides as it unfolds. Through this process, the cognitive resources necessary for high-level STME are constructed. In this way, high-level STME is grounded in low-level STME.[23] In addition, high-level STME embraces the feelings aroused by, but not attributed to music. Along with the musical expressivity recognized through low-level STME, those feelings are used when listeners attempt to make meaning of their experience as the music unfolds. This contributes to the formation of cognitive resources needed for cognitively complex emotions.

4. Conclusion

The STME I support is expanded in that it encompasses high-level as well as low-level STME providing an account of musical expressivity not only for primitive feelings but also for cognitively complex emotions. It is expanded also in that it embraces the feelings aroused by but not attributed to music. The expanded nature of the STME I propose can be found in another respect too. As mentioned above, what motivated Cochrane's STME initially is the problem *X*. Cochrane concedes that the situation that *X* identifies is a problem and proposes his own STME as a solution for it. In contrast, I solve *X* in an expanded way arguing that while there are cases where *X* is indeed a problem, there are other cases where it is not. For the cases where music can afford cognitively higher resources, *X* is solved simply by being denied.

These expanded aspects of the STME I propose are also the aspects that are enhanced compared to Cochrane's STME. Another enhanced aspect mentioned above is that the STME I propose does justice to listeners' active role and represents the relational and emergent nature of musical expressivity well. Notably, this is the aspect enhanced compared not only to Cochrane's STME but

also to other theories on musical expressivity. As mentioned above, many theories on musical expressivity have overlooked the relational and emergent quality of musical expressivity committing itself to finding its *locus*. This directly points to my final argument that the STME I propose is an enhanced theory compared not only to Cochrane's STME but also to other theories on musical expressivity. The STME I suggest, embraces the strong points of Cochrane's STME and overcomes its problems. As such, it could be especially valuable in several respects. First, it pays due attention to the role the arousal of emotions plays without facing the trouble that plagues *the* arousal theory and validly captures the embodied nature of our experience of musical expressivity. Second, it provides an account of the mechanism underlying listeners' experience of movement in music and the process through which it is transformed into emotional experience and then into the recognition of musical expressivity. Third, it represents the dynamic interaction between music and listeners and the role it plays in listeners' recognition of musical expressivity. Fourth, through this dynamic interaction, it properly presents the relational and emergent quality of musical expressivity. Fifth, it successfully solves the problem *X* from a broader view that many philosophers from the analytic tradition have considered most puzzling. However, it certainly is not perfect as it is. As shown above, the mirroring process adopted in STME is too limited and should be expanded to include mirroring for emotions and sensations. Moreover, it needs a stronger account for why some feelings are attributed to and aroused by music, while others are just aroused by, but not attributed to music. Nevertheless, I believe these five points are enough to underpin the STME I propose as an enhanced theory of musical expressivity.

Notes

[1] I follow Cochrane's terminology using 'expressivity', but with clear intention to avoid the controversy around the terms 'expression' and 'expressiveness'. In the current aesthetics scene, 'expressiveness' is used deliberately to refer to the perceptual property of artworks and sharply distinguished from 'expression' which needs agency. 'Expressivity' is not sensitive to the distinction between 'expressiveness' and 'expression'. I use 'expressivity' since my argument in this paper does not necessitate such distinction.

[2] It is remarkable that the question Cochrane himself formulates does not correctly represent his epistemological concern. The appropriate form would be 'how (purely instrumental) music could be possibly recognized as expressing emotions when it neither presents the situations that cause emotions, nor has a body.' Cochrane's discussion is focused on purely instrumental music. So will my discussion, since it is purely instrumental music which generates philosophical puzzle.

[3] This strategy is not new, though. Cognitivists or appearance emotionalists recognized problem *X* earlier and attempted to solve it by appealing to the way we recognize emotions in others. See Davies, *Musical Meaning and Expression*; Kvy, *Sound and Sentiment: An Essay on the Musical Emotions, Including the Complete Text of the Corded Shell, The Arts and Their Philosophies*).

[4] For the account of simulation theory of mindreading and theory theory of mindreading, see Goldman, *Simulating Minds*; Heal, *Mind, Reason and Imagination: Selected Essays in Philosophy of Mind and Language*; Stueber, *Rediscovering Empathy: Agency, Folk Psychology, and the Human Sciences*.

[5] This distinction is most explicitly articulated in Goldman's theory, but also supported by many theories of empathy which are based on simulation process. See Stueber, *Rediscovering Empathy*; Zahavi, 'Basic Empathy and Complex Empathy'.

[6] For more detailed account, see Goldman, *Simulating Minds* 47-48.

[7] I object to the idea that the recognition of musical expressivity based on purely perceptual information is necessarily confined to low-level STM. Music can provide context needed for the recognition of cognitively complex emotions through the processes in which the auditory information presented by music is accumulated and reconstructed. I believe that high-level STM could occur through this process. In other words, music can afford high-level STM by providing music-structural situation and context. This will be discussed in more detail later on.

[8] Cochrane's STME is thereby connected to *persona theory* which has been supported by music theorists and philosophers from analytic tradition since the late 20th century. His view that "if the music expresses the actual feeling of an emotion rather than merely its appearance, then there should be someone to whom that emotion belongs" (Cochrane, 'A Simulation Theory of Musical Expressivity' 203) echoes the argument of the robust form of *persona theory*. We can find the most robust form of *persona theory* in Levinson's arguments. See Levinson, 'Musical Expressiveness as Hearability-as-Expression' 91-108; Levinson, 'Musical Expressiveness', 90-125.

[9] The position of *the* arousal theory can be formulated in the statement that the ground for the description that a certain kind of music, *M* expresses a certain kind of emotion, *E* is that *M* arouses *E* in listeners.

[10] Certainly, it should be clarified at a more fundamental level why while some feelings are attributed to as well as aroused by music, other feelings are aroused by but not attributed to music.

[11] Due to this, *the* arousal theory has been thought to make musical expressivity arbitrary and unreliable.

[12] So, the resemblance between the structure of music and the appearances characteristic of emotions should be understood at an overall embodied level in a comprehensive way involving sensory-motor, somatosensory, and visceral sensations.

[13] Notably, since cognitivists or appearance emotionalists expelled felt emotions from discourse on musical expressivity, many philosophers have attempted to reinstate them with strong insight that felt emotions should be involved in our experience of musical expressivity. For such attempts, see Matravers, 'Expression in Music'; Ridley, *Music, Value, and the Passions*, 1995, Robinson, *Deeper than Reason: Emotion and its Role in Literature, Music, and Art*, 2005.

[14] Mirroring response is a process in a simulator based on 'mirror neurons' which were first recorded in *macaques* as discharging both when they perform an action themselves and when they just observe a similar action done by other *macaques* without performing it themselves. Today, the view that there is a mirror system which plays the same role in human beings is gaining more and more support. Furthermore, studies show that there are mirroring response not only for actions but also for emotions and for sensations. For mirror neurons, see Gallese, Fadiga, Fogassi, and Rizzolatti, 'Action Recognition in the Premotor Cortex'; Pellegrino, Fadiga, Fogassi, Gallese, and Rizzolatti, 'Understanding Motor Events: A Neurophysiological Study', 176-180; Kohler, Keysers, Umiltà, Fogassi, Gallese, and Rizzolatti, 'Hearing Sounds, Understanding Actions: Action Representation in Mirror Neurons'. By the way, it is not true that low-level STM necessarily occurs through mirroring processes. For other mechanisms through which low-level STM is operated, see Goldman, *Simulating Minds*, 124-129.

[15] For instance, while cognitivists or appearance emotionalists claim that musical expressiveness is established on the resemblance between the structure of music and the appearances characteristic of emotions, they don't provide any detailed account of what makes such resemblance possible but simply call it as *metaphorical* resemblance in that it is held across different modalities. See Kivy, *Sound and Sentiment*, 54-56. The rare attempts to explain the mechanism through which listeners experience movement in music are found in Nussbaum and Robinson. Nussbaum, *The Musical Representation: Meaning, Ontology, and Emotion*; Robinson, *Deeper than Reason*.

[16] For mirroring response for emotions and sensations, see Bastiaansen, Thioux, and Keysers, 'Evidence for Mirror Systems in Emotions', 2391-2404.

[17] Imaginings encompass from those involved in our recognition of 'ascending' and 'descending' melodic lines or 'arrival' to tonic which are immediate and unconscious to those involved in our cognitive understanding and interpretation which are conscious and reflective. These imaginings are *embodied* in its nature in that they are all founded on listeners' bodily engagement. Among them, it is conscious imagining which is considered to be at the core of high-level simulation.

[18] For the excellent account of listeners' emotional experience as intimately entangled with music, see Walton, 'Projectivism, Empathy, and Musical Tension'; Walton, 'Listening with Imagination: Is Music Representational?'.

[19] In this argument, I'm not necessarily committed to *persona theory*. However, my argument doesn't necessarily exclude musical persona. Cone, who introduced musical persona into music analysis for the first time, defines persona as a being who subsists only in virtue of artwork (Cone, *The Composer's Voice*, 160). In the case of music, musical persona exists only in the structural unfolding of music and therefore the mental states or processes attributed to musical persona will be attributed to music all the time. So, persona theory is not incompatible with my view.

[20] The ontological status of musical persona is similar to that of fictional character in novels in a sense. First, both are not real, but fictional. Second, both have their own history and context even though those of musical persona are relatively less consistent. However, in a stricter sense, they

are different. First, while the identity of a fictional character is determinate and stable in paradigm cases, that of a musical persona is not determinate nor consistent. Second, while the mental states or processes of a fictional character can be directly described independently of readers' responses to them, those of musical personas cannot. It is in this sense of differences that the mental states or processes of musical persona are not completely independent of those of listeners.

[21] The resemblance between mental states or processes may be an outcome of the fact that the mental resources of the simulator are reused successfully. Both 'resemblance' and 'reuse' have been considered to be crucial in simulation. Some emphasize the former and others the latter. For a discussion, see Gallese and Sinigaglia, 'What Is So Special about Embodied Simulation?'.

[22] The most salient case is appearance emotionalism and *the* arousal theory is another case. In appearance emotionalism, musical expressiveness is considered consisting in the structure of music as its perceptual property. In contrast, *the* arousal theory finds the locus of musical expression in listeners' emotional responses.

[23] This intimate relation between low-level and high-level simulation might be regarded as a distinct feature of simulation occurring especially in the musical experience.

References

Bastiaansen, Jojanneke A.C.J., Marc Thioux, and Christian Keysers. "Evidence for Mirror Systems in Emotions." *Philosophical Transactions of the Royal Society B,* vol. 364, 2009, pp. 2391-2404. https://doi.org/10.1098/rstb.2009.0058

Cochrane, Tom. "A Simulation Theory of Musical Expressivity." Australian Journal of Philosophy, vol. 88, no. 2, 2009, pp. 191-207. https://doi.org/10.1080/00048400902941257

Cochrane, Tom. "Music, Emotions, and the Influence of Cognitive Sciences." *Philosophy Compass*, vol. 5, no. 11, 2010, pp. 978-988. https://doi.org/10.1111/j.1747-9991.2010.00337.x

Cone, Edward T. *The Composer's Voice.* California University Press, 1974.

Davies, Stephen. *Musical Meaning and Expression.* Cornell University Press, 1994.

Di Pellegrino, G., et al. "Understanding motor events: A neurophysiological study." *Experimental Brain Research,* vol. 91, no. 1, 1992, pp. 176–180, https://doi.org/10.1007/bf00230027

Gallese, Vittorio, et al. "Action Recognition in the Premotor Cortex." *Brain*, vol. 119, no. 2, 1996, pp. 593-609. https://doi.org/10.1093/brain/119.2.593

Gallese, Vittorio, and Corrado Sinigaglia. "What Is So Special about Embodied Simulation?" Trends in Cognitive Sciences, vol. 15, no. 11, 2011, pp. 512-519. https://doi.org/10.1016/j.tics.2011.09.003

Goldman, Alvin I. *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading.* Oxford University Press, 2006.

_____. "Mirroring, Simulating, and Mindreading." *Mind & Language*, vol. 24, no. 2, 2009, pp. 235-252. https://doi.org/10.1111/j.1468-0017.2008.01361.x

_____. "Two Routes to Empathy: Insights from Cognitive Neuroscience." *Empathy: Philosophical and Psychological Perspectives*, edited by Amy Coplan and Peter Goldie, Oxford University Press, 2011, pp. 31-44.

Heal, Jane. *Mind, Reason, and Imagination: Selected Essays in Philosophy of Mind and Language.* Cambridge University Press, 2003.

Karl, Gregory, and Jenefer Robinson. "Shostakovich's Tenth Symphony and the Musical Expression of Cognitively Complex Emotions." *The Journal of Aesthetics and Art Criticism,* vol. 53, no. 4, 1995, pp. 401-415. https://doi.org/10.2307/430975

Kivy, Peter. *Sound and Sentiment: An Essay on the Musical Emotions, Including the Complete Text of the Corded Shell, The Arts and Their Philosophies.* Temple University Press, 1989.

Kohler, Evelyne, et al. "Hearing Sounds, Understanding Actions: Action Representation in Mirror Neurons." *Science,* vol. 29, 2002, pp. 846-848. https://doi.org/10.1126/science.1070311

Levinson, Jerrold. "Hope in The Hebrides." *Music, Art, and Metaphysics: Essays in Philosophical Aesthetics,* Cornell University Press, 1990, pp. 336-375.

_____. "Musical Expressiveness." *The Pleasures of Aesthetics: Philosophical Essays,* Cornell University Press, 1996, pp. 90-125.

_____. "Musical Expressiveness as Hearability-as-Expression." *Contemplating Art: Essays in Aesthetics*, Oxford University Press, 2006, pp. 91-108.

Matravers, Derek. "Expression in Music." *Philosophers on Music: Experience, Meaning, and Work*, edited by Kathleen Stock, Oxford University Press, 2007, pp. 95-116.

Nussbaum, Charles. *The Musical Representation: Meaning, Ontology, and Emotion*. The MIT Press, 2007.

Ridley, Aaron. *Music, Value, and the Passions*. Cornell University Press, 1995.

Robinson, Jenefer. *Deeper than Reason: Emotion and its Role in Literature, Music, and Art*. Oxford University Press, 2005.

Stueber, Karsten. *Rediscovering Empathy: Agency, Folk Psychology, and the Human Sciences.* The MIT Press, 2006.

Walton, Kendall L. "Projectivism, Empathy, and Musical Tension." *In Other Shoes: Music, Metaphor, Empathy, Existence.* Oxford University Press, 2015, pp. 118-150.

Walton, Kendall L. "Listening with Imagination: Is Music Representational?" *The Journal of Aesthetics and Art Criticism,* vol. 52, no. 1, 1994, pp. 47-61. https://doi.org/10.2307/431584

Zahavi, Dan. "Basic Empathy and Complex Empathy." *Emotion Review*, vol. 4, no. 1, 2012, pp. 81-82. https://doi.org/10.1177/1754073911421387