

This article was downloaded by: [Cochrane, Tom]

On: 25 May 2010

Access details: Access Details: [subscription number 922472718]

Publisher Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Australasian Journal of Philosophy

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713659165>

### A Simulation Theory of Musical Expressivity

Tom Cochrane <sup>a</sup>

<sup>a</sup> Swiss Center for Affective Sciences (CISA),

First published on: 18 June 2009

**To cite this Article** Cochrane, Tom(2010) 'A Simulation Theory of Musical Expressivity', Australasian Journal of Philosophy, 88: 2, 191 – 207, First published on: 18 June 2009 (iFirst)

**To link to this Article:** DOI: 10.1080/00048400902941257

**URL:** <http://dx.doi.org/10.1080/00048400902941257>

## PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

# A SIMULATION THEORY OF MUSICAL EXPRESSIVITY

Tom Cochrane

This paper examines the causal basis of our ability to attribute emotions to music, developing and synthesizing the existing arousal, resemblance and persona theories of musical expressivity to do so. The principal claim is that music hijacks the simulation mechanism of the brain, a mechanism which has evolved to detect one's own and other people's emotions.

## 1. Introduction

Listeners to music can often make confident judgments regarding its expressive content. Yet they may well be unable to articulate how the music has this effect. Thus when explaining how it is that music is expressive of emotions, one must carefully distinguish the experience that allows one to form judgments of expressive content from the causal basis of this experience. Several existing theories have something to say about both questions, and confusions can often arise regarding their compatibility as a result. In particular, I will argue that the principal alternative theories—the arousal, resemblance and persona theories—can all be synthesized with regard to a detailed picture of the causal mechanisms involved. This synthesis can moreover point us towards the account of the experience best supported by the causal story. Thus the goal of this paper is not to analyse in detail the experience of musical expressivity—for instance, the ways in which this experience subjectively varies, or how distinctive and fleshed out the perceived emotion seems to be. The goal is rather to present a coherent and plausible causal story that takes us up to the point at which the experience emerges.

Despite this focus, however, several arguments that I make about the causal mechanisms involved are partially justified by basic and fairly uncontroversial facts about the experience that they generate. We should, after all, be guided by the experience that the causal story is intended to support. In particular, we should note that, whilst it is certainly possible to judge a work as sad or happy in a purely conventional or inferred way, the central experience of musical expressivity is to have an immediate impression, sometimes of an extremely vivid nature, of an occurrent emotional state. We seem to hear in the music an emotion that is temporally

developing, that modulates in quality and varies in intensity in the ways that everyday emotions do.

Thus two basic features of the experience of musical expressivity are assumed here: first, that we have an impression of an occurrent, temporally developing emotional state, and second, that this emotional state seems somehow to be a property of the music, i.e. our experience is focused on the qualities of the music. Cases where tunes judged to be say, cheerful, owing to their similarity to tunes conventionally regarded as cheerful, are then taken ultimately to be parasitic upon experiences of occurrent emotions in music of the kind I have outlined.

An additional grounding feature of this account is its assumption of a theory of emotions that takes patterns of bodily changes to be essential. Several other accounts of expressivity, it seems to me, are needlessly encumbered by theories in which conceptual judgments are taken as central to emotional states, where bodily changes and the feelings of those changes are relegated to a peripheral role at best. Yet it is much easier to account for the strength and immediacy of musical expression when you regard feelings rather than judgments as central to emotions.

In particular I adopt a perceptual theory of emotions of the kind presented by Prinz [2004] and Deonna [2006]. These theories recognize that bodily changes are necessary for emotions. Yet they also recognize that the function of an emotion like fear is to represent danger, allowing the subject to avoid or manage that danger. Thus perceptual theories further claim that bodily changes can represent the meaningful, intentional content of the emotion. Bodily changes such as increased heart rate, or even overt behaviour, like running away, can represent danger because they are set up by evolution or learning to be causally initiated by dangerous situations. This allows them to reliably track the presence of danger for the subject.

Different perceptual theories then make sense of how bodily feelings could really be *about* danger in different ways. On the view I prefer, the intentional object (danger) should only be identified with features of the external environment in so far as those features are felt to physically impact on the bodily organism. The situation, say a bull charging towards me, is less the direct cause and content of my fear than is the imagined impact of the bull upon my body, which triggers bodily responses to flee or resist. By triggering these reactions, the charging bull *highlights* the danger, my bodily vulnerability, but the sense of impact and my bodily reaction to this impact more directly form the content of my representation of danger. So the view is that emotions perceive the status of the organism by means of bodily feelings and reactions.<sup>1</sup> This, I think, makes sense of the claim that bodily feelings are doing the perceptual work.

Now on this view, bodily changes include changes to the respiratory system, circulatory system, digestive system, musculoskeletal system, the endocrine system, and the release of neuro-transmitters (such as serotonin). Any and all bodily changes that generate feelings can potentially be

<sup>1</sup>See Cochrane [unpublished] for a detailed exposition and defence of this view. Note that the perception of the status of the organism is necessary and sufficient for the emotion, though a conceptual appraisal of the situation that triggers the bodily reaction may be considered part of the complete emotional state.

incorporated within the emotion because one's overall bodily feeling is causally as well as experientially associated with the status of the bodily organism. So, for instance, facial and vocal expressions, despite their probable evolution for communicative purposes, also partly constitute emotional states because they are bodily changes that feel a certain way. It is thus unsurprising that by regulating facial and vocal expressions one can regulate emotions, affecting both subjective self-reports and other signs of physiological arousal (such as skin-conductance) which are presumably associated at a sub-personal level with those behavioural expressions.<sup>2</sup> On this view we might also admit all kinds of tactile sensations into the emotional state. Feelings such as pressure, movement and heat are typical features of emotional experience. Hence even things like the feel of a cold winter's day or a fur coat could potentially be incorporated within the emotional state because they directly affect the sense of impact on one's body and could thereby modify the experience of an emotion.

Apart from the bodily changes which constitute emotions, we must also include as their realizing base the neural functions responsible for initiating and controlling these changes. In particular, according to neurologists such as Antonio Damasio [2000; 2004], the brain continuously maps the overall state of the body. This serves several functions. First, it provides centralized regulation for the suite of emotional responses. Second, it is most likely the immediate basis of the conscious feeling of the emotion. And third, it enables bodily changes (and the feeling of those changes) to be anticipated or simulated in the absence of actual physiological changes. This mechanism, which Damasio calls the 'as-if loop', enables subjects to react more quickly than we might expect some of the changes (such as hormonal changes) to occur, as well as to map the temporal development of one's bodily status. As I shall detail below, it is also heavily implicated in our capacity to empathize with others.

Given this picture of emotions, we are now left with the problem of how (purely instrumental) music could possibly express emotions when it neither presents the situations that cause emotions, nor has a body. The solution that several theorists have come up with is to compare the way we recognize emotions in music with the way we recognize emotions in *other* people. The general reasoning behind this is that we are able to recognize the emotion of another without being aware of what causes it. We are also able to make such recognitions based purely on audible or visual information. The various theories of expression then elaborate this comparison in different ways. With regard to the causal story, I think they all contribute a part of the picture, which in combination with recent empirical data we can now piece together. Accordingly, this presentation must range across a variety of different theories and the overall picture may thereby appear rather complex. Yet we can summarize the theory by its appeal to our capacity for empathy, in particular, the kind of empathy whereby we recognize someone's emotion by perceiving their expressive behaviour. In a manner most compatible with the above view of emotions, this kind of empathy is

<sup>2</sup>See Laird [2007] for an extensive review.

grounded in the simulation of bodily patterns. For this reason I've labelled it a simulation theory of musical expressivity.<sup>3</sup>

## 2. The Necessity of Arousal

The first piece of our causal story is provided by the arousal theory of musical expressivity, a sophisticated version of which is offered by Matravers [1990]. According to this theory, the way to overcome the problem that music presents an occurrent emotion without presenting a person who possesses that emotion is simply to find a real person to whom we can attach it. In particular, arousal theory suggests that the listener is able to apply an emotional label to the music on the basis of his or her own emotional reaction. Then, in order to preserve the fact that our response is still guided by and focused on the music, Matravers makes an analogy with our reactions to emotionally expressive people. In cases where we perceive that another person is undergoing an emotion, he claims that we typically mirror that person's emotion or feel an emotion complementary to it. The same is then said to be true of expressive music.

I will not address here the criticisms that arousal is insufficient to properly discern the expressive qualities of music. The most common criticism is to attack the necessity of arousal for judging the expressive quality of music. Many critics complain that they are perfectly capable of recognising the emotional quality of a piece of music without actually being aroused by that emotion, or any other emotion complementary to it. Certainly we might admit that people *can* be aroused by music, but that is far from saying that it is *necessary* for us to be aroused. It does not even seem to be the paradigm case of musical expressivity.

Matravers responds by saying that this criticism assumes too strong an idea of arousal, as if the feeling must overwhelm the listener; instead, merely the barest beginnings of feeling are sufficient for arousal. I think this is right, but the answer needs refining. In order to recognize an emotion in another person, and similarly in music, it is necessary to be aroused by either attenuated bodily changes or, which is more primary, a neural simulation of such changes.

Supporting this conclusion is empirical evidence that when people become unable to experience a particular emotion, they develop a corresponding inability to recognize that emotion in others. Antonio Damasio describes the case of 'S' who, owing to damage to her amygdala, simultaneously lost the ability to feel fear and recognize the expression of fear in others [2000: 62–5]. A similar inability to feel disgust and recognize it in others has also been observed in patients with Huntington's disease [Goldman & Sripada 2005]. More evidence comes from studies of subjects suffering from Parkinson's

<sup>3</sup>Note that simulation theory has been applied to artistic expression before, most notably by Currie and Ravenscroft [2002] and Currie [2004]. Yet the kind of simulation that Currie and Ravenscroft appeal to is based on adopting pretend beliefs and desires, which, when processed by one's own cognitive machinery, output certain emotional states that one can then attribute to the target. This is the kind of empathy that one may engage in when the other person is absent, and is highly suitable for literary presentations.

disease, a disorder that affects the brain's production of the neurotransmitter dopamine. Since levels of dopamine rise when engaging in aggressive behaviour, this neurotransmitter has been linked with the capacity for feeling anger. As such it has been noted that when Parkinson's patients stop taking L-Dopa, the drug that enables dopamine production to be restored, they correspondingly lose the ability both to be angry as well as to recognize that emotion in others [Lawrence et al. 2007]. Being able to selectively turn on and off both the arousal and recognition of anger by regulating one specific neurotransmitter entails that it is a necessary pre-requisite of both phenomena, and hence that they share neural processes.

Furthermore, I agree with Matravers that the arousal involved need only be of a highly attenuated nature. There might only be the beginnings of muscle tension, expressive movements, or other somatic changes. More importantly, we can appeal to Damasio's as-if loop. It is only necessary for the brain to recreate or simulate the bodily changes or expressive behaviour involved in an emotional state to trigger the feeling of those changes. Given the function of this mechanism to regulate our physiological responses, such simulation may or may not then go on to trigger the associated actual changes.

Still the critic may complain that even an attenuated or simulated bodily response should generate feeling and he just doesn't seem to have that feeling. However, this ignores the phenomenology of empathic states. When we recognize (at least strong) emotions in others, we do not need to infer from their expressions to any 'hidden' inner feelings. Rather we get a direct impression of the feeling *in* their face or body; we seem to perceive the emotion itself externalized. So the idea is that when recreating the aroused response of another, this arousal is usually experienced not as something we feel in ourselves, but as something belonging to the other person.

This kind of 'projective' perception requires some elaboration. First of all, we note that the observer's arousal is geared towards tracking the emotion of the other person. It is perfectly intelligible by causal theories of representation, such as Fred Dretske's [1981; 1986], that the perceptual state should represent the original cause, the other person, as having the emotion, rather than any of the intermediate stages involved, i.e. my various physiological arousal mechanisms. Compare this to touching a hot or vibrating surface, where despite the fact that we only perceive these qualities due to the transduction of heat or vibrations into our fingertips, we perceive the heat or movement as belonging to the surface, not as belonging to our fingertips. Similarly, we can distinguish between cases of perceiving the arousal of another and feeling personally aroused in just the same way as we can distinguish cases of touching a hot surface and having hot fingers after having removed one's hand. In both cases, the additional contextual and perceptual information (e.g. seeing the other person, or the hot surface), as well as the differential capacity to stop attending to the phenomenon, helps to locate the quality appropriately.

So similarly in the case of music, the listener's arousal is supposed to track the expressive qualities of the music. This should lead to a direct impression of the emotion *in* the music rather than a self-conscious impression of

personal arousal (though such an impression is certainly possible, especially in very intense cases). This seems to be a faithful description of what it's like to recognize the expressive properties of music.

It is also this argument from the phenomenology of empathy that convinces me that arousal (in some form) occurs at the *moment* of recognition, rather than merely similar neural machinery underlying both recognition and arousal. If we are getting a sense of an *occurrent* emotion expressed by the music, or the other person, then, given that we don't directly receive the sensations of others, what else could provide the phenomenology of the feeling?<sup>4</sup> Even if we employed a memory of a past emotion, recreating the felt quality of a bodily pattern would require at least a neural recreation of arousal. If no arousal takes place, it is difficult to see what else could provide a sense of the feeling in the music, or the other person.

So I claim that to recreate another's bodily changes and projectively perceive the feeling that results as belonging to the other is what it is to empathize with someone in this way. One could not be *directly* aware of the other person's emotion (though one might infer it by other means) without undergoing this process.<sup>5</sup> Matravers is thus on the right track when he claims that arousal is necessary to recognising emotions in music. Being (at least sub-personally) aroused by an emotion that *mirrors* the one expressed by the music will allow us to perceive the music as possessing that emotional quality. However, it seems clear that feeling personally aroused is not a necessary part of the experience of perceiving the emotional content of music. At most, arousal is a necessary part of the causal process. Thus the kind of arousal that I am appealing to here may not be that which would be endorsed by a traditional arousal theorist.

Furthermore, by focusing more on the listener than the music itself, arousal theory does not provide a very informative explanation of exactly what it is about the music that encourages its treatment in emotional terms. How is it that music can generate the arousal of the listener? In order to answer this question we must turn to resemblance theory. Again, our answer will appeal to a deeper explanation of how patterns of bodily changes are simulated in our empathic activities.

### 3. Arousal Relies on Resemblance

There are several types of resemblance theory that differ according to what emotional phenomenon music is supposed to resemble. One idea, provided by both Peter Kivy [1980] and Stephen Davies [1994], is that music resembles the ways in which humans give expression to their emotions. Since

<sup>4</sup>Later I provide more defence for the claim that music captures the felt sensation of emotions.

<sup>5</sup>Note that since the arousal I identify can take the form of simulated bodily changes or attenuated bodily changes, this view is compatible with the possibility that one have an additional and simultaneous emotional response which one identifies as one's own (e.g. *schadenfreude* or sympathy). For instance, the other's response may be simulated whilst one's own is realized with actual physiological changes, or different emotional mechanisms may simply be active at the same time. One may then experience a mixed emotional feeling, or one's attention may successively shift from one aspect of the experience to another.

bodily gestures and vocal utterances can be perceived as expressive of emotions, then things that seem like bodily gestures or vocal utterances could also be perceived as expressive of emotions.

Beginning with the resemblance between music and vocal utterances, we can see that the contour of a melody may directly resemble the rise in pitch at the end of a question, or the emphasis on certain words. However, these aspects of vocal expressions may not be as significant in providing their emotional effects as their more detailed tonal qualities. It is known for instance that infants respond to the emotional *inflection* in their mother's voice without having to understand the propositional content of her words. Psychologist Mechthild Papousek [1996] has shown that when speaking to infants, mothers will use sharp, staccato contours to express disapproval and slower, falling pitch contours to soothe.<sup>6</sup> The expressive qualities of these kinds of details have been verified in empirical studies [e.g. Scherer et al. 2003]. For example, increases in the fundamental frequency (F0) of the voice and its degree of variation, an upward F0 contour, increased articulation rate and intensity all indicate greater arousal. Specific emotions can also be distinguished by the particular pattern of variables. For example, anxiety is characterised by an increase in F0, but low variation and low intensity. Music is clearly able to imitate as well as to exaggerate all of these basic non-verbal features.<sup>7</sup>

Yet although the resemblance to vocal utterances goes some way towards grounding the expressive qualities of music, it is insufficient to account for all cases of musical expression. For example, the broader melodic, harmonic and rhythmic features of a piece of music seem expressive in a way that does not resemble any of the variables mentioned above. Consequently, Davies and Kivy also appeal to resemblance with bodily movements and posture (for Davies [2005] this is primary). Melodic lines can seem graceful or heavy, they can jump or droop and chords can seem tense or gentle. The means by which music may resemble motion includes variations in volume, articulation, rhythm, texture and pitch. These combined resources seem sufficient to capture virtually any type of movement imaginable.

How this connection between musical and bodily movement is made requires a bit more explanation. But first we should explain how resemblance would solve our problem. I've endorsed the view that physiological arousal, or at least a neural simulation of such arousal, is necessary to recognize emotions in others. If music resembles expressive behaviour, we can now go some way towards explaining how this arousal occurs, again by looking at the case of empathy. First of all, given that emotions are necessarily constituted by patterns of bodily change, arousal requires that some of the same bodily changes (or a neural simulation of those changes) happening in the other person are replicated in the empathizer. Since this process must be triggered by the perception of the other person, it is clear that the empathizer can only imitate those bodily changes that he is able to perceive. The most obvious perceivable changes

<sup>6</sup>Cf. Storr [1992: 23].

<sup>7</sup>Music also has the capacity to imitate other specific cues such as sighs, tremors, hoarseness, weeping or laughing, though only in a highly stylized way, and as such not particularly accurately.



are to the other's expressive behaviour, such as his posture, bodily gestures, facial and vocal expressions.

Recall that expressive behaviour counts as one kind of bodily change that generates emotional feelings. Hence if the empathizer behaves expressively, this will generate a range of bodily changes that are characteristic of the overall emotional pattern. These could in turn trigger other physiological changes which typically accompany such expressive behaviour because of sub-personal association mechanisms. However, whether or not these internal bodily changes are also imitated, the imitation of expressive behaviour should generate sufficient arousal to allow recognition of the emotional state.

The phenomenon of emotional contagion, such as panicking when those around you start to panic, or laughing when those around you are laughing, despite not knowing the reason in either case, shows that we can be aroused by another's emotional state as a result of observing it.<sup>8</sup> Emotional contagion is explained as the result of unconsciously imitating the expressive behaviour of others [Hatfield et al. 1994]. Yet this imitation need not be overt. It has been observed that adult humans tend to tense the muscles required to perform an action when viewing another person performing that action [Fadiga et al. 1995]. There is also subtle activation of people's facial muscles when perceiving emotional expressions [Dimberg et al. 2000].<sup>9</sup> Finally, the by now well-known discovery of 'mirror neurons' has helped to explain how imitation is achieved. These neurons, which have recently been confirmed to exist in human brains,<sup>10</sup> fire both when the subject performs an action and when it perceives another performing that same action [Gallese & Goldman 1998]. Mirror neurons indicate a neural level of action mirroring, allowing the perceiver to recreate the motor plans of the other person. As such, they most likely ground a person's capacity to actually imitate the other's actions.

We are also in a position to hypothesize a plausible connection between Damasio's as-if loop and the mirror neuron system. Recall that the as-if loop simulates bodily changes and that these include changes in expressive movement. The activation of mirror neurons which monitor the expressive movements of others may activate the neurons responsible for monitoring (and simulating) the same expressive movements in oneself. Note, however, that the central claim about mirror neurons is that they activate *both* when an action is engaged in and when it is perceived. So it may well be the case that the very same neurons realize both the mapping of one's own expressive behaviours and the mapping of another's expressive behaviours. The as-if loop and the mirror neuron system may partially overlap. In either system, neural activity can stimulate actual motor activity. Similarly, links within the as-if loop to neurons mapping/simulating other physiological changes are liable to trigger a recreation of such physiological changes and perhaps the changes themselves. Below I will also suggest ways in which music could

<sup>8</sup>Davies [forthcoming] has recently connected musical expression to what he calls 'attentional' emotional contagion in explaining how people can be aroused by music. Cf. also Robinson [2006: Ch. 13].

<sup>9</sup>Cited in Damasio [2004: 312].

<sup>10</sup>Slack [2007].

more directly activate the simulation of physiological changes other than expressive behaviour.

So we are gradually building a picture in which mirroring expressive behaviour stimulates neural activity which arouses an emotional feeling, and which is then perceived as belonging to the other person. Similarly in the case of music, we hypothesize that the aural presentation of movement is mirrored, resulting in a sense of feeling belonging to the music.<sup>11</sup> We must now address some potentially problematic differences between the cases of music and empathy.

First of all, how exactly is this mirroring achieved? In the case of empathy, it is necessary that the visual or audible data of the other person's behaviour must be somehow converted into motor plans for producing the same behaviour. However, forming a motor plan of another's visually or aurally perceived behaviour requires a translation of that information into a form that can be directly mirrored. One does not directly mirror the visual look of a person but a first person sense of behavioural movement. Similarly, the perceived resemblance between music and motion is more problematic than that between music and vocal utterances because it must operate between different sense modalities, between sound and either kinaesthetic motion or tactile feeling.

Kivy claims that we connect modalities, attributing spatial positions to pitch, for instance, because of constant associations gathered from everyday life, such as engines whining higher as they use more energy to lift something from the ground [1980: 55]. However, we also hear deep booming sounds from the sky when it thunders. So why should the association between low sounds and low positions be so consistently perceived? It seems that the way we describe pitches as spatially higher or lower is simply entrenched in our language (and hence in our thought). In other cultures, high and low pitches are described as weak and strong (among the African Bashi), white and black (among the Lau of the Solomon Islands), or small and big (among the African Basongye) [Merriam 1964].<sup>12</sup> Despite the differences here, these alternative descriptions make sense to us. They are all highly analogous to our descriptions of high and low, and indicate a universality of applying intermodal metaphors to musical pitch.

We can also provide a deeper explanation for this convention by appeal to the general intermodality of our perceptual processing. First of all, the fact that our own behavioural movements are presented to us in multimodal form can help explain how we make the same association when perceiving the behaviour of others. We not only see, as well as feel, our own bodily movements, we also hear, as well as feel, our own vocalizations. Moreover, our perception of the world is generally geared towards multimodal presentation. Physical movements, as well as the textures of surfaces, are often presented to us visually, audibly and in tactile form.

<sup>11</sup>Neuroscientists Molnar-Szakacs and Overy [2006] similarly hypothesize that recognizing emotions in music is grounded by the mirror neuron system. However, they focus on mirroring the intentional actions required to produce the movements we perceive in music rather than the more generalized mirroring of bodily patterns (some of which could be generated by deliberate actions).

<sup>12</sup>Cited in Davies [1994: 231–2].

It is also known that the senses can influence each other in our recognition of objects. It has been observed, for instance, that when subjects are presented with a single visual flash accompanied by two audible beeps, or two tactile taps, they have a corresponding impression of two visual flashes. The illusion persists even when subjects know that only a single flash is presented, suggesting that it is a feature of early stages of perceptual processing [Violentyev et al. 2005].

A more extreme example of the connection between modalities is the phenomenon of synaesthesia, where subjects report sensations of colours when hearing sounds, or shapes when tasting food. These multimodal stimulations are fairly rare, although certain 'pseudo' synaesthetic associations are much more common. A well-known test of this association is to have people look at two pictures, one similar to an inkblot and the other like a piece of shattered glass and to ask which one is called 'bouba' and which called 'kiki'. Nearly all subjects name the inkblot 'bouba' and the jagged picture 'kiki' [Ramachandran & Hubbard 2003].

Thus there is good evidence for all sorts of reciprocal connections between the sense modalities. It seems that many different forms of sensory information (frequency amplitude, brightness, hue, texture, solidity) are reduced to a sense of movement or three-dimensional shape, flexible between any particular form of presentation.<sup>13</sup> Furthermore, this reduction, or convergence of sensory data, is not achieved by any additional neural module. Although the brain is organized into linear columns of neurons that perform the various stages of sensory processing, there are also lateral connections between modalities at every stage. As such there is a continuous bi-directional flow of information within both these linear and lateral streams. The apparent result is that the input from different sense modalities is integrated right from the beginning of processing, not as a final stage.

So, in just the same way as we intuitively label a jagged visual shape 'kiki' and an inkblot 'bouba', our brains systematically link sounds and spatial movement or shape. It seems conceivable that we could link high sounds with low positions and dark colours rather than the reverse. Yet we can agree with Kivy that the greater amount of real objects and processes that link high sounds with high positions and light colours makes it more likely that our normal intermodal associations will develop. And once these associations are neurally fixed, they will be applied across the board.

#### 4. Resemblance Must Connect to Feelings

We have grounds for automatically perceiving a resemblance between music and movement. Davies and Kivy, accordingly, go a long way towards explaining the expressive capacities of music, and the fact that music may resemble both bodily movement and vocal utterances. However, I think that

<sup>13</sup>Antonio Damasio also describes experiments where both adults and children spontaneously describe the movements of a chip moving on a screen in emotional terms [Damasio 2000: 70].

these resemblances are a means for music to provide a deeper resemblance to the feeling of an emotion, not merely its outward appearance. This is partly motivated by my view of emotions. Emotions are constituted by patterns of bodily changes and their experience is centrally characterized by the feeling of those bodily changes. Given this, it is intuitive to suppose that if music is so good at expressing emotions, then it should be because it captures the experience of undergoing bodily changes.

More importantly it is the nature of expression to give us a sense of what the expressed state is *like*. Davies claims that hearing an emotion in music is no more involved than recognising sadness in the mask that traditionally denotes tragedy [Davies 1997: 97]. But it is quite possible to stare at that mask and get the impression of genuine feeling. Of course I can treat it as a symbol referring to sadness in the same way as the word 'sad' does. But in the case we are interested in, my recognition is supposed to rely on my sensitivity to a resemblance between qualities of the visual form and qualities of the emotion itself.<sup>14</sup> Emotions are not appearances, they are mental states. So Davies and Kivy fail to recognize that what makes sad facial expressions *expressive* of sad emotions (even when not actually driven by emotion) is a connection that is forged between visual qualities and the qualities of the emotional state.<sup>15</sup>

In addition, by restricting the resemblance capacity of music to the outward appearance of emotions, Davies and Kivy unduly restrict the facets of emotions that music is able to express. One example is the visceral sickly feeling that is so well captured by the quiet unsynchronized glissandi of a violin section. In this case, the expressive effect relies on making an intermodal connection between sound and the tactile sense of pressure and motion. Recall that on the view of emotions introduced at the beginning of this paper, these feelings can directly manipulate an emotional experience. Yet the intermodal connection must go beyond the mere sense of vibration if it is to convey *sickliness*. So although one need not simulate expressive behaviour here, one must still simulate the bodily changes that would generate this sense of sickliness. And since it is possible to imagine feelings of sickliness, it is plausible that changes responsible for such feelings can be automatically simulated under the guidance of music.<sup>16</sup>

In general neither timbral nor harmonic tension are adequately captured by the appearance theorist. Yet these sorts of characteristics add a great deal to the expressive qualities of the music. I can think of no vocal or behavioural gesture that captures the peculiar nutty quality of a bassoon for instance. But the timbre of a bassoon suggests a certain feel of solidity that affects the overall emotional feeling.<sup>17</sup> Moreover there are additional dynamic aspects to our experience of emotions that music could be taken to

<sup>14</sup>If one defers to the resemblance between the mask and a real sad face, then one simply pushes the question back to why real faces are expressive of emotion.

<sup>15</sup>Cf. Robinson [2006: 309].

<sup>16</sup>Davies has said to me that the effects of music due to timbre may simply be another channel to expression. However, there is no clear reason to motivate a distinction like this other than to protect the appearance resemblance theory. Timbre is very closely related to both harmonic texture and dynamics, all of which have an easily discernible resemblance to tactile and proprioceptive feeling. The arguments from intermodality and simulation provide a unified account of resemblance.

<sup>17</sup>Vocal timbre similarly conveys a sense of the speaker's body.

resemble. For example, sometimes the thoughts we have during emotions have characteristic dynamic qualities such as being unable to concentrate on one thing for more than a few moments, or the inability to change one's mind when depressed. So one could potentially be guided by the dynamic qualities of the music to simulate (probably more consciously) these sorts of thought dynamics. Another possibility is more unusual: that one imagine the temporal development of a long-term emotion like love compressed into a few minutes, whilst preserving certain structural features such as an initial hesitancy that gradually becomes consistently passionate. Ultimately it seems just as reasonable to characterize the entire phenomenology of an emotional state as languid, restless or clumsy as it does a behavioural gesture [cf. Pratt 1931].<sup>18</sup> So why should we limit music's resemblance to just the outward form of emotions?

On the basis of appearance theories, we may also wonder why music is commonly accorded the distinction of most emotionally expressive art as opposed to (unaccompanied) dance, which can resemble the gestural aspects of emotions far more accurately. The appearance theorist may respond that music has the advantage over dance that it can resemble vocal utterances in addition to bodily gestures. Yet the quality of the expressiveness of music seems far deeper than an extra sort of resemblance could account for. It seems that music can absorb listeners and carry them along with the progress of an emotional state in a way quite foreign to other art forms.

There is a very simple reason why music is superior to the other art forms in its expressive capacity. It is because sound is more like tactile feeling than any other sense modality, and tactile feeling is a central part of emotional feeling (taking tactile feeling to include feelings of pressure, heat and motion within the body as well as at its surface). In many ways the two senses overlap one another. It is worth noting that our sense of hearing evolved from a refinement of our sense of touch. The evolutionary ancestor of the eardrum is a bone in the sides of fish that functions to sense pressure variations in water.<sup>19</sup> Sound, more than sight, is experienced in terms of vibration (although sight also relies on vibrations, it is not experienced as such). A loud sound literally feels a certain way. In addition, friction and movement typically generate sound, meaning that sound is a constant accompaniment to our kinaesthetic and tactile experiences. Finally, both feeling and hearing parse the world in similar ways. They share similar qualities of both abstraction and immediacy. The way that feelings structure experience is not in terms of discrete disconnected objects so much as ongoing actions and textural contrasts. Similarly sound (and particularly music) is understood in terms of continuous streams of movement and timbral contrasts. As Kendall Walton notes, we are much more likely to get emotionally involved in sounds if only because we don't objectify and

<sup>18</sup>Cited in Budd [1985: 39]. Budd similarly recognizes the resemblance to the general phenomenology of feeling [1995: 207].

<sup>19</sup>Nussbaum [2007: 51–4] discusses this point at length, noting the 'strong affinity' between sound and tactile feeling.

distance ourselves from them the way we do with our visual experience.<sup>20</sup> As a result, music literally *resonates* with feeling.

Overall, then, I agree with Malcolm Budd and Carroll Pratt that music sounds the way that emotions feel. Via a variety of resources, music is able to fully capture the dynamic and visceral qualities of emotional feelings, which we are able to simulate using the resources we possess for monitoring our own and others' emotions. Since the inner experience of emotions is primarily characterized by the temporal development of these qualities, it is no wonder that music is the pre-eminent art form of emotional expression.

### 5. Persona Enables Resemblance

If we agree that we are sensitive to a resemblance between music and emotional feelings, there is one final piece of the puzzle required in order to fully account for the causal story of musical expression. It is a philosophical truism that anything can resemble anything else. So why is music taken to resemble *human* movement rather than any other kind of movement? I have argued that music hijacks the simulation mechanism which recognizes the actions and emotions of other people. But why does the simulation mechanism get activated at all? In addition, my rejection of the appearance theories of Davies and Kivy leaves my account open to the problem that their theories resolved. 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. The arousal theory was rejected for failing to concentrate on the music itself. So how is it that we can properly focus on the music itself whilst still holding that it expresses an actual emotion?

I believe the gap here can be filled by persona theory. Persona theory argues that when we hear a piece of music as emotionally expressive, we necessarily imagine or have an illusion that a person is appropriately connected to that emotional expression.<sup>21</sup> The person we imagine need not be any actual person; it may only be connected to that specific piece of music.

Now the fundamental purpose of the persona is to make the recognition of emotions in the music intelligible. It should therefore be impossible to hear an emotion in the music without having also imagined a persona. But what makes a listener think of a persona? There are two possible strategies here. The first is that it is a fairly automatic illusion. The second is that it is a more deliberate imaginative activity. As justification for the first possibility, it is very likely that whenever we approach a work of art we have a

<sup>20</sup>Walton [1997: 78], noting our tendency to experientially separate sounds from their sources, says,

We reify or objectify feelings and sensations, as we do sounds, and we conceptualize them and our relations to them in similar ways. We think of feelings of exuberance or anguish as entities distinct from their sources, and sometimes as leaving their sources and surrounding or entering us. A feeling, like a sound, may come over me. It may permeate my consciousness.

<sup>21</sup>E.g. Cone [1974], Vermazen [1986] and especially Levinson [2005].

background belief that it has been deliberately constructed by a human being. As such we will tend automatically to interpret that work as the product of certain mental states, and derive the nature of those mental states from the characteristics of the work. In neurological terms, this means that we confront works of art in just the same way as we confront people, which, I would suggest, primes our simulation capacity from the start. For example, if I see a person stub his toe, I don't have to deliberately engage my simulation mechanism in order to appreciate the pain that this will engender. Rather I immediately get a sense of the painful consequences of this action. So, similarly, when we listen to music, our background belief that a person is responsible for the work could prime the simulation mechanism. The music then only need give us the merest hint of a person (i.e. by resembling person-like movements) to trigger the mechanisms geared towards understanding people.<sup>22</sup>

The alternative is to assume that the listener does not approach the music with the background belief that it is produced by a human and their simulation mechanism is not primed. This is more common when listening to purely natural sounds such as a tap dripping. Again, we still immediately get a sense of movement from the sound of the dripping water, and if we concentrate upon that sound we can even get a tactile sense of the vibrations that this movement would generate. However, we do not perceive any emotional qualities in this sound. At this point then, the listener may now *deliberately* imagine that a person is responsible for producing this sound. For instance, that a person is flicking the drops of water from a brush. This will then trigger the simulation process, and generate a sense of the way it feels to perform those actions that would produce those sounds in that precise manner. Depending on the properties of the sound, this may seem to have emotional qualities or not. For instance, the sound of the tap dripping may now appear to have a nonchalant quality it didn't have before.

The above example provides an informal test of the simulation theory I have offered. Empirical verification is also available given the direct connection I have made between our capacity for empathy and our capacity to recognize emotions in music. In particular, people with autism are usually characterized as lacking the ability to understand other minds. This is attributed either to a lack of a theory of mind or an inability to simulate the internal states of others. So we should anticipate that autistic subjects are unable to understand expressive music.<sup>23</sup> Little has been done to confirm this hypothesis, yet in one experiment conducted, autistic subjects were in fact able to successfully identify the emotions presented in music [Heaton et al. 1999]. Unfortunately, this experiment only asked autistic subjects to identify emotions in music that they were already able to identify in other people. If they can recognize when a person is sad or happy then we should expect them to recognize the same in music.

<sup>22</sup>Kivy [1980: 59] similarly points to a general 'animating tendency' to imbue natural objects with human characteristics. As he says, 'far from being difficult to hear or see things as animate, it is, apparently, difficult not to'.

<sup>23</sup>Similarly, where people with Huntington's disease confuse expressions of fear and anger, we should expect them to also confuse music that expresses fear or anger.

It is uncertain what it is exactly that autistic people are unable to do, and there are likely to be several interacting factors that lead to the disorder. If autism is caused by a dysfunction in the simulation mechanism (such as a lack of mirror neurons) then autistic subjects should not be able to process the emotional content of music at all beyond a level they are capable of achieving for people. If instead the problem is more to do with *attributing* the results of simulative processing, then we may expect that whilst autistic subjects could be contagiously aroused by the expressive qualities of music, any conscious recognition of emotion would be phrased in egoistic terms. Autistic subjects would not ascribe the emotion to some other persona, and possibly not even as caused by the music.

## 6. Conclusion

In summary, the causal story I have presented is that music expresses emotions by hijacking the simulation mechanism that has evolved to detect one's own and other people's emotions. This simulation mechanism is triggered either by a background belief or an active imagination of the agency involved in generating the sound. The simulation mechanism utilizes the intermodal connection between the sounds and movement that the brain makes at all times. These movements are then mirrored from a first person perspective, which, if they display an emotional pattern, will arouse a simulation of that emotion in the listener. So far all of this can happen unconsciously. The listener may even be fully aroused by an emotion without recognizing that it is caused by the properties of the music. This would be like a case of emotional contagion. If, however, the listener is paying conscious attention to the music then they will perceive that music as having the properties of feeling that their simulation process has generated. They will then be disposed to verbally identify the expressive content of the music accordingly. So because the judgment is conscious, and one is generally aware that the properties of the music are responsible for one's judgment, this is more like a case of empathy than one of emotional contagion.

This brings us to the end of the causal story of musical expression and the emergence of the conscious experience of identifying emotional content in music. To finish, two final remarks can be made regarding which account of the experience of music this causal story seems to best support.

First of all, my account is guided by the idea that a pattern of feeling is perceived as a property of the music. We don't require two ideas; one of the music and another of the feeling. This is because the brain automatically makes an intermodal connection between the two forms of sensation, combining them into a single percept of the dynamic qualities of the music. As a result, we should look for comparable phenomenological accounts, for example, the visual perception of the solidity of an object. Just as when I see a wooden beam I also perceive its solidity, so when I perceive the music I also perceive its feeling. This is analogous to the way we perceive emotional feelings in the facial or vocal expressions of other people. There are of



course discernible features within those faces, which it is possible to attend to in a purely technical manner. That is, it is possible to distinguish the facial expression and the feeling. But the more common experience is just to perceive the emotion *in* the face. Similarly we hear the emotion *in* the music.<sup>24</sup>

Second, because the simulation mechanisms I have invoked to explain expressivity are functionally geared towards tracking the emotions of people, it is plausible that the results of such processing would give the impression of a person. So it looks like a version of the persona theory, such as that defended by Jerrold Levinson [2005], is supported by the causal considerations I have presented. Levinson's concept of a persona is fairly minimal so there may be several ways in which the persona could be characterized or identified in the imagination of the listener. One point worth noting, however, is that if one adopts the theory of emotions that I have employed here, it looks as though to have a sense of feelings is to have a sense of a person's *body* connected to that feeling. This is what emotional feelings *are*, meaningful presentations of the impact of the world on the bodily organism or self. Hence one would need to emphasize the imaginative connection to an embodied rather than disembodied persona.<sup>25</sup>

Swiss Center for Affective Sciences (CISA)

Received: May 2008

Revised: November 2008

## References

- Budd, Malcolm 1985. *Music and the Emotions: The Philosophical Theories*, London: Routledge & Kegan Paul.
- Budd, Malcolm 1995. *Values of Art: Pictures, Poetry and Music*, London: Penguin.
- Cochrane, Tom unpublished. *Enactive Emotions* (available on request).
- Cone, Edward T. 1974. *The Composer's Voice. Ernest Bloch Lectures*, Berkeley: University of California Press.
- Currie, Gregory 2004. *Arts and Minds*, Oxford: Oxford University Press.
- Currie, Gregory & Ian Ravenscroft 2002. *Recreative Minds*, Oxford: Clarendon Press.
- Damasio, Antonio 2000. *The Feeling Of What Happens: Body, Emotion and the Making of Consciousness*, London: Vintage.
- Damasio, Antonio 2004. *Looking for Spinoza*, London: Vintage.
- Davies, Stephen 1994. *Musical Meaning and Expression*, Ithaca and London: Cornell University Press.
- Davies, Stephen 1997. Contra the Hypothetical Persona in Music, in *Emotion and the Arts*, ed. M. Hjort & S. Laver, Oxford: Oxford University Press: 95–109.
- Davies, Stephen 2005. Artistic Expression and the Hard Case of Pure Music, in *Contemporary Debates in Aesthetics and the Philosophy of Art*, ed. Matthew Kieran, Oxford: Blackwell.
- Davies, Stephen forthcoming. Infectious Music: Music-Listener Emotional Contagion, in *Empathy: Philosophical and Psychological Perspectives*, ed. P. Goldie & A. Coplan, Oxford: Oxford University Press.
- Deonna, Julien 2006. Emotion, Perception and Perspective, *Dialectica* 60/1: 29–46.
- Dimberg, U., M. Thunberg & K. Elmehed 2000. Unconscious Facial Reactions to Emotional Facial Expressions, *Psychological Science* 11/1: 86–9.
- Dretske, Fred 1981. *Knowledge and the Flow of Information*, Cambridge, MA: MIT Press.
- Dretske, Fred 1986. Misrepresentation, in *Belief: Form, Content and Function*, ed. Radu Bogdan, Oxford: Oxford University Press: 17–36.
- Fadiga, L., L. Fogassi, G. Pavesi & G. Rizzolatti 1995. Motor Facilities During Action Observance: A Magnetic Stimulation Study, *Journal of Neurophysiology* 73: 2608–11.

<sup>24</sup>Cf. Trivedi [2001].

<sup>25</sup>I would like to acknowledge the help I received from Greg Currie, who supervised the PhD from which this article was derived. I'd also like to thank the organizers of and participants in the conference, 'Aesthetic Psychology' (Durham University, September 2007) where I presented an earlier version of this paper, as well as two anonymous referees for this journal for their helpful criticisms.

- Gallese, V. & A. Goldman 1998. Mirror Neurons and the Simulation Theory of Mind-reading, *Trends in Cognitive Sciences* 2/12: 493–501.
- Goldman, A. I. & C. S. Sripada, 2005. Simulationist Models of Face-based Emotion Recognition, *Cognition* 94/3: 193–213.
- Hatfield, E., J. T. Cacioppo & R. L. Rapson 1994. *Emotional Contagion*, Cambridge: Cambridge University Press.
- Heaton, P., B. Hermelin, & L. Pring 1999. Can Children With Autistic Spectrum Disorders Perceive Affect in Music? An Experimental Investigation, *Psychological Medicine* 29: 1405–10.
- Kivy, Peter 1980. *The Corded Shell. Reflections on Musical Expression*, Guildford: Princeton University Press.
- Laird, James 2007. *Feelings: The Perception of Self*, Oxford: Oxford University Press.
- Lawrence, A., I. Goerendt, & D. Brooks 2007. Impaired Recognition of Facial Expressions of Anger in Parkinson's Disease Patients Acutely Withdrawn from Dopamine Replacement Therapy, *Neuropsychologia* 45/1: 65–74.
- Levinson, Jerrold 2005. Musical Expressiveness as Hearability-As-Expression, in *Contemporary Debates in Aesthetics and the Philosophy of Art*, ed. Matthew Kieran, Oxford: Blackwell.
- Matravers, Derek 1990. *Art and Emotion*, Oxford: Clarendon Press.
- Merriam, Alan 1964. *The Anthropology of Music*, Evanston: Northwestern University Press.
- Molnar-Szakacs, I. & K. Overy 2006. Music and Mirror Neurons: From Motion to 'e'motion, *Social Cognitive and Affective Neuroscience* 1/3: 235–41.
- Nussbaum, Charles O. 2007. *The Musical Representation: Meaning, Ontology, and Emotion*, Cambridge, MA: MIT Press.
- Papoušek, Mechthild 1996. Intuitive Parenting: A Hidden Source of Musical Stimulation in Infancy, in *Musical Beginnings: Origins and Development of Musical Competence*, ed. I. Deliège & J. Sloboda, Oxford: Oxford University Press: 89–112.
- Pratt, Carroll C. 1931. *The Meaning of Music*, New York: McGraw-Hill.
- Prinz, Jesse 2004. *Gut Reactions: A Perceptual Theory of Emotion*, Oxford: Oxford University Press.
- Ramachandran V. & E. Hubbard 2003. Hearing Colors, Tasting Shapes, *Scientific American* (May 2003).
- Robinson, Jenefer 2006. *Deeper Than Reason: Emotion and its Role in Literature, Music, and Art*, Oxford: Oxford University Press.
- Scherer, K. R., T. Johnstone & G. Klasmser 2003. Vocal Expression of Emotion, in *Handbook of Affective Sciences*, ed. R. J. Davidson, K. R. Scherer, & H. Hill Goldsmith. London: Oxford University Press.
- Slack, G. 2007. Found: the source of human empathy, *New Scientist* 196/2629 (10–16 November 2007): 12.
- Storr, Anthony 1992. *Music and the Mind*, London: Harper Collins.
- Trivedi, Saam 2001. Expressiveness as a Property of the Music Itself, *The Journal of Aesthetics and Art Criticism* 59/4: 414.
- Vermazen, Bruce 1986. Expression as Expression, *Pacific Philosophical Quarterly* 67: 196–234.
- Violentev, A., S. Shimojo & L. Shams 2005. Touch-induced Visual Illusion, *Neuroreport* 16/10: 1107–10.
- Walton, Kendall 1997. Listening with Imagination: Is Music Representational? in *Music and Meaning*, ed. J. Robinson, Ithaca: Cornell University Press: 57–82.