

# **Shared Emotions in Music**

**Thomas Cochrane B.A. Hons, M.A.**

**Thesis submitted to the University of Nottingham  
for the degree of Doctor of Philosophy**

**March 2007**

## Abstract

In this thesis I show that groups can share token emotional states by performing music together. First I argue that emotions are perceptions, representing the self's dynamic relation with the world. This representation is achieved by patterns of bodily changes, functioning independently of conscious feeling. Moreover, emotional expressions should be included in this analysis because they contribute to the pattern of bodily changes. This entails that we can 'think through' our emotions by manipulating our behavioural expressions. I then argue that empathy relies on our tendency to neurally mirror the expressive behaviours of other people, resulting in a simulation of emotional arousal.

Turning to music, I argue that music hijacks our simulative capacities and thus that recognising emotions in music is like recognising emotions in people. The fact that the brain processes patterns of sound, vision and touch intermodally as patterns of movement underlines this claim. All this allows me to argue that musicians can use music to physically extend the cognition of their emotions. Here the music may not just influence their bodily changes, but may be processed alongside those changes as an elaboration of the overall pattern. On some occasions, the music may even take the dominant role in this respect. Thus emotional representations are best described more neutrally, though bodily patterns remain the central case of emotions.

I then analyse joint listening to music, arguing that our perceptual activities may be interdependently structured, mutually fixing the character of the object, as well as encouraging similar emotional responses. In order to show that the intrinsic content of mental states can be shared, I then look at the theory of collective intentions. This provides a model for embodying the content of a mental state in the agreement between individuals. I apply this model to ensemble musical performance.

## **Acknowledgements**

There are several people who have been invaluable in helping me to complete this thesis. I have in particular to thank my supervisor Greg Currie for keeping me on track, for his always useful suggestions and criticisms and his complete intolerance for any vague overstatements of the issue. And of course for “what? You mean your supervisor doesn’t meet with you every two weeks as well??”

Secondly I must thank my fiancé and soon to be wife Ju who I first met just as I was starting this PhD. She has helped to make the whole experience a good one, she has supported me, taken care of me, and sometimes let me try and explain my theories to her.

I should also thank the Jacobsen Fellowship for funding during 2004-2005. And of course my parents for their help and support, especially in the two years prior to starting when I was mostly sitting around at home trying to figure out the future of humanity.

Finally I’d like to thank everyone in the Nottingham University philosophy department, for discussion, support, work and ideas.

**Tom Cochrane**

**March 2007**

## Table of Contents

<b>Introduction</b> .....	Page 5
<b>Chapter One: Emotions as Perception</b> .....	Page 12
<b>Chapter Two: Expression and Empathy</b> .....	Page 50
<b>Chapter Three: Emotions in Music</b> .....	Page 86
<b>Chapter Four: The Real Properties of Music</b> .....	Page 124
<b>Chapter Five: Music and the Extended Cognition of Emotion</b> .....	Page 166
<b>Chapter Six: Joint attention to Music</b> .....	Page 209
<b>Chapter Seven: Shared Emotions in Performance</b> .....	Page 247
<b>Conclusion</b> .....	Page 307
<b>Appendix: Experiment in Shared Emotions no. 1</b> .....	Page 311
<b>Bibliography</b> .....	Page 316

### **CD: Experiment in Shared Emotions No. 1.**

**Track 1: Extended Cognition (solo keyboard)**

**Track 2: Agreement**

**Track 3: On being Recorded**

**Track 4: On the Weather**

**Track 5: On our Friendship**

**Keyboard:** Thomas Cochrane. **Drums:** Felix May. **Saxophones:** Matthew Wright.  
**Recorded live, Sunday 4th March 2007. Produced by Thomas Cochrane.**

**This recording is intended to provide examples to help illustrate the thesis.  
It is not intended to be assessed on its success as a piece of music,  
or as providing any theoretical claims additional to the text.**

## Introduction

When we can put on headphones and listen to music as if it's spontaneously pouring from our own brains, we might forget that music is essentially a social activity. Music is created by humans for humans. Cross-culturally, music is used to accompany all manner of social rituals from weddings to funerals, festivals, religious services, healing ceremonies, games and storytelling. Music can send us to sleep, send us into battle or accompany work in the field or factory. In addition all societies use music for dancing to, either as part of rituals or for its own sake. In general music is a ubiquitous means by which we mediate our social interactions. It is no surprise then that people use music to assert their collective identity as well as their distinctness from other social groups (consider for instance, national anthems or genre preferences amongst teenagers).

Claims are also made that music can symbolically express our social relationships. For example a concerto may symbolise the dialogue between the individual and the social forces that surround him. Similarly the individually expressive yet harmonised voices of Bach's polyphonic music are said to symbolise an ideal balance between individuality and collectivity (cf. Small 1998). Yet music may not just *represent* social solidarity in an idealised form, but actually help to generate it to begin with. The group cohesion encouraged by shared musical production is even cited as an evolutionary reason for the development of music. Psychologist Isabelle Peretz writes:

Music possesses two design features that reflect an intrinsic role in communion (as opposed to communication, which is the key function of

speech). Pitch intervals allow harmonious voice blending when sounding together, and temporal regularity facilitates motor synchronicity. (Peretz 2001: 115)

Contrast this with spoken language: Though speaking may be directly expressive of certain mental states (such as beliefs), it cannot allow many people to talk all at once without destroying meaning and coherence. Words must be distinctly produced and perceived in a way that musical patterns need not be. In addition, I argue that music can express and constitute emotions to at least the same degree that words can express and constitute beliefs or reasoning processes. As such the twin qualities of music to enable both intense expressivity and blending make it almost uniquely suitable for the sharing of mental states.

Yet my coming to this thesis was not motivated by conceptual concerns such as this. In the first place I was seeking to explain and justify various experiences that I've had whilst playing music of intense absorption with the sound I was making, as well as a feeling of sharing a basic sense of life with my fellow musicians. Other musicians I asked reported similar experiences. Some similarities could also be found in reports of religious ceremonies (being overtaken by the 'spirit' of the occasion). Hence there seemed to be a definite phenomenon that could serve as the basis for investigation.

Ultimately I argue that music can allow people to share emotional states, not their conscious experience. Yet this is certainly enough to justify the socially cohesive powers of music. However the level of absorption in music that stimulates the

experiences that inspired me seems to be a rare occurrence, typically a result of serendipity more than careful planning. Nevertheless I hope to show exactly why it happens and how it happens. With any luck, this might help to make it a less rare occurrence. At the same time of course, there are a number of philosophical goals to this thesis. On the route to justifying my claims, I explore several issues concerning the nature of mind, emotions, empathy, music and collective behaviour, many of which are further illuminated by the possibility of shared emotions in music.

It is worth making a clarification here about what I mean by ‘sharing’. Sometimes we say that we share an emotion and we mean that we both possess qualitatively similar (though numerically distinct) emotional states. Sometimes, these emotions are also directed towards the same state of affairs (cf. Goldie 2000: 192). However, the sense of sharing I am interested in here is more radical than either of these possibilities. By sharing an emotion, I mean that there are two people, one mental state. Now emotions are not like cake, where you can each have a portion of the totality. Rather the sharing involved is one of joint possession or control over the emotion; in the manner we might share ownership of a house. The individuals involved may contribute in different ways to forming the state, but overall they jointly possess the whole thing as a totality.

Clearly, arguing that people can share mental states rejects a very basic assumption that people have about minds; that minds are only possessed by individuals. However there is a general movement in philosophy towards viewing the mind as something extended beyond the boundary of skin and skull. This is known as *externalism* or *extended cognition*. With these theories we have begun to appreciate

how much thinking is an activity that is realised in our physical manipulations of the environment. Some claims are also made about the way that other people help us to realise our mental states. Yet even these theories tend not to address the core qualitative kinds of mental states, perceptions and emotions. This is where my thesis takes a step further.

In recent years our understanding of emotions has also developed. In chapters one and two I defend and refine Jesse Prinz's (2004) theory of emotions, which argues that emotions are essentially perceptual states, constituted primarily by patterns of bodily changes that represent our relations with the world. I use this theory to ground my explorations of emotional expression and empathy in chapter two. Here I argue for an especially intimate link between expressive behaviour and emotional states. One important claim is that we can use expressive behaviour to 'think through' our emotions. Another is that other peoples' expressive behaviour can provide us with an immediate sense of their inner feelings.

Of course, music is renowned for its ability to express emotional states. There is a long philosophical tradition of trying to explain how music is able to capture emotional states so effectively. In chapters three and four I review this research, firstly explaining the basic capacity of music to express emotions, and then justifying to what extent the expressive properties of music are really present in the music. I argue that the theory of emotions and empathy I hold helps to synthesise many of the prevalent theories of musical expression.

Yet although the model of emotions I defend is ultimately intended to apply to music, it is justified independently of any ideas about musical experience. In general it is unwise to base a theory of X (i.e. emotions) on some application of X in a limited field (i.e. expression in music). However I would argue from the example of music to a theory of emotions if it were clear which theory of the emotions music supports. Unfortunately there is still much debate as to how music manages to express emotions, or indeed whether what it expresses is much like a real emotion at all.

Overall, there is a relation of metaphysical supervenience between emotional expression in music and emotions in general. This relation is supported by an argument that our experience of music is a by-product of more basic emotional functions that have evolved for other purposes. Yet before the end of this thesis, I argue that music *does* capture the essential elements of emotions, and that it *does* provide insights into the nature of emotions, as well as the nature of mind generally. I even think that music allows the development and transformation of our emotional capacities, particularly of course, in regard to the possibility of shared emotions. So although I base my theory of music on a theory of emotions, I eventually use that theory of music to reciprocally develop theories of emotion (and theories of mind generally). Hence once the basic relation between music and the emotions has been established, the phenomena surrounding the creation and reception of music can more securely reflect on the emotions and other mental processes that underlie them.

So is it the case that my protracted argument for shared experience in music will stand or fall according to the strength of the theory of emotions upon which it is founded? If our theory of emotions was supplanted, some of the insights of my

argument should remain intact inasmuch as they make sense of the phenomenal experience of music rather than the cognitive processes underlying it. However my account of music is tailored to the model of emotions I offer below to such an extent that it would have to be significantly altered to fit another model. In such a circumstance, any other model of the emotions would need to permit two general premises in order to justify my overall theory: Firstly, that music provides direct access to the inner character of emotional states. Secondly, that deliberate emotional expression can be (at least partly) constitutive of an emotional state. It is certainly not the case that *any* theory of the emotions will permit these two claims. However I argue that this is true for any account of the emotions that accords bodily changes a central role.

Overall, the intimate relation between music and emotions allows me to argue in chapter five that a musician can use music to extend the cognition of his emotional state. I argue that jazz improvisation is particularly suitable in this regard because it allows the musician to sincerely *commit* to the music as directly responsive to his actual occurrent emotion in a way that classical performance tends not to encourage.

Having justified the idea that music can partially constitute a musician's emotional state, the final two chapters of my thesis then explore the social issues surrounding musical perception and production. In chapter six I address the nature of joint attention (where we both attend to the same thing in the environment at the same time) and argue that it is best explained as interdependently structuring our perceptual activities. The same interdependence can then apply to when we jointly listen to music and converge on similar emotional responses.

In chapter seven I then explore the possibility of sharing not just the overall structure of a mental task, but also the intrinsic content of a mental state. First I explore the phenomenon of collective intentionality, which has attracted some significant research in recent years. The question here is whether collective intentions are simply the combination of the intentions of the individuals involved, or whether the group forms a distinctive entity of thinking and acting in its own right. I argue that we should indeed think of collective intentions in this latter way. The discussion of collective intentions then provides a model for the sharing of emotions in music. Music can allow a group of musicians to mutually control and manipulate the intrinsic content of their emotional states, where their emotions are partially embodied in the sounds they produce. Finally, I end the thesis by suggesting some practical methods by which groups may come to enjoy shared emotional states in this manner.

## Chapter One: Emotions as Perception

The purpose of this first chapter is to set up and defend a model of the emotions, which in later chapters I can apply to the more specific case of emotional expression in music. In general the debate about how emotions should be understood has been dominated by two positions; cognitivist and non-cognitivist. Both are broad categories of theories but in general, cognitivists regard cognitive states such as judgements as essential to emotions where non-cognitivists do not, focusing instead on bodily changes and the experience of those changes. Yet there is some disagreement concerning what actually counts as a cognitive state, as well as what kinds of roles bodily changes can instantiate. Hence in this chapter I argue that the dichotomy has outlived its usefulness. The evidence demands some form of hybrid or synthesis of these two positions.

The particular synthesis that I wish to defend here is Jesse Prinz's (2004) perceptual theory of emotions. Prinz follows non-cognitivists such as William James, Carl Lange and more recently Antonio Damasio in placing bodily changes central to emotional states. However, he also makes a plausible case for how these bodily changes can be intentionally directed towards the environment, thus incorporating some of the insights of the cognitivist position. Overall Prinz argues that emotions are constituted by bodily changes that perceptually represent 'core relational themes'. These core relational themes are then certain formal or general aspects of the situation, common to all cases of the emotion, which bear on the well-being of the subject in some way. I refine this notion somewhat and argue that bodily changes represent the subject's *dynamic* relation with the world. This to me makes more

sense of the appropriateness of particular bodily changes in representing the situations they represent.

So the general goal of this chapter is to establish the perceptual nature of emotions. I also look at how conceptual interpretation and the focus of attention can affect emotions in ways analogous to other kinds of perceptual states. Yet I argue that bodily changes can represent situations independently of any conscious awareness of the subject. This claim will be important in establishing that musicians can share emotions without having to share their subjective experiences of that emotion.

Having established a basic model of the emotions in this chapter, I then go on in chapter two to develop this model somewhat. In particular, the recognition that the bodily changes generated by behaviour should be included within the emotional state forces us to reevaluate the relation that emotions have with both cognition and action. This then impacts on our understanding of emotional expression and our ability to recognise the emotions of others. So in chapter two I argue that whilst the central function of emotions is to represent certain aspects of one's situation, they also play a richer role in our interactions with the world and other people.

### **Cognitivism versus Non-cognitivism**

In order to see what motivates Prinz's perceptual theory of emotion, it is necessary to provide some background on the debate between cognitivism and non-cognitivism. The most famous non-cognitive theories are those of William James (1884) and Carl Lange (1885) who independently proposed what Prinz calls 'somatic feeling' theories. Here they identify somatic changes in the body as essential to emotional

states. Narrowly construed, the somatic system is what processes information about the muscles of the body. However the term is usually meant to cover a wide range of bodily responses. These include changes to the respiratory system (such as breathing rate), circulatory system (heart rate), digestive system, musculoskeletal system (including facial expressions) and the endocrine system (the secretion of hormones). Lange focuses on the vasomotor system in particular, which regulates blood flow, where James refers to changes in the viscera, facial expressions, and behaviours like crying or impulsively striking out.

Prinz describes the James-Lange positions as somatic *feeling* theories because they both stipulate that an emotion must have a phenomenological feel in order to count as an emotion. So altogether they claim that perception of some object causes a change in the body, the phenomenal feeling of which constitutes the emotion. Antonio Damasio (2000) has a similar view, but claims that the brain may register bodily changes in a way that constitutes an emotion without these changes being consciously experienced.<sup>1</sup> For the moment I take a neutral stance and assume that on the non-cognitivist view, the bodily changes involved in emotions are registered and organised at *some* level, whether this level is conscious or purely neural. Then at the end of the chapter, having got a better sense of the relation between bodily changes and the overall emotional state, I examine whether they need be consciously experienced or not. Also to avoid confusion, throughout this thesis I use the term ‘feelings’ to refer to the phenomenal experience of bodily changes and ‘bodily pattern’ to refer to the registration and organisation of bodily changes in the brain.

---

<sup>1</sup> Damasio believes that emotions have evolved from the homeostatic functions of the brain, which constantly track and alter the state of the body to keep it within survival limits (e.g. not too cold or hot). This view is quite compatible with my claim below that emotions represent the dynamic relation between self and world, inasmuch as that relation bears on the *integrity* of the body or self.

So non-cognitivists distinguish emotional states from other mental states by looking at the characteristic patterns of pangs, tingles and surges that they seem to involve. In contrast, cognitivist theorists have been impressed by the fact that emotions usually involve a response to the person's situation or can be embedded within purely imaginative episodes. Some cognitivists insist that emotions are identical to thoughts or judgements (e.g. Solomon 1976), where others include bodily changes as necessary components of emotional states (e.g. Lazarus 1991). Yet they all agree that emotions go beyond mere bodily changes.

Typically cognitivists argue that emotions are a kind of propositional attitude.<sup>2</sup> A simple cognitivist characterisation of an emotion would be as follows: I represent the world in some way, for example I believe that I have forgotten my mother's birthday. I then have an attitude towards that representation, for example a belief that I have transgressed a moral imperative. There are two propositional attitudes here, but it is only the second that constitutes my emotion of guilt. Hence the emotion of guilt or any other emotion is never just a simple occurrent state, but should be fully characterised as 'guilt *that* x', where 'x' involves some sort of appraisal of the world as it relates to my well-being. Cognitivists then analyse these propositional attitudes in various ways. They may be reduced to beliefs, evaluations, judgements, construals, imaginations, desires or combinations thereof. For instance an emotion of anger may

---

<sup>2</sup> A notable exception is Robert Solomon:

If Mary loves Mary and hates spinach, the objects of his emotions are Mary and spinach, respectively, not propositions. If Fred believes that spinach is good for you (and that, perhaps is why he loves it), the object of his *belief* (but not his emotion) is the proposition that spinach is good for you. (Solomon 2003b: 181)

However Solomon still retains the basic insight that emotions are *about* something.

be composed of a belief that I have been insulted combined with a desire (or action tendency) to revenge myself.

It is not the case that the possessor of an emotion must regard these propositional attitudes as somehow true. I can for instance ‘entertain’ a thought that there is danger, which constitutes my fear state, without actually believing that there is danger (Greenspan 1988). However, arguments in favour of cognitivism typically point out the meaningful aspects of emotions. Emotions are usually not only directed at relevant situations but can also be more or less appropriate or warranted by those situations. In contrast they argue, how can a bodily change be meaningfully *about* something?

Showing that the bodily changes involved in emotions are in fact meaningful is key to Prinz’s theory. For the moment however we can admit that meaningful bodily changes are at least possible. For example, Prinz offers the experience of feeling sickened by some moral offence. We say that the sick feeling is not just triggered by the idea of the moral offence but that we feel sickened *over* that offence. Yet we would be hard pressed to say that feeling sickened is a propositional thought of any kind (Prinz 2004: 27).

In the above case, the sick feeling, (or the emotion of disgust) is directed at a propositional object, the moral offence. However, the sick feeling may be considered independently of the propositional object that it happens to attach to. The same feeling may be triggered by some other thought or continue even after the propositional object has been forgotten. These considerations reveal a general

question that needs to be addressed: Of the many potential aspects of emotional states, which of them are essential or fundamental and which are merely contingent causes or effects? Also, if we regard emotions as being constituted by multiple components, how do these components hang together? Is there some key function that these components jointly deliver?

It is by locating necessary and sufficient conditions for emotions that we can begin to answer these questions. Accordingly one of the main arguments for the necessity of bodily changes to emotional states is to be found in William James's paper 'What is an Emotion?' (1884). Here he asks the reader to imaginatively subtract all the bodily symptoms from an emotional state and see if what remains could intuitively be called an emotion:

Can one fancy the state of rage and picture no ebullition of it in the chest, no flushing of the face, no dilatation of the nostrils, no clenching of the teeth, no impulse to vigorous action, but in their stead limp muscles, calm breathing, and a placid face? The present writer, for one, certainly cannot. The rage is as completely evaporated as the sensation of its so-called manifestations, and the only thing that can possibly be supposed to take its place is some cold-blooded and dispassionate judicial sentence, confined entirely to the intellectual realm, to the effect that a certain person or persons merit chastisement for their sins. In like manner of grief: what would it be without its tears, its sobs, its suffocation of the heart, its pang in the breast-bone? A feelingless cognition that certain

circumstances are deplorable, and nothing more. Every passion in turn tells the same story. (James 1884: 194)

It seems that when we take away the conscious feelings of bodily changes, we remove the emotion's distinctive character, its intensity, as well as its immediate motivational force. So the strength of this argument depends on showing that we ordinarily think of emotions in terms of the feelings of bodily changes. In a similar vein, Lange cites our linguistic practices whereby many of our emotion terms also refer to bodily changes, for example to 'shudder' with fear (Lange 1885: 678). The practice of using bodily metaphors for emotional states also appears to be cross-cultural (Heelas 1986 and Wierzbicka 1999 cited in Prinz 2004: 139). Various cultures associate the emotions with particular body parts such as the liver, heart, stomach or the intestines. This indicates a common intuition that there is some link between emotions and the body, if not definite correlates.

However, is this link to bodily changes true of all emotional states all the time? The range of states we might describe as emotions is broad and without sharp boundaries. Yet Prinz claims to find felt aspects to even highly intellectual passions, such as those involved in solving mathematical puzzles (Prinz 2004a: 49). Without these feelings he claims, the phenomenology of the emotion is exhausted. Of course, the problem with these kinds of arguments is that different people have different imaginative abilities. If someone claims to be able to imagine an emotional state without any corresponding feeling we cannot simply deny the validity of that imagination without additional justification. In addition, if we agree that emotions need not be consciously experienced, the argument from phenomenology might not

be considered relevant, since the phenomenal experience of emotions may be entirely incidental to their purpose. Then again, we might expect that our self-reflective representations of emotions in conscious awareness would generally target their most central component.

Overall, I do not find James' argument for the necessity of bodily feelings to be conclusive, though a recognition that most people do in fact think of emotions in terms of physical feelings lends some support to the non-cognitivist position. However one potential counter-example to the above argument is that emotions such as love or loneliness can persist for many years without constant bodily perturbation. But here it is reasonable to claim that these more complex states may be called emotions because they reflect the *disposition* of the subject to have certain distinctive bodily states aroused, and these dispositions can continue to exist whether or not they are currently activated. Furthermore, Prinz argues that if someone claimed to be in love and yet never experienced *any* felt response, we would regard this person as 'disingenuous or confused' (Prinz 2004a: 50) rather than actually undergoing the emotion.

Compare these considerations to the proposed necessity of thoughts: Prinz argues that if emotions are essentially propositional attitudes then the possessor of the emotion must possess the concepts from which these attitudes are constructed. For example, to fear that there is danger looming (a propositional attitude purportedly essential to fear) requires the concepts of danger and looming, which also require concepts of potentiality and harm. The problem with this requirement is that we do not think infants and non-human animals possess these kinds of concepts, yet we

observe states in them that seem highly continuous with our own emotional states (Zajonc 1984). Lazarus has responded to this objection by claiming that for all we know, infants and animals could actually be making cognitive appraisals, though presumably this would severely restrict the sophistication of the cognitions involved (Lazarus 1984).

There are other ways in which we may disconnect emotional states from cognitions. For instance, sometimes emotional states are triggered by appraisals; say anger that we have been deliberately deceived, that we later find were mistaken. Even though we now believe there has been no deception, it is possible for the state of anger to remain despite the change of appraisal. If this is the case then the appraisal is not a necessary component of the emotion. The cognitivist may respond that perhaps the appraisal hasn't in fact changed. People are often recalcitrant in modifying their attitudes, even in the face of conflicting evidence. Yet in this case, the nature of the appraisal seems less like a detailed and explicit belief and more like an underlying intuition about one's relation to the world, something that a set of bodily changes could potentially capture.

Perhaps the most convincing evidence against the necessity of cognitions comes from studies showing that the neuro-anatomical processes involved in emotional states can function separately to those involved in conceptual thought. For example, an emotion such as fear can be triggered by a low level representation of a coiled snake-like object. In this case the image is received by the retina which sends a signal via the optic nerve to the thalamus, a part of the brain capable of registering basic visual properties such as shape but not capable of recognising objects. The

thalamus then sends a signal directly to the amygdala, which in turn organises a range of bodily responses such as fleeing or freezing, changes in heart rate, breathing as well as facial expression. The thalamus also sends a signal to the neocortex, which is finally able to recognise the object. When the object is finally recognised, it may turn out to be just a piece of rope. Yet the low level perception of the object is enough to trigger to full range of fear responses *before* the neocortex becomes involved (LeDoux 1998).

Now someone may insist that these responses are not true examples of *fear*, but rather something like an ‘affect programme’ (cf. Griffiths 1997). Yet it seems that the subject would be able to report afterwards that it felt like fear, because all the bodily changes that generated the phenomenal feeling were present. Moreover, the subject would have reacted in a fearful way; their heart rate would have gone up, they would have frozen or run away and so on. So from both a first person and third person perspective there would be good reason to think it was a genuine episode of fear. Then since we can identify both cortical and subcortical fear episodes as fear, there must be something other than cognitive appraisals that is unifying these episodes. The obvious candidates are the bodily changes or feelings involved.

Though there have been doubts about how distinctive or determinate the bodily changes involved in emotions are (e.g. Robinson 2006: 208), in general our ability to identify specific biological correlates for various emotional states is improving all the time. This is particularly true since neurological states have begun to be recognised alongside other bodily correlates. Thus Damasio boldly claims,

There is nothing vague, elusive or nonspecific about emotional responses, and there is nothing vague, elusive or nonspecific about the representations which can become feelings of emotions. The substrate for emotional feelings is a very concrete set of neural patterns in maps of selected structures. (Damasio 2000: 282-283)

In order to justify this view, Damasio cites various neurological evidence showing that specific areas of the brain underlie particular emotional states. For instance Damasio describes the case of 'S' who as a result of damage to her amygdala lost the ability to feel fear. In this case the reasoning skills of S were intact except for her ability to judge faces as more or less trustworthy in a way that other people were able to (Damasio 2000: 62-65). Furthermore she was unable to recognise faces as expressive of fear as well as imitate these expressions, though her performance was unimpaired for other emotions such as anger or surprise. Now, the amygdala controls bodily responses, it does not seem to be involved in anything like cognitive appraisal. So it seems clear that this system for registering and organising bodily changes is most essential for the emotion.

This argument is additionally backed up by empirical evidence (to which James also refers) that when people have lost sensational feelings due to spinal injuries, their ability to experience emotions has been correspondingly limited.<sup>3</sup> However, the evidence here is conflicting and it is not the case that the subject's emotions have been *completely* knocked out. Hence there is some debate over whether what

---

<sup>3</sup> Particularly relevant for my purposes, James cites the case of a woman who appearing to have lost her proprioceptive sensations and as a result her emotional responses, reports that she has also lost the capacity to enjoy music (James 1884: 200).

emotional capacities remain are as a result of their judgemental capacities or because other internal bodily systems are still functioning, as well as normal muscular feelings from the neck up (cf. Damasio 2000: 289-290).

There is also another possibility for the non-cognitivist to appeal to here: Damasio (2000, 2004) argues for the existence of an 'as-if body loop' which is a system whereby the brain can create a map of the state of the body independently of actual bodily changes. The existence of this mechanism is plausible because the various bodily changes that generate emotions should be registered and organised at *some* level, or else the subject would never experience a unified emotional state but only a loose collection of distinct responses that did not add up to any particular content. Moreover, the subject may not be able to organise a unified behavioural response to all these changes. Then, since the registering function, what I call the bodily pattern, is not identical with the bodily changes themselves, there is no reason to think that it cannot be activated independently of the bodily changes, in the same way as the neural registration of visual information can be activated independently of the eyes, as in visual imagination. Since the activation is not supported by actual bodily changes, but presumably by some other part of the brain (Damasio points to mirror neurons in the frontal cortex, which I discuss in chapter two), we may also expect that it will only generate attenuated somatic images, just like visual imagination.

Damasio claims that the as-if loop functions to *anticipate* bodily changes by making the brain imagine or hallucinate that it is undergoing those changes. Susan Hurley (2005) argues for a similar anticipatory mechanism. Since the brain is presumably where the experience of those changes is most directly generated, the mechanism

therefore generates the phenomenal feelings of those bodily changes. The anticipation of the bodily changes involved in *emotional* responses then allows the generation of the feelings involved in those emotions. Hence this mechanism may explain the limited preservation of emotions in subjects with spinal injuries because we can effectively imagine or hallucinate bodily changes. The existence of the as-if loop is also helpful in solving another problem for the non-cognitivist; that we are able to have emotional responses more immediately than we would expect some of the bodily changes to occur (such as hormonal responses) (Damasio 2004: 118). I also cite this mechanism in chapter two because it is strongly associated with our capacity to simulate *other* peoples' emotional states. All these capacities make the development of an as-if loop particularly useful for an organism, and so intelligible from an evolutionary standpoint.

Meanwhile, we have grounds to suppose that bodily changes are necessary for the existence of emotions, but are they also sufficient? Some evidence that they are comes from the fact that drugs that alter bodily states such as alcohol or adrenaline, can correspondingly affect emotional states. In addition there are experiments where psychologists find ingenious ways to force participants into different facial expressions in order to see whether a bodily change can induce an emotional change. One example involved getting subjects either to grip a pencil between their teeth- producing a grin, or hold it with their lips- producing a grimace, and then had them rate cartoons. The grinning subjects tended to rate the cartoons as more amusing, indicating that their mood was elevated by their facial expression (Strack, Martin & Stepper, 1988). In a similar experiment conducted by Zajonc, subjects were given two different stories to evaluate. One of the stories contained lots of 'ü' sounds,

which forces the face into a slightly frowning expression where the other contained more 'ee' sounds which produces more of a smile. Subjects correspondingly rated the 'ü' story as less pleasant (Zajonc, Murphy, & Inglehart, 1989).

It has been repeatedly shown that mechanically adopting the familiar muscular positions associated with the basic emotions can affect autonomic nervous system arousal levels as well as the subject's own stated reactions. It is this fact that seems to underpin the phenomenon of emotional contagion, where by unconsciously imitating the bodily attitudes of others we come to adopt their emotional states as well (Hatfield et al. 1994, see also chapter two). It is this kind of consideration that grounds James's famous declaration that "we feel sorry because we cry, angry because we strike, afraid because we tremble" (James 1884: 190). The bodily states involved seem to cause the emotional states rather than the other way round. It is hard to see how these bodily states could cause emotional states only by way of cognitivist style appraisals.

### **Feelings with Meaning**

Thus bodily changes and the bodily patterns or feelings that register them seem to be intrinsic to emotions, whether they are generated by physical movements and muscle tension, or more visceral changes. However, even non-cognitivists agree that this is not all there is to emotional states. Jesse Prinz states for example:

There is a deep intuition that emotions are meaningful. They are not simply arbitrary feelings. Instead they inform us about our relationship to

the world, they embody our convictions, and they factor intelligibly into our decisions in life. (Prinz 2004: 16)

Moreover there are all kinds of bodily patterns that we undergo such as pain or tiredness that we do not class as emotional states. On their own, these bodily patterns are not sufficient for the occurrence of an emotion. Rather, what seems to determine whether these patterns are emotional patterns is the additional *context* in which they occur. For instance, a sensation like tiredness would have quite a different significance if it occurred just after having broken up with one's girlfriend rather than after completing a long run. Thus we are driven towards incorporating some aspects of the cognitivist position. Yet this need not require an appeal to additional cognitive components so much as recognising the more complex role that bodily patterns can play.

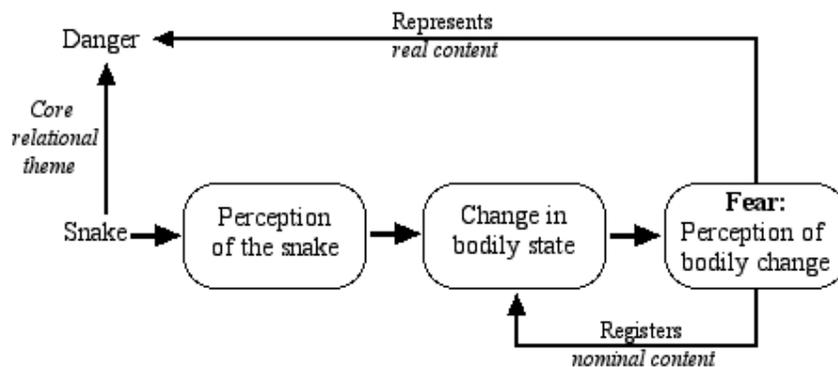
In order to account for the meaningful aspects of emotions, Prinz appeals to Lazarus's appraisal theory and its notion of core relational themes. A core relational theme is a relation between oneself and the emotion elicitor that bears on well-being. For example, the core relational theme represented by anger is 'a demeaning offence against me and mine'. Happiness is 'making reasonable progress toward the realization of a goal'. Sadness is 'having experienced an irrevocable loss' and anxiety is characterised by 'facing uncertain existential threat' (Lazarus 1991). The wording of these descriptions may not be exactly appropriate to every emotional episode, but they seem to capture the gist of our emotional categories. They target what Prinz calls the *formal* object of an emotional situation, which remains constant between particular object situations (Prinz 2004: 62-63, cf. Kenny 1963). As such

they should be included within an analysis of the emotional state itself rather than as causes or effects of that state.

The key difference between Prinz and Lazarus is that Prinz argues that bodily patterns rather than cognitions are able to represent these core relational themes. He claims “emotions are intentional in their own right, independent of any representations that happen to accompany them” (Prinz 2004: 62). They are thus what he calls ‘embodied appraisals’.

In order to make sense of how the registration of bodily changes can be appraisals, Prinz makes use of Dretske’s (1981, 1986) notion of representation. According to Dretske, representations are states that carry information and which are also capable of being in error. If a state reliably co-occurs with a situation, then it can represent that situation. However, it must also be the case that the state in question has been set up by learning or evolution to have the function of representing what it represents. Otherwise smoke would represent fire because it is caused by fire and thus reliably indicates the presence of fire. In the case of emotions then, it would be rather odd if our emotions were set up by evolution simply to register bodily changes. What difference to a creature’s survival could it make to register a heightening of muscle tension for instance? Hence Prinz concludes that although our emotions reliably indicate both bodily changes and core relational themes, they are only set up to *represent* core relational themes, due to the utility of these appraisals for the organism.

Note that Prinz is not arguing for any kind of *transitivity* of representation here. Bodily patterns do not represent core relational themes by *representing* bodily changes. This would make Prinz vulnerable to counter-examples in which by representing clouds we also represent the chemical processes on the water surfaces of the earth that cause them (Hatzimoysis 2003a: 109). Rather bodily patterns represent themes by *registering* bodily changes. Compare this to visual perception: The visual representation of an object results from the stimulation of various lower level receptors for colours, light intensity and edges which we eventually perceive as a unified visual scene. However, it is not the case that our object perception represents this lower level activity, which then represents the object.



Prinz's perceptual model of emotions (2004: 69)

So the subject perceives the core relational theme in virtue of bodily changes, just as one may see a tree in virtue of retinal activity. The bodily pattern that registers these bodily changes is then analogous to the visual centres in the brain. So the bodily pattern and bodily changes supporting it *constitute* rather than *cause* the emotional state. Prinz then describes core relational themes as the 'real contents' of emotions (their representational meaning) where the bodily changes are their 'nominal contents' (like the words used to represent that meaning) (Prinz 2004: 68).

Now we might observe here that if the bodily pattern in the brain can independently generate the experience of an emotion, why not say that this is what really constitutes the emotion and not the bodily changes that it registers? First of all, I think there is an issue here concerning wide and narrow content. Whilst it is true that the bodily pattern is necessary to all cases of emotions and in some cases sufficient, it is not sufficient for *all* or even most emotional states. In particular the bodily pattern cannot by itself distinguish between anticipated, currently ongoing, or as I show in chapter two, simulated emotional states.

Second as I mentioned above, the bodily pattern is likely to be highly attenuated where it is not supported by constant bodily changes. In general, the various bodily changes are what generate and maintain the content of the emotion, where the bodily pattern merely unifies that content. The bodily pattern is like the manager of a company, where the work done by the company as a whole is what constitutes the mental state. Certainly the bodily pattern is a very important stage in the system where various content converges. This is an important step before going on to interpret that content, direct the attention of the subject, stimulate behaviours and connect to other mental states. Yet the bodily changes are doing most of the representational work for that mental state.

There is an analogy here with visual perceptions. It might be possible to have something like a visual experience when only the visual centres in the brain are active. Yet the functioning of the eyes is still needed to distinguish between real and imagined visual states. Moreover constant interaction is required between the visual

centres and retinal activity to generate and maintain the complete visual representation. Hence we should at least include retinal activity as part of a *system* that constitutes the visual state.

If we take a direct realist view of perception here, then strictly speaking, only the system as a whole (including the core relational theme) is the emotional state. In this case, all the various stages of activity are *supporting* that mental state (by generating, maintaining or even distorting content). Yet this activity is still distinct from the vehicle of the mental state, such as the body or brain. So it is misleading to say that the various stages of activity are merely the causal substratum of the state. Rather the various stages of activity are where the content is instantiated. So they should be included in an analysis of the mental state in a way that the basic physical substance of the body and brain should not be. These considerations will be important when in chapter five, I argue that music can partly constitute a musician's emotional state.

Having fixed the 'nominal' contents of emotions, what kinds of things can count as instantiating the 'real' contents of emotional states?<sup>4</sup> One of the observations motivating the cognitivist view was that emotions do not just respond to objects in the world, but also to cognitively complex states like beliefs or imaginings. However according to Prinz's theory, these states remain as causes rather than components of the emotion. So in order to show how emotional responses to beliefs or imaginings are equivalent to responses to direct environmental objects, Prinz describes a process

---

<sup>4</sup> Due to my direct realist intuitions, I am uneasy about this distinction between real and nominal contents, which has indirect realist implications. Unfortunately I do not have the space to attack this distinction here. I would just like to suggest that perhaps the system of bodily patterns provides the subject with *direct* access to the core relational theme. So at the perceptual level, the actual core relational theme is part of the subject's experience. See also my discussion of the appropriateness of specific bodily patterns to their contents below.

whereby emotional states may become calibrated towards more elaborate cognitive causes. This calibration process allows emotions to become more sophisticated as well as more or less warranted or irrational. In the most basic cases, we are innately programmed to respond emotionally to certain causes, such as fear in response to loud noises, looming objects or the sudden loss of support, all of which are reliable indicators of imminent danger. However, as we develop intellectually and have more experiences of fear, we come to unify the features that elicit that emotion using the concept of danger. Through repeated association, the concepts that were originally only associated with the emotional state can become causes in their own right. At this point an emotion of fear may be triggered by a realisation that we are in danger (which accesses that concept) even in circumstances lacking the original perceptual elicitors (Prinz 2004: 76).

Prinz also claims that higher social emotions such as jealousy are the result of more basic embodied appraisals becoming recalibrated towards new *types* of situations. When a bodily response (such as anger) becomes set up to occur as a result of judgements of infidelity, the emotion may now represent infidelity as a core relational theme. Hence similar nominal contents may come to track different real contents. Another important consequence is that it is possible for different cultures to encourage the calibration of different higher emotions, whilst still being ultimately founded on the same basic bodily responses. Prinz describes a situation whereby cultures may ‘hypercognize’ a particular sort of emotionally eliciting situation, and distinguish a new emotion unique to that culture as a result. All that is required is that the new emotion reliably serves to pick out situations of a particular formal kind (Prinz 2004: 99, 141).

So is it the case that all emotional bodily patterns involve appraisals of some situation (real or imaginary)? One possible counter-example is moods, which usually seem to lack any particular content. A possible explanation for mood is that we are not aware of the real content of the emotional state. So for example a bodily pattern informs us that we are anxious about something, yet the real cause is hidden to us. However, Prinz offers a different explanation; that moods are appraisals about how things are for us in general, and that they accordingly represent our more long-term goals (Prinz 2004: 182-188). Hence moods count as genuine emotional states, though they form a distinct sub-group within the class of emotions in virtue of their more generalised objects.

Yet even if we restrict their content to the formal or general aspects of situations, the idea that bodily patterns represent situations may still seem implausible. This is because bodily patterns may just seem insufficiently complex to accurately represent those formal aspects. Prinz notes however that bodily states need not be as complex as the core relational themes that they represent, they simply have to be set up so as to reliably track these formal objects. Prinz uses the analogy of the devices placed in cars to warn drivers of the presence of police radars. The device emits only a simple beep, but in doing so is able to represent a complex situation. Similarly, a bodily state does not have to be able to *describe* the core relational theme it represents, nor does it necessarily have to be preceded by any cognitive representation of comparable complexity (Prinz 2004: 65). Hence bodily patterns are able to intentionally represent core relational themes despite their fairly unstructured character.

Whilst Prinz's argument here seems satisfactory for explaining the overall capacity of bodily patterns to represent the formal aspects of a situation, it does leave a puzzle as to why the particular patterns represent the situations that they do. That is, how are the various emotional states distinguished as belonging to this or that formal object? Is there anything about the state itself that makes it appropriate to target that sort of situation? Certainly our phenomenal feelings do not *seem* like arbitrary or merely conventional signs of the situations they target. Rather there seems to be something distinctly sad about the feeling of sadness, and angry about the feeling of anger. We might respond that the appropriateness of these feelings may just be a consequence of having had the association between the feeling and the situation entrenched by experience. Yet feelings are not like words, where for instance, we can repeat a word to ourselves until it becomes meaningless. Rather repeatedly imagining the feeling of fear just underlines the rightness of that feeling for that situation.

The appropriateness of feelings and the bodily patterns that underlie them seems to be a result of their *natural* meaning, having acquired their function for evolutionary and biological reasons rather than arbitrary convention. The most obvious way to then explain why the various patterns have their natural meaning is to look at the particular behaviours that they generate, at least in the most primitive instances of emotions. For instance, we saw how in the case of fear the amygdala caused the subject's heart and breathing rate to increase, as well as cause him to freeze or to run away. These reactions are appropriate ways to deal with a situation that merits fear, since they either prepare the subject for a certain response (i.e. by increasing the

amount of blood flow to the muscles) or are the response itself (i.e. avoiding something dangerous).

Of course we recognise that the subject need not necessarily act on their emotional state. In more developed cases of fear, such as a fear of public speaking, it may be completely inappropriate to run away. For this reason most theorists do not include the behaviours, or even the tendencies to certain behaviours as essential to emotional states (and why not also the tendency to suppress the behaviour?). Yet as I argue in detail in chapter two, this will not stop us from including the *bodily changes* involved in those behaviours or the preparation for those behaviours as essential to the emotional state. Moreover, it seems equally plausible for the bodily changes involved in *suppressing* certain behaviours to be part of the content of emotions. For instance, the tension involved in stopping oneself from striking could become part of the characteristic bodily pattern of anger.

Furthermore I would dispute the notion that patterns of bodily changes have an unstructured character. First of all we can compare the relative levels of tension in the various emotional feelings, ranging from a calm flat feeling, to a sense of vibration, to explosive bursts of energy. We can also look at the various parts within the bodily pattern according to where it is located in the body and the range of simultaneous changes. Secondly it is important to emphasise that emotions are not just static representations, but constantly iterated dynamic processes, responding to an ever changing environment as well as our actions within that environment. I am not just afraid by the proximity of the bull, but also the fact that it is charging towards me. Hence we must also look at the overall temporal profile of that pattern.

Here we can distinguish the way that the different emotions begin and end, such as whether they fade in (like sadness or anxiety) or involve more sudden shifts (such as surprise and disgust). We can also see whether once the bodily pattern has become established, it remains relatively steady, comes in waves or pulses, or has contrasting sub-sections. This is why I call the emotional state a bodily *pattern* as opposed to say a bodily ‘colour’. Like colours, emotions may contrast with each other along various dimensions. Yet emotions also have distinct temporal and spatially oriented *parts*.

Overall these kinds of profiles might not enable us to pick out a small number of distinctive emotion types. Rather we may only see gradual differences in degree along the various dimensions. Yet there is certainly enough contrast here to reflect all kinds of dynamic structural patterns. Then when we look more closely at the perceptual nature of emotions, I argue that we can more accurately specify what it is these patterns reflect.

### **Emotional Perceptions**

So far I have explained why bodily patterns act like perceptual states of the body, relating the subject towards certain aspects of the environment. However, how far can this analogy with perceptual states be pushed? Prinz says for instance, “we can feel the offensiveness of external situations resonating through our flesh” (Prinz 2004: 227). Yet I am unsure whether this characterisation is phenomenally accurate. I might certainly feel that my anger is a reasonable response to the situation I am faced with, but the bodily pattern of anger seems to be experienced as occurring within (Prinz 2004: 61). In contrast, the contents of other perceptual states, such as the perception of colour, place those qualities as existing out there in the

environment. One possible reply here is that we are more apt to locate the emotional content within our bodies simply because emotional reactions to situations are so subjectively variable. For instance we can ask “is it me, or is it hot in here?” and thus locate the sensation of heat accordingly. So perhaps if emotional reactions were more stable across different observers, we might be more inclined to experientially locate their content out in the world.

Another possible disanalogy between emotions and perceptions is that emotions integrate the information from several sense modalities and involve a sense of the organism to which they belong. So in possessing dedicated sense organs and a more exclusive focus on the environment, our regular perceptual faculties are more direct and less relational than emotional states. However, some perceptual states do in fact register relational properties between organism and environment. Prinz uses the example of the perceptual quality of being ten feet away from something. In this case, there are many distinct sources of information (such as colours, heights, the sharpness of lines etc.) that are organised to form the perceptual content. Hence it is not the case that perceptions are always more direct than emotional states (Prinz 2004: 226). In addition, the phenomenon of synaesthesia, where people report seeing sounds and hearing colours, shows that it is possible to receive perceptual sensations via other sense modalities (Prinz 2004: 231). Furthermore, despite the fact that emotions organise information from several sensual faculties, they do have dedicated input systems in virtue of the reception and organisation of bodily changes in the brain. Just because emotions don't have their own sense organ on the *surface* of the body, it doesn't entail that they are not a distinctive perceptual faculty.

From the point of view of the conscious subject, a more significant similarity between emotions and other kinds of perception are that both are perspectival and both can be hallucinatory (as when we fear that a situation is dangerous when in fact it isn't). So as well as emotions having the same kind of epistemic status as other forms of perception, the information they provide also seems to share the same structure. To elaborate; both emotions and other perceptual functions have an initial stage of collecting raw data (such as edges in vision or hormonal changes in emotions). There is then a stage at which this data is organised, or bound into a unified picture, placing the subject at the centre of his experience of the world. Finally, there is a conceptual stage in which the object of the perception is recognised or stored apart from any particular perspective. For instance, the abstract idea of what sadness feels like or what a door looks like that enables us to recognise the object as an instance of that category.

Again, in drawing this comparison with perceptual states, it might seem problematic that we can respond emotionally to purely imagined scenarios or considerations. This is because if a person imagines a situation that then triggers an emotional reaction, it seems that the bodily pattern is not *mediating* between the subject and the object anymore in a truly perceptual manner. However, we can accommodate this more complex case by saying that the perspective on an imaginary object is transformed when the bodily pattern gets involved. This is similar to engaging a different sense modality (such as touch instead of sight) and as a result getting a new perspective on an object. In the same way an embodied appraisal transforms the imaginary situation into an emotional object by revealing its formal property or core relational theme as

an emotion elicitor. Hence bodily patterns may still mediate between the person and this newly appreciated emotional object.

However, the case of having a *genuine* emotion triggered by a purely imaginary scenario seems quite unlike other perceptual faculties. This is because of a general representational feature of our perceptual imaginations in which the only details we perceive in an imagined scene are details that we must consciously add to that scene. We cannot for instance, ‘discover’ the redness of the clothes that someone is wearing in our imagined image, or how hot it is, or how far away we are from an object, or how noisy it is, without making up those details as we go along. In contrast, we can ‘discover’ a new emotional response.

Prinz responds to this objection (personal communication) by claiming that there are some instances in which one may discover a perceptual quality as a result of imagination. He cites the case of imagining oneself standing on a high ledge, and then feeling a sense of wobbling (i.e. a proprioceptive perception). Now I am not entirely convinced that this isn’t a motor response rather than a perceptual response. Neuroimaging may resolve which is the case. Either way, the responses are still something focused on the body rather than outer representations of the world. It is comparable to feeling an involuntary twinge of pain as a result as imagining knee surgery. The body is reacting to maintain its integrity in some way.

Overall, even if we have to say that emotions are unlike other perceptions in this particular respect, it still seems a strong case can be made for treating emotions as essentially another kind of perception. Yet I think that Prinz undermines the

perceptual nature of emotions where the core relational themes that they track are understood in rather conceptual ways. Instead, emphasising the strong connection that emotions have with physical movement and actions reveals what it is that emotions perceive in a more concrete way than the kinds of core relational themes that Prinz and Lazarus refer to. All perceptual functions help to locate the subject within his environment. Similarly, I claim that emotions locate the subject within his dynamic response to the world, or in other words; how things are going for him. This is not to say that emotions represent *all* kinds of shifting relations between subject and world, otherwise this would include clearly non-emotional situations such as walking around. Rather the dynamic relations involved are ones that impact on the bodily integrity or capabilities of the subject. This is the causal principle that makes them *dynamic* changes as opposed to changes *simpliciter*. For example, the phenomenal feeling of joy is a feeling of energetic capability. It represents the world as providing certain freedoms of movement, as something we can run, dance or skip through. We may then literally respond by running or dancing or skipping, or by more generally acting in such a graceful or energetic manner.

In my view then, we should understand the core relational theme for an emotion like anxiety not as ‘facing uncertain existential threat’ but rather something more like the physical sense of losing support. The reason for this is simple; it is a simple way to sum up what fear actually feels like. In primitive cases of anxiety there may literally be a feeling that one is about to fall. But due to the same calibration that Prinz appeals to, what counts as support, and losing that support can be construed in more abstract ways. So one might be anxious about losing one’s job, in that the job acts as something that ‘supports’ one’s long-term material well-being.

There are even dynamic qualities to our thought processes. For instance, Malcolm Budd notes that, “[a] depressed state is characterised by relatively slow and confined mental processes, lack of energy, lack of determination, passivity, and difficulty in changing one’s state” (Budd 1995: 207). Equally, one may spend a sleepless night feverishly turning over in one’s mind different ideas about what will happen the next day. Or when one is ecstatically happy, one may be unable to concentrate on anything for more than a few seconds. Thus emotional representations, even when dealing with highly conceptual processes, seem to be well characterised by these dynamic qualities or contrasts.<sup>5</sup>

This kind of understanding of the perceptual content of emotional states would then certainly affect the way we categorise and distinguish one emotion from another. For instance, we may class together various cases of fear as losing support either gradually (anxiety), or more suddenly (panic). But there may also be other kinds of fear that are better classed as the sense that something is approaching that may strike you, (which again could be understood either literally or more abstractly). Technically then, these would be two distinct emotions, though we might still place them close to one another in that they satisfy a more general sense of ‘losing stability’.

Prinz’s final position is that emotions are *valent* embodied appraisals. He argues that the broad characterisation of emotions as either positive or negative corresponds to

---

<sup>5</sup> Dynamics could also include a sense of modal contrasts rather than purely temporal contrasts. Consider for instance the case of thankfulness. When one feels thankful, one’s actual situation may not have changed at all. But one may still consider how things *might* have been worse, and thus a sense of things proceeding smoothly in contrast to that possibility.

additional *markers* given to the emotion which dictate either approach or withdrawal behaviours towards their elicitors. Prinz admits that these aspects of emotions are not perceptual because they do not register bodily states so much as issue imperatives for action (Prinz 2004: 228). I am not so keen on the idea that emotions have simple valence markers such as these. I concur with Robert Solomon (2004a) that emotions have a variety of dimensional opposites such as healthy/unhealthy or virtuous/vicious rather than simply positive or negative. Accordingly, approach and avoidance behaviours would be just one of these opposites. Moreover, on my view of the content of emotional perceptions, a sense of oneself approaching or withdrawing is readily incorporated as a dynamic aspect of the bodily pattern itself.

If we make this link to driving certain behavioural reactions, is it the case that in helping the subject to decide how to proceed, that embodied appraisals are more cognitive in character than our other perceptual states? In synthesising the cognitivist and non-cognitivist positions, Prinz may also have blurred the boundaries between the notions of sensation and cognition or percept and concept. The problem is that it is not especially clear how we should define cognition. Does cognition apply to any form of 'thinking' (conscious or not) or does it require the use of concepts? It would seem that making a judgement or evaluation is a paradigm cognitive process and embodied appraisals seem to fall into that category. Yet even Robert Solomon, who was erstwhile one of the more extreme cognitivist theorists, said recently that bodily states can potentially be characterised as judgements because 'knowing' something is often a matter of habits and practices performed rather than consciously articulated discriminations (Solomon 2003b).

Prinz defines cognitive states and processes as “those that exploit representations that are under the control of an organism rather than under the control of the environment” (Prinz 2004: 45). He concludes that emotional states are percepts rather than concepts because they are automatic responses over which we have no control. That is, they are processes that work from the ‘bottom up’ in reacting to the environment. There is some debate over the possibility of Machiavellian or pre-meditated emotional states (e.g. Solomon 2003b, Griffiths 2003), which on Prinz’s terms would indicate a degree of cognitive control. We can for instance, deliberately focus on an injustice in order to ‘work up a rage’ or allow ourselves to become upset in order to gain sympathy. Prinz seems to admit this much, yet he maintains that whilst it is possible to conceptualise emotions, the everyday experience of emotions need not involve any concepts. We can imagine being angry, recognise a state of fear or remember a happy episode. Yet we can equally conceptualise a state like seeing red without undermining its basic perceptual status.<sup>6</sup>

My own position is that emotional states are genuinely perceptual but are more sophisticated than our other perceptual abilities. In the following chapter I present reasons for thinking we can potentially cognise our emotional states in a way quite unlike any other perceptual function. However at this stage, we can recognise that because emotions organise and respond to information from several sense modalities at once they seem more like a ‘meta-modality’ than a regular perceptual function. Most of all though, in encouraging certain behaviours and providing information about well-being, emotions seem to *evaluate* the world rather than simply map it. In

---

<sup>6</sup> According to Prinz (personal communication), recognition counts as perceptual because it is typically a passive reaction to the world. He qualifies however that it often uses resources that in general underlie cognitive processes (i.e. concepts). Hence again, there is some ambiguity in the distinction between perception and cognition.

this way the emotions signify the beginnings of a unified self. They seem occupy a middle ground between pure sensations, like colours or pain, and more definitive mental states like beliefs, memory and the imagination. Thus we can appreciate what de Sousa means when he says that, “emotions are like Descartes’ pineal gland: they function where mind and body most closely and mysteriously interact” (1987: xvi).

The dichotomy between cognitive and non-cognitive is no longer particularly useful when traditionally non-cognitive states can occupy the role of cognitive processes. We may similarly have to accept that there is a rather vague boundary between feelings and cognitions, percepts and concepts. One of the guiding themes of this thesis is the extent to which our mental processes are actively engaged with, and potentially externalised or embodied within the environment beyond the individual. Both emotions and other forms of perception do not passively construct a picture of the world but rather actively interrogate it. As I argue in later chapters, we typically perceive the world in terms of affordances for action and our cognitive processes are often integrated with our physical manipulation of the environment. Understanding emotions as perceptions is just the first move away from viewing the mind and emotions as purely internalised cognitive processes.

### **Awareness**

In order to round off this analysis I would like to explore the role that awareness plays in emotional states. So far I have maintained a distinction between bodily patterns and the phenomenal feelings of bodily changes, yet this distinction must now be properly justified. Intuitively we might think that awareness is essential to all emotional states. When we talk about our emotions for instance, we usually mean the

way things *feel*. However, throughout this chapter I have been emphasising a more functional understanding of emotional states. As such, I argue that as long as bodily changes occur in response to the appropriate objects, and are neurally registered in such a way as to orchestrate at least some of the characteristic behavioural responses, (including emotional expressions) then the emotion has been sufficiently distinguished for us to suppose its existence.

Consider the following example: You are facing a firing squad, looking down the barrels of the guns pointing at you and waiting for the order to fire. You clearly recognise the imminent danger, yet you do not consciously experience any feelings of fear. You might even congratulate yourself on being so incredibly brave. Yet when you look down at your body you realise that your body is shaking and you have wet yourself. Accordingly you realise that actually you are terrified and you simply didn't experience it at first. Now in this case, your body is representing the danger well enough. Perhaps we could argue that this is a mixed emotion, and that in some respects you are calm as well as terrified. Yet even if it is a mixed emotion, you are still partly undergoing fear that you are not conscious of.

There are also times when we undergo an emotional response and then become unconscious of it for a while simply because our attention is fixed elsewhere. Imagine for instance that you are slumped in your chair, weeping over the death of your dog, when your eye happens to catch the shopping list you made earlier and you find yourself trying to remember the things you need to buy. A moment later your attention is drawn back to your dog and your sadness. During this period of distraction, tears still streamed down your face. Moreover, none of the bodily states

you are experiencing seems to have changed from what they were just a few moments before. So if one is still acting in an emotionally expressive way, and none of the elicitors of the emotion have changed, then there is no good reason to suppose the state ceased to exist whilst you were unconscious of it.<sup>7</sup> Jesse Prinz takes a similar line here. He states:

There is no reason why one cannot represent core relational themes unconsciously. They can prepare us for behavioural responses, they can initiate thinking processes, they can embody cultural values, and they can motivate moral conduct. None of these effects is intrinsically bound to consciousness. (Prinz 2004: 202)

There is also empirical evidence in favour of unconscious emotional states. An experiment by Strahan et al. (2002) found that people were more likely to choose upbeat music to listen to after having being subliminally presented with pictures of sad faces. Yet when asked, the subjects did not report any change in their mood (Strahan, Spencer & Zanna 2002, cited in Prinz 2004: 203). In this case then, people appear to be unconscious of their own emotionally driven behaviour.

The possibility of unconscious emotions is also made more plausible when we consider them as perceptual states. There may be emotional states similar to blindsight, in which subjects report no conscious visual experience and yet are able to identify the objects presented to them and successfully navigate around rooms. Of

---

<sup>7</sup> A similar argument is made for pain, which we might think is even more closely tied to its phenomenology. The example that Rosenthal (1991) gives is that one may forget a headache for a moment whilst reading an absorbing novel, though one continues to engage in pain behaviours, such as rubbing one's temples (cited in Prinz 1994: 201).

course, we may achieve varying degrees of detailed awareness of the kinds of situations that typically arouse our emotions, as well as the impact our own personality, or the exact quality of the feeling involved. However, the self-conscious attention to our own emotional feelings is not the same as the basic occurrence of an emotional state. The idea of understanding emotions as perceptions is that we represent the emotionally eliciting situation *in* making an embodied appraisal. Hence there is a minimal level of awareness, or intentionality that simply is the bodily pattern. Though this is not necessarily the same as there being something that it is like, a particular phenomenal quality for that emotion.

Overall, Prinz believes that consciousness targets just one stage of the neural processing of emotion. As I mentioned earlier, emotions are structured like other perceptual states in having three stages of processing; initial data, organised perspective, and then conceptual recognition. Prinz then argues that the conscious awareness of emotions tends to target the point when the various bodily changes are organised into what I've called the bodily pattern. Consciously experiencing this overall pattern rather than the individual components then facilitates discrimination of the emotional state. It also relates the subject to the emotional elicitor in an immediate way that can motivate action. Yet focusing the attention of the person onto the bodily changes or environmental situation counts as only one of the characteristic responses to an emotional state rather than an essential element.

However, this is not to say that the focus of attention cannot have a significant impact on how the emotional state develops from that point on. One theory in particular accords awareness a central role in deciding *which* emotion one is

undergoing. Schachter and Singer's labelling theory (1962) holds that our bodily feelings must be interpreted or labelled by us in order to determine what emotion they represent. They argue for this theory based on an experiment where subjects were unknowingly given adrenaline and then placed into two different situations, one involving an actor becoming enraged by a questionnaire, the other involving an actor playing childishly. The idea was that although subjects were in the same physiological states, their emotional states would differ according to the subjects' different interpretations, which had been primed by the expressive behaviour of the actors. The emotional states of the subjects were gauged by observing their expressive behaviour as well as their self-reports after the experiment. Those in the anger-primed situation were observed to agree with the actor who expressed anger and those in the amusement-primed situation showed amusement.

However in both cases, subjects afterwards reported their own states as amused. Schachter and Singer dismiss these reports as the result of a desire to please the experimenter. Yet it seems just as likely that subjects in the anger-primed situation were trying to placate the actor. In addition, even though the adrenaline injections would put the different subjects in a similar physiological state to begin with, these states may have changed over the course of the experiment. Jesse Prinz also criticises Schachter and Singer's experimental design because it assumes that different expressive behaviours signal different emotional states. Yet it also assumes that different expressive behaviours may nevertheless signal the same *physiological* states (Prinz 2004: 70-71). There seems no reason why expressive behaviour would fail to track physiological arousal like this, so these assumptions not very plausible.

However there is an intuitive appeal to Schachter and Singer's labelling theory. In my own experience I have often had feelings that I have been unable to definitively label one way or the other. By judging the situation more deliberately, I have then decided (or possibly misrepresented) for myself what emotion that feeling signals (cf. Goldie 2004: 93). Alternatively, when we are feeling tired we may be more easily upset or moved to anger. In these cases an unemotional feeling is transformed into an emotional appraisal based on the slightest excuse. So even if it is not the case that the conscious interpretation of feelings may differentiate the very same feelings as signaling different emotions, interpretation may still encourage congruent appraisals that can then arouse more distinct emotional states. This is also similar to other perceptual states, where for instance believing that one is looking at a person can cause one to interpret vague visual information to fit that interpretation, and thus identify eyes, a nose and so on even if one is not in fact looking at a person.

Another consequence of conscious awareness of one's bodily states is that it is possible to have emotions *about* one's bodily states. For instance, I can be sad that I am sick. A more dramatic case are panic attacks, in which awareness of the heart beating faster causes the subject to panic that they are having a heart attack, which causes their heart to beat faster and so on. So an unfortunate feedback effect is set up that usually can only be resolved by consciously controlling one's breathing rate. Future cases of panic may even be set off by a judgement that one is beginning to panic. Here we see one example of emotional recalibration, i.e. panic about panic, that although unwarranted seems to necessarily require the conscious attention of the subject to occur.

So it seems that consciously attributing an emotional state to oneself can potentially arouse or alter an emotional state. Moreover, the attention of the subject can focus on different aspects of the emotional state, such as the bodily feelings, the situation, the long-term context (e.g. is this love?) or how it reflects one's overall character (e.g. am I a brave person?). Awareness can help to link these components together as well as to cause errors of misattribution. It is also certainly *useful* to us when emotions grab our attention because they cause us to focus our resources on dealing with the situation at hand. As such the role of awareness is another sign of the dynamic and iterative nature of emotions in locating ourselves within our relation to the world and helping to guide a response. However, it is quite possible for emotions to function independently of the awareness of the subject. In general emotional states are distinct from the phenomenal experience of those states.

## Chapter Two: Expression and Empathy

In the first chapter I argued that emotions are bodily patterns that perceptually represent the subject's dynamic relation with the world. In this chapter I develop this theory further and explain the nature of emotional expression and empathy. This is required for my explanations of how music can express emotions in the following chapters. Understanding how this is possible depends on how much musical expression can instantiate the same roles as ordinary types of expression. Thus my purpose here is to establish how ordinary emotional expression works, unmediated by special tools or training. In particular I am interested in how exactly expressions of emotion relate to the emotional states that they express. This involves finding properties or essential functions that all expressions share (given that expressions form a single coherent class). Possessing these properties should then explain what makes expressive behaviours expressive.

Overall I argue that expressions of emotion such as facial expressions, non-verbal utterances and other bodily gestures are simply ways in which emotions happen, or surface level *modes* of the emotion itself. That is, we should consider expressive behaviours to be as much part of the emotional state as more internal bodily changes. The main reason for this is that behavioural expressions function in the same way as somatic changes to generate bodily patterns and feelings. Behavioural expressions can even replace more internal somatic changes as the main source of the subject's emotional experience. As such, the role of expressions in communicating one's emotion to others is *secondary* to this more central characteristic. This is not to say that the evolutionary reason for the development or preservation of expressive

behaviours is not dominated by their communicative role. The point is just that expressive behaviours are able to serve this role so well because they are parts of the emotion itself. At the same time, the fact that expressions are a means to generating emotions has important implications for our ability to directly control our emotional states. As a result I argue that emotions can sometimes be cognitive rather than purely perceptual states.

Having established what expressions of emotion are, I then explore the way that they allow other people to recognise our emotional states. The central claim here is that when we perceive another person's expressive movements, it triggers a simulation process in the brain, whereby that movement is processed from a first-personal perspective. As a result, seeing a person making an emotional expression causes us to simulate making that expression ourselves. Since expressions of emotion are part of emotional states in helping to generate bodily patterns, a *simulated* version of the emotion that the other person is undergoing is therefore aroused in the observer. This simulated arousal then allows the emotion to be recognised.

Finally I look at more sophisticated forms of empathy, in which our ideas about the person who is undergoing the emotion and his situation come into play. Here my goal is not to explore the various intricacies of this kind of empathy in detail, so much as show how it relates to the more basic recognition of emotion. This also rounds off my analysis of emotions by identifying some additional factors that can influence an emotional state. Altogether my theories of emotion, expression and empathy then lay the grounds for showing how music can capture emotional states in chapters three and four. In the final chapters of my thesis, the model of emotions will

also be used to analyse simultaneous and related emotional experiences, ultimately enabling us to make sense of the possibility of shared emotions in music.

### **Expression and Emotional Personality**

When trying to account for the nature of emotional expression, I am mainly interested in what are called ‘primary’ expressions. Examples of these are facial expressions and the tonal qualities of the voice as well as other vocal productions such as laughing, crying or sighing. I also include observable changes in posture and muscle tension, and very impulsive actions such as punching the air in triumph, jumping for joy, or clenching the fists in anger. Behaviours of increased attention or reorientation towards the emotional elicitor may also be regarded as expressive of emotion. Though these perhaps reflect the initial perceptual role of emotions more than their expression. Facial, vocal and bodily expressive movements are the most relevant to this analysis because they seem to be the most basic and direct; they need not be specially learnt, and may be found across many different cultures, amongst both infants and adults. Actions such as throwing or hitting something can also be considered in this light, but only in so far as the identity of what was thrown or hit has no special relevance for the person undergoing the emotion.<sup>1</sup>

One of the reasons that Jesse Prinz argues that emotions are perceptions is because they are processes that we passively undergo. In contrast, since emotional expressions are a sort of behaviour they might be described as something that we

---

<sup>1</sup> There are more complex forms of expression where the subject tries to achieve some further end with these actions, such as smashing a favourite vase of someone who has made them angry, or hitting something that symbolises the other person. These more complex activities might be described as partly attempting to satisfy desires that *result* from emotional states (i.e. to injure the offending person) rather than purely expressive.

actively *do*. However a distinction should be made between deliberately using one's behaviour to express an emotion and having one's behaviour reveal or betray an emotional state irrespective of one's desire or intention to do so (cf. Vermazen 1986). Yet a purely binary distinction may be less appropriate than placing expressions along a scale of more or less deliberate behaviours. This is because even within the primary expressions there are several levels at which the motives and intentions of the subject can get involved (cf. Goldie 2000: 126).

At the most deliberate end of the scale there are expressions of emotion that are completely insincere and reflect neither inner feeling nor genuine appraisal, such as a smile to a hated colleague with whom we want to maintain good relations. Other expressions may be genuinely aroused and yet calculated to serve a further end, such as crying in order to gain sympathy. In this case we may simply allow others to perceive the expression, or it may be exaggerated, or more subtly we may make a show of trying to repress the emotion. The ways in which the intentions behind this kind of communicative act may be obscured, distorted or revealed are potentially quite complex.

Moving further towards the less deliberate end of the scale; differences to emotional expression may be caused by genuine attempts to hide our emotions from others when such expressions would result in undesirable social consequences. For example, to laugh when attending a funeral would be socially disastrous in most cases, at least in this culture. This is just one of the ways in which culture may legitimate or encourage certain forms of expression in circumstances that would not apply when we are alone. Another alternative is that an expression may have no further end or

motive and yet may still be permitted or repressed according to one's background emotional personality (regardless of the current social context). For example a self-consciously masculine person may repress the urge to cry when he is sad. This phenomenon should be accessible to cultural influence, which can encourage certain types of emotional personality. Peter Goldie also claims that general emotional character traits may be signaled by habitual expressions or mannerisms. He describes for instance "the permanent expression of disgust at all of human nature in the face of the woman on the bus" (Goldie 2000: 150). This habitual expression is then reckoned to correspond to a generalised appraisal of the world.

Finally we have completely sincere and uncontrolled emotional expression. One of the characteristics of this kind of expression is that it tends not to be self-conscious. In my own experience for instance, I find that as soon as I focus my attention on my own emotional expression, it seems somehow insincere to maintain it in just the same way. Part of the reason for this may be that by revealing the emotion to the person experiencing it, emotional expression can stimulate further emotional reactions. I mentioned in the last chapter how we may become angry that we are sad, or afraid that we are angry. It is also worth noting that non self-conscious expression is more likely when we are alone and so need not worry about revealing our emotions to others.

Given the above possibilities, the way in which expressions of emotion may be communicative, or self-revelatory of emotional states seems to be one of *variables* of emotional expression. It does not seem to be essential to all expressive acts. Rather emotional expressions may just be ways in which emotions happen, or in insincere

cases where they appear to happen. However even if expression does not always occur in social contexts, it may yet have developed for that reason in evolutionary terms. In his famous cross-cultural and cross-species study, *The Expression of Emotions in Man and Animals* (1872) Darwin argued that our expressions of emotion could be explained by their social utility, if not for our species then at least for our evolutionary ancestors. For instance he claimed that the bristling of hair when we are afraid served to make our more furry ancestors look bigger. Similarly to bear the teeth in anger indicates the creature's ability to fight back. Darwin struggled to convincingly explain other expressive behaviours such as shoulder shrugging or crying. However, these may simply have more direct communicative utility. So the purpose of crying is to signal distress, especially for infants, who are utterly reliant on their caregivers and lack other means to communicate. On the other hand, shoulder shrugging seems much more artificial. It expresses 'giving up' or supplication by making a show of raising the shoulders as if preparing to act, but only so that the *release* of the shoulders and thus the effort involved, may be signaled more effectively (cf. Dewey 1894: 568).

Alan Fridlund (1994) also argues that facial expressions are predominantly communicative. So the expression of anger serves as a warning to others, smiling acts as an invitation to approach, and frowning is a sign of supplication. We should also note that some expressions of emotion are very hard to fake. The best example is the difference between a faked smile and a genuine smile, which involves certain muscle movements around the eye that are not amenable to conscious control. Thus expressions are generally reliable guides to the emotional state of the subject. Fridlund also cites the fact that our strongest emotional episodes are not always

expressed. For instance, Olympic athletes often do not smile just after they have won gold medals so much as when they are on the award podium in front of others (cited in Prinz 2004: 111). This would indicate an increased tendency to express emotions in social circumstances, and thus that expressions are generally geared towards a communicative role.

Yet social situations may also lead to more expressive behaviours simply because they tend to be more arousing than solitary situations. In the case of the Olympic athlete, standing on the podium receiving the adulation of thousands of people may well be an even greater stimulus than winning the medal in the first place (and of course, they could be exhausted directly after winning). Paul Ekman (1997) also contests Fridlund's theory by citing evidence that Japanese people are more likely to express negative emotions when in private. Since we may equally have a tendency to *suppress* certain expressions in public, our explanation of the communicative tendencies of expressions will have to be more complex. It would be better to say that our expressions of emotion may be *modified* by our communicative or social roles.

It is still likely that a background awareness of the communicative nature of expressions is leading Japanese people to suppress them more. Moreover, although we often express emotions in private, this need not undermine the claim that expressions have generally evolved for communicative purposes. Their occurrence in social situations may simply have carried over into private situations because it has relatively little costs for the organism to do so. It does not do the organism any harm to express emotions in private, and even in regard to conserving energy, it may

require more effort to suppress emotions in private than just to let them happen all the time.

Yet the point I want to raise is that if we use expressions to communicate our emotional states or related intentions, we may well do so because we recognise their natural meaning as modes of emotion. There is no reason to suppose that understanding emotional expressions as communicative is exclusive or even prior to understanding them as modes of emotion.

Darwin was interested in proving that emotional expressions were universal across different cultures as a means to show that the emotions generating them were also universal. Ekman uses similar reasoning such that if the same expressions are used in the same circumstances then the inner emotion is likely to be the same as well. I am not so much interested in the universality of emotional expressions here as in exploring the link between the outer expression and the inner state. It does not seem immediately evident to me that similar expression guarantees similar emotion, unless that is, expressions directly contribute to the bodily pattern of emotional states.

We have already come across some evidence that may indicate such a relation between expressions and emotional states in chapter one. The experiments in which subjects were emotionally aroused as a consequence of being forced into different facial expressions reveal that expressions can cause emotional states rather than the other way round. This at least suggests that emotional states and expressions have an especially immediate link. James makes a great deal of this kind of evidence and argues that it allows emotional states to be deliberately modified. He claims:

Everyone knows how panic is increased by flight, and how the giving way to the symptoms of grief or anger increases those passions themselves. [...] Refuse to express a passion, and it dies. (James 1884: 197)

James goes on to claim that persistently faking the outward appearance of cheerfulness can arouse the genuine emotion. Hence it is argued that not only can expression intensify an emotional state but can also arouse completely contrary states.

The relation between emotional arousal and vocal expression has also been explored, although not in as much depth as facial expression. In general it is more difficult to deliberately control our voices than our facial expressions (although we cannot turn off our facial expressions the way we can our voices). However Ekman cites the case of one woman who was able to deliberately modify her vocal tone and as a result her arousal level (Ekman 2003: 36). Ekman (agreeing with expression theorist Silvan Tomkins) claims that we are inclined to make sounds whenever our emotions are aroused and that there are different sounds for each emotion (Ekman 2003: 59). As such vocal productions are likely to bear the same relations to emotional states as facial expressions in terms of arousal. Though our emotions may not be so amenable to conscious control via this means.

In general, our ability to control our expressive behaviours may depend on which emotion one is undergoing and which feature of the body it involves. For instance, it

is generally far less easy to control one's behaviours when one is scared as opposed to happy. Similarly in regard to specific behaviours, the facial expression of disgust may be more involuntary than clenching one's fists in anger. Moreover, as in the case of the woman Ekman cites, some people may find it easier to control their expressions, or various aspects of their expressions than others. This may reflect a natural disposition of different people, but it should also be possible given practise to learn better control. This is one of the central methods of cognitive-behavioural therapy for instance.

There have been various studies that aim to show whether expression intensifies or relieves internal somatic changes and whether as a result is it therapeutically beneficial to suppress or encourage emotional expression. Cacioppo et al. (1992) review a series of studies that show that increased expression corresponds to increased measures in various internal systems such as blood pressure, heart rate and skin conductivity. They conclude however, that individuals may differ in characteristic patterns of responses along *all* of these variables. In particular, some individuals show *lower* ANS activity relative to others in response to emotional elicitors, despite showing *greater* levels of facial expressiveness. Other individuals show no visible expression at all, and yet measures of their internal changes show intense activity.

So although it is not the case that expressive behaviour is necessary for all emotional

states, it may be sufficient.<sup>2</sup> Drawing a similar conclusion to Cacioppo et al., Buck (1980) distinguishes between increases of arousal at the level of the individual subject, compared to that between different subjects. He concludes that as our expressions of emotion increase in response to emotional elicitors, in general our inner arousal levels are correspondingly intensified. This is according to the within-subject measure. However, it is not the case that those people who are rated by others as characteristically more expressive tend to undergo greater levels of inner stimulation. Rather the opposite is true. Different people may undergo comparable levels of emotional intensity, yet one manifests it as internal changes where the other manifests it as outward behaviours, and yet another has a more equal ratio of both. A distinction is made accordingly between *internalisers* and *externalisers*, with a category of generalisers in-between.

But what allows us to say that these different kinds of changes (expression and ANS activity) are in fact comparable in terms of emotional arousal? The Cacioppo et al. study derives comparable intensity in terms of the gain, or rate of increase of activity in each variable. Having established the ranges for that variable in the population as a whole, they can then compare the relative extremes at which different people show activity. As a result they are also able to show that some people have lower levels of arousal both in expressiveness *and* ANS activity than others, and so tend to have less intense emotions.

---

<sup>2</sup> That expressions are sufficient for emotional states would indicate that completely insincere expression (where there is no corresponding bodily pattern at all) is actually impossible. As such our understanding of insincere expression should more accurately refer to a conflict between the arousal generated by the false smile (for example) and any other bodily changes or appraisals that the subject may have.

In order to account for the different patterns in which emotional arousal can occur in different people, we may wish to stipulate not just a close causal connection between emotions and expression, but a *constitutive* relation. If some people show a great deal of expressive behaviour with correspondingly little ANS activity, it is unreasonable to assume that the internal activity is merely causing the expression. For externalisers, the expressive behaviour may be the dominant characteristic of their emotional state. There is also a very plausible reason why expressions could partly constitute emotional states: What the set of emotional expressions I listed at the beginning of this section have in common is that they all generate feelings and directly contribute towards the overall bodily state of the subject. Since according to the perceptual theory, the contents of emotional states are essentially bodily patterns, we should include the bodily changes involved in expressive behaviours in these patterns. Clearly facial expressions, vocal productions, bodily posture and movements all involve bodily changes that can be directly felt. In the next section I will also look at ways in which the audible or visual aspects of expression may be automatically ‘translated’ into bodily patterns. Finally, the acts of throwing or hitting something also provide a means to physical sensation. This more indirect method then allows all kinds of tactile sensations to potentially become incorporated within the emotional state.

Hence even though expression is a type of behaviour, it is not necessarily performed for any further instrumental purpose. Rather I claim that expressive behaviours bear the same relation to emotional states as other bodily changes in generating bodily patterns and feelings for the subject. It is this aspect of expressions that is central rather than their communicative role, at least so far as the expresser is concerned.

The main function of expressive behaviour is then exactly the same as the function of emotional states, that is, to represent the dynamic relation between subject and world. Thus the answer to what it is that makes expressive behaviours expressive, what makes them reveal or provide evidence for the emotion of the subject, is simply that they are parts of emotional states. The main difference between emotions and expressions is merely that expressions occur in areas closer to the surface of the body (as opposed to say, visceral changes).

This then has several important consequences: Firstly it allows emotions, and thus our appraisals of the environment, to be observable by others, and so available for social interaction in all the ways that were mentioned above. We can then admit that due to the great utility of this for the organism, evolution has preserved and refined this mode of emotion. A second important consequence is that it makes these aspects of emotion more amenable to conscious control. For that reason they may be used by the subject to satisfy various intentions, and can be tailored to fit the subject's wider goals, background personality and social context.

This consequence is particularly important because it implies that emotions have a cognitive aspect. Recall Prinz's definition of cognitive states and processes as "those that exploit representations that are under the control of an organism rather than under the control of the environment" (Prinz 2004: 45). Since we can control our emotional expressions and as such our emotional patterns, this therefore entails that emotions can be cognitive processes. Moreover a striking result of including the expressions into the bodily patterns of emotions is that it allows for more complex emotions. From an experiential point of view, internally generated sensations are not

especially fine grained compared to the feelings that movement can generate. Expressions can enhance the force, temporal range and nuance of the feelings the subject experiences. So by controlling our expressions we can directly manipulate our feelings to have more complex emotional reactions overall.

But in what sense does elaborating the bodily content of an emotion enable a more subtle or sophisticated emotional reaction? Well to begin with, many cases of expressive elaboration will be heavily mixed up with one's resulting behaviour. For instance, instead of allowing oneself to be overcome with rage, deliberate expressive behaviour may channel or focus that feeling in a productive and highly controlled activity, which in turn generates a sense of power or capability. However not all appropriate emotional responses need involve some kind of immediate modification of one's situation. Many involve contemplation of the emotional state and the struggle to accommodate it within one's life projects. We may ask ourselves, is this emotion rational or beneficial to my goals? Should I therefore try to sublimate it or allow it to blossom? Expressive behaviour is a way in which we can ask ourselves these questions. By directly modifying the bodily pattern and monitoring the results, we can experiment with the emotional appraisal; adjusting one's feeling as seems appropriate. Thus expression can allow the more sensitive appraisal of one's situation.

So to the extent that we can control and elaborate our bodily patterns, emotions are cognitive rather than perceptual processes. There are no perceptual states that we can elaborate at will in the same way. It is not like deliberately refocusing one's attention onto some other object. It is not even like perceiving an object in a different way as a

consequence of one's conceptual interpretation of what it is. A comparable occurrence in a perceptual function like vision would be making a red object appear to have green spots simply by deciding to 'look' in a different way. Overall then, we must conceive of emotions as more complex combinations of perceiving and cognising. Though the perceptual function of emotions is still central, since it is true of all cases of emotion, many emotional states have a cognitive function as well.

Hence as I mentioned in chapter one, it again seems that the boundary between perceptions and cognitions is not especially sharp. In the field of emotions at least, we should regard the stages of perception and cognition as more integrated and overlapping processes. In general our perceptual faculties are actively directed towards exploring and interpreting the environment. Emotions are just a more cognitively sophisticated example of this. Again we should emphasise that emotions are constantly iterated dynamic processes; where the behavioural responses to emotionally inducing situations generate bodily patterns, which generate further behaviours, which also feel a certain way and so on. As such expressive behaviour is another way in which our emotions are pitted against an environment that is experienced as physically resistant to our goals. They are a significant part of shaping one's dynamic response to the world.

A third important consequence of the surface nature of emotional expressions is that it allows individual and cultural differences in emotional states to develop as a consequence of encouraging or prohibiting certain kinds of expressive behaviour. For instance Prinz notes that if we exaggerate a facial expression habitually, then that

bodily change will come to partially characterise that emotion (Prinz 2004: 143).<sup>3</sup> So he claims that two people may have different bodily patterns that are elicited by a single type of emotional object and hence that different bodily patterns can represent the same core relational theme. Given my conception of emotions where bodily patterns are regarded as more directly appropriate to the situations they represent, I cannot agree with Prinz here. Rather, I would argue that two slightly different bodily patterns consequently represent the situation as having slightly different emotional meanings. They would indicate a different way of interacting with the world. Nevertheless, I think Prinz and I would agree that this fine tuning process means that the changes generated by expressions can determine, and be determined by the background character and cultural environment of the subject. For instance a brave man may habitually suppress expressions of fear, resulting in less intense experiences of fear generally.

In general then, people can ‘wear’ their emotions in different ways, some undergoing their emotions more in terms of outward expressions than others. However, all these different forms of expression-emotion sit within a single arena of bodily patterns and feelings. There is no principled reason to distinguish an increase in heart rate from a pattern of facial tension in terms of generating bodily patterns.

It should be noted however that incorporating expressive behaviour into emotional states does *not* make my theory of emotions a behavioural theory in the traditional sense, since it still concentrates on bodily patterns rather than dispositions towards certain instrumental actions. As I described in chapter one, the relation that bodily

---

<sup>3</sup> Both Prinz and Ekman agree that expressions are components of emotions rather than effects (Ekman 1977 cited in Prinz 2004: 134).

patterns have to various behavioural responses (which includes interactions with other people) are important in explaining the *appropriateness* of the pattern to its emotional object. However expressive behaviours are components of emotions *only* in terms of being a special kind of bodily change that generates feelings for the subject. In so far as they are instrumental in achieving certain goals, expressive behaviours are consequences rather than constitutive of emotional states. As instrumental actions, behaviours may even be the response to the world that the bodily patterns are representing. So instrumental actions can cause, and be caused by emotional states. They are not part of the emotion. Hence in some cases we could have two perspectives on the very same thing. In one respect an action could be part of the emotion, and in another, a cause or consequence of the emotion. Accordingly we should distinguish between behaviours that are purely expressive and behaviours that are purely instrumental responses, and those that are a combination of both.

### **Recognising Another's Emotion**

So far I have identified two main aspects of emotional states; the bodily patterns that are generated by both internal and expressive bodily changes, and the situations or thoughts that these patterns track. In any given emotional state, since both of these aspects may vary, they both need to be described in order to fully identify and distinguish that state. However a complete description of an emotion must also consider the person to whom the emotion belongs. Sometimes our experiences track the emotional states of *other* people. These are empathic states.

When exploring how we recognise and understand other peoples' emotions it is important that our explanation be compatible, and preferably complementary with

the account of emotions I have presented so far. The most immediate consequence is that if we do not recognise the bodily pattern that another person undergoes, we may not have properly grasped the nature of their emotion. So I will try to justify this claim, as well as to explain *how* it is that we can get an experience of the way another person's bodily patterns feel. In addition, I have argued that emotions are a form of perception. So it seems that *fully* understanding the emotion of another would also require grasping the situation that their emotion is directed towards, whether this is real or imaginary. Yet emotional states affect the way in which we perceive the situations that elicit that emotion. For example, the emotion may cause us to focus more closely on the aspects of the situation that instantiate the formal object of the emotion (cf. de Sousa 1987). Hence it is not enough that I recognise the situation that the other is directed towards, then recognise the bodily pattern, and then simply add these components together. I have to somehow find a way into the loop where confronting a situation determines an emotional response, and the emotional response determines the situation one takes oneself to be confronted by.

So we have three distinguishable considerations; the bodily pattern (which generates feelings), the situation, and the way the pattern and situation interrelate. There is however, a further level of complexity to grasping another's emotion, which is when considerations of character (such as being brave) and non-psychological context (such as being the Prime Minister) come into play. In both this chapter and the last, we have seen ways in which the character and cultural background of the person can impact on their emotional state. Hence any given process of understanding another's emotion will be more accurate and complete the more of these considerations are incorporated.

Now it is sometimes claimed that incorporating ideas of character and context is essential for empathy. This is because empathy is commonly defined as imagining *being* the other person, (or at least being in their position) and using this imaginative process to understand or anticipate the other's behaviour. As such empathy is regarded as a highly developed imaginative capacity, apparently relying on rather more sophisticated cognitive processes than those required for the basic experience of emotions. However, at this stage I only concentrate on the direct recognition of another's emotion and then develop my account to include considerations of character and context in the following section. I believe that the basic recognition of emotion is the core case of emotional empathy. It still provides a minimal sense of what it's like to be another person, since we replicate a feeling of their bodily patterns. Unlike other cases of empathy however, this is not really an imaginative project, or a sense in which I relocate myself in another's position or character. When recognising someone's emotion, I generally retain the perspective that he is separate from me. Also I do not wish to deny that in more developed empathic projects, we may begin with the characterisation and *then* get a feeling of the bodily pattern involved. Nonetheless, if we have direct access to the bodily pattern of another, it is this that we will 'check' our imaginative process against to see if it has been reliable.

So how do we achieve this direct awareness of the other person's emotion? The simple minded answer is that we just see the emotion in the facial expression or bodily movements of the other person, or hear it in the tone of their voice. It is most faithful to the phenomenology of emotion recognition that the way these expressions

are presented to me in my experience is not as causal correlates of the emotional state, but rather as the emotional state itself externalized.<sup>4</sup> It just doesn't seem that the emotions of others are *hidden* (at least not always) as is sometimes assumed by theories of mind reading. Accordingly, it doesn't seem like I have to *decode* or *infer* the emotion from the facial expression of another, where for example, a down-turned mouth 'means' sadness and an upturned one 'means' happiness.

Inferring emotions from facial expressions would be similar to verbal reporting because it does not require any sense of how the emotion *feels* in order to recognise it. Even if I had never felt sadness or happiness myself, I could come to learn that certain causal inputs typically result in down-turned mouths, reports of "I feel sad" and various behavioural outputs like inactivity or crying.<sup>5</sup> However, beyond the most basic emotional states, the visual (or vocal) expression of an emotion can be extremely complex and subtle. It would be very difficult to recognise many expressions as a clear sign of any emotional state without getting a feeling of the bodily pattern associated with them. This is true both in the sense that without a feeling of the bodily pattern, we are unlikely even to recognise many emotions, as well as the sense in which we cannot help but associate a bodily pattern with the perceived expression.

So the view I wish to develop here is that our brains are organised such that they non-inferentially associate the visually and audibly perceived expressions of emotions with their corresponding bodily patterns. This association requires three

---

<sup>4</sup> This idea is also defended by J. L. Austin (1979).

<sup>5</sup> This may in fact be how autistic subjects recognise emotions, and hence explain why they are so bad at it.

distinctive stages; first the visual or audible data must be perceived as patterns of bodily movement. Second, the subject must either imitate, or mirror at a neural level, the movement that he perceives. Finally, this imitation or mirroring must arouse the subject to some minimal extent, allowing them to recognise the emotion that the other expresses. Altogether this constitutes a simulation process, whereby we recognise the emotional state of another by replicating it in ourselves. This account is therefore allied to the Simulation Theory, which is a theory that we understand the minds of other by taking on pretend versions of their beliefs and desires (see following section).

In order to see what motivates my account of emotion recognition, it is perhaps best to examine the stages involved in reverse order. Why for instance, should we think that recognising the emotion of another requires the emotional arousal of the empathiser? There are several reasons we can appeal to here: Firstly, if no arousal takes place, it is difficult to see what else could provide an appreciation of the bodily pattern of another person. If, as I claim, we are truly getting a feeling of the bodily pattern of another person, and given we don't directly receive the sensations of others, what else could provide the phenomenology of the other's feeling? Even if we employed a memory of our own past emotions here, recreating the felt quality of a bodily pattern would require at least a neural recreation of arousal.

On the positive side, we know that it is possible to be aroused by another person's emotional state purely by observing it. There are cases of emotional contagion in which simply observing another person undergoing an emotional state can cause the same emotion to be aroused in oneself, even when one is not aware of the cause of

the original subject's emotion. Consider for example the infectious panic of a crowd, or laughing when those around you are laughing, despite not knowing the reason in either case (cf. Hatfield et al. 1994). Moreover if we can be aroused by the emotional states of other people, then we would also expect that such arousal would allow the recognition of that state, just as we are able to recognise our own emotional states. Hence to a large extent, emotional contagion and the recognition of our own emotions could already provide most of the mechanisms we need for empathy. It would certainly be surprising if empathy did not make use of these available processes.

Of course, these cases could not be *completely* similar to empathy because when we empathise, we recognise the emotion as belonging to the other person, not ourselves. As I described above, phenomenally we seem to perceive the emotion *in* the other person's body. Yet if the empathiser's arousal is tracking the emotional state of the other, then we should expect it to represent the arousal as a property of what causes it, as described by Dretske's theory of representation. Equally, if the empathic process is initiated with the explicit goal of determining the other's emotional state, then it is reasonable to suppose the results of that process will be imputed to the other person. So the story we are developing here is that we have at least a minimal level of emotional arousal, which because it is set up to track the emotional state of another person, we don't perceive as belonging to ourselves, but rather projectively perceive as belonging to the other person. This triggers a non-inferential judgement that we just see the emotion in the face or hear it in their voice. Moreover, since within this context the empathiser's arousal accurately and reliably tracks the emotion of the other, it is a *veridical* perception of the other's emotion.

Finally, the most conclusive reason for the connection between arousal and empathic recognition comes from empirical evidence that when people become unable to experience a particular emotion, they develop a corresponding inability to recognise that emotion in others. In chapter one I mentioned the case of ‘S’ who lost the ability to feel fear as well as recognise the expression of fear in others (Damasio 2000: 62-65). A similar inability to feel disgust and recognise it in others has also been observed in patients with Huntington’s disease (Goldman & Sripada 2005).

In addition, some of the best evidence comes from studies of subjects suffering from Parkinson’s 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 recognise that emotion in others (Lawrence et al. 2006). Being able to selectively turn on and off both the arousal and recognition of anger by regulating one specific neurotransmitter strongly entails that it is a necessary pre-requisite of both phenomena, and hence that they share neural processes.

Yet despite these considerations, it might still be objected that when we recognise the emotions of others we just don’t seem to feel aroused, at least in most cases. However, this objection assumes too strong an idea of arousal. Firstly, arousal need only be of a highly attenuated nature. There might only be the beginnings of muscle

tension, expressive movements, or other somatic changes. Alternatively we can appeal to Damasio's as-if loop, which I described in chapter one, whereby the brain makes a map of the bodily pattern in the absence of any actual bodily changes. It may only be necessary for the brain to 'plan' to engage in the bodily changes or expressive behaviours involved in that emotional state to trigger the emotional arousal. This modeling of the bodily changes could also trigger associations of previous experiences of that emotional state. Finally as I described above, any feeling of emotion should be imputed to the person we observe. Hence provided the emotional arousal is fairly limited, we would expect that it would not be consciously attributed to oneself. Then on those occasions when empathy results in a stronger level of arousal this could simply lead us to attribute the emotion *both* to the other and ourselves.

So we are gradually developing a picture of how empathic recognition can occur, in which limited arousal, or a neural model of the bodily pattern allows the emotion of the other to be recognised. Moving backwards through the process we should now explain how exactly this arousal comes about. Arousal requires that some of the same bodily changes (or a neural plan of those changes) that are going on in the other person be replicated in the empathiser. In addition, this process must be triggered by the perception of the other person. It is clear then that the empathiser must imitate those bodily changes that he is able to perceive in the other person. The most obvious perceivable changes will be the expressive behaviours of the other, such as their posture, bodily gestures and facial expressions. As I argued in the previous section, if the empathiser adopts these behaviours, this will generate a range of bodily changes that are characteristic of the overall emotional pattern. These in turn could trigger the

internal bodily changes that usually accompany those expressive behaviours. However whether or not these internal bodily changes are also imitated, the imitation of expressive behaviours should generate sufficient arousal to allow recognition of the emotional state.

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: 86-89, cited in Damasio 2004: 312). But how exactly is this imitation achieved? It is necessary that the visual or audible data of the other person's behaviour must be somehow converted into motor plans for producing those same behaviours. It is also evident that humans have the capacity for this from birth, since it has been observed that infants are able to imitate facial expressions such as tongue protrusions within an hour of birth (Meltzoff & Moore 1983: cited in Meltzoff and Gopnik 1996).

Since infants cannot see their own faces, this indicates an innate capacity to associate the proprioceptive sense of moving the muscles in the face with a presentation of that movement visually. This association primes the imitative action, which then takes several attempts to achieve. If imitation takes several attempts, the infant also requires a means to recognise whether imitation has been successfully achieved or not. This requires some representation to be held which incorporates both visual and proprioceptive aspects of the action. So when attempting the action, a motor plan uses the representation to predict what should be perceived and this is matched against the perception of what actually happens (cf. Hurley 2005).

There has also been an empirical discovery about what could ground these imitative tendencies. The existence of 'mirror neurons' has been observed in the brains of monkeys (and so presumably exist in humans as well). Mirror neurons have aroused a great deal of interest because they fire both when the monkey performs an action and when it perceives another monkey performing that same action (Gallese & Goldman 1998). Marcel Kinsbourne (2004) argues that what is happening here is that mirror neurons are involved in representing the action neutral of any particular perspective, and then another part of the brain must mark that action as belonging either to oneself or another (Hurley 2005 makes a similar argument). Where this action is of an expressive type then, it may be that we get an immediate sense of emotion regardless of who it belongs to and only at a later stage distinguish whether it belongs to oneself or another.

Mirror neurons indicate a neural level of action mirroring, allowing the perceiver to recreate the motor plans of the other person. As such, mirror neurons most likely ground a person's capacity to actually imitate the other's actions. Assuming these mirror neurons are present from birth, they should also ground the innate imitative tendencies we see in infants. 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.

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 do not only see as well as feel our own bodily movements, we also hear as well as feel our own vocalisations. 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.

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 even persists when subjects know that only a single flash is presented, suggesting that it is a feature of early stages of perceptual processing (Violentyev, Shimojo & Shams 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, however certain 'pseudo' synaesthetic associations seem to be universal. For instance, it is common to link a high sounding note with a spatially high location. In one test of this association, Vilayanur Ramachandran had people look at two pictures, one similar to an inkblot and the other like a piece of shattered glass and asked which one was called "bouba" and which called "kiki." He found that 98% of people named the inkblot "bouba" and the jagged picture "kiki". He says:

Perhaps that is because the gentle curves of the amoebalike figure metaphorically mimic the gentle undulations of the sound "bouba" as

represented in the hearing centers in the brain as well as the gradual inflection of the lips as they produce the curved “boo-baa” sound. In contrast, the waveform of the sound “kiki” and the sharp inflection of the tongue on the palate mimic the sudden changes in the jagged visual shape. The only thing these two kiki features have in common is the abstract property of jaggedness that is extracted somewhere in the vicinity of the TPO, [an area of the cortex at a junction between the temporal, parietal and occipital lobes][...] In a sense, perhaps we are all closet synaesthetes. (Ramachandran & Hubbard 2003)

Of particular relevance here, Antonio Damasio (2000) also describes experiments where subjects spontaneously described the movements of a chip moving on a screen in emotional terms:

Some jagged fast movements will appear ‘angry’, harmonious but explosive jumps will look ‘joyous’, recoiling motions will look ‘fearful.’ A video that depicts several geometric shapes moving about at different rates and holding varied relationships elicits attributions of emotional states from normal adults and even children. The reason why you can anthropomorphise the chip or an animal so effectively is simple: emotion, as the word indicates, is about movement, about externalized behaviour, about certain orchestrations of reactions to a given cause, within a given environment. (Damasio 2000: 70)<sup>6</sup>

---

<sup>6</sup> Damasio also notes that damage to the amygdala inhibits the ability to attribute emotions to visual patterns. The patients instead describe them in an “accurate, matter of fact manner” (Damasio 2000: 343-344 ft. 21).

Thus there is good evidence for all sorts of reciprocal connections between the sense modalities. Moreover, the different forms of sensory information can all be reduced to a sense of movement or shape, neutral between any particular form of presentation. This reduction, or convergence of sensory data is not achieved by any additional neural module. Although the brain is organised 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.

There is also evidence that the brain prioritises multimodal stimulation. Receiving stimulus from two modalities at once results in greater activation in *both* of the areas responsible for dealing with each kind of sensory input of its own. In addition, some individual neurons in the prefrontal cortex function solely to register the fact of coincident visual and auditory stimuli or visual, auditory and tactile stimuli (Aou et al., 1983, Tanila et al., 1992). This prioritisation is hardly surprising since the brain is adapted to deal with real objects and events in the world, which tend to possess multiple sensory aspects.

So overall, it seems that our ability to empathise with another person's emotion is grounded in the brain's multimodal connections, which interpret the various sensory information as patterns of movement. The third-personal presentation of movement then triggers a first-personal imitation or neural mirroring of this movement.

Imitating the movement, or even just forming a motor plan to do so then results in the arousal of the observer. Furthermore, it should be noted that although this simulation process is triggered automatically, it is not always done unconsciously. This is where the recognition of emotions is distinct from emotional contagion. Both processes involve getting the feeling of another's bodily patterns. However emotional contagion is caused by the unconscious mirroring of another person's emotional behaviour which then arouses the same emotion in oneself, though the self-conscious reasons we then ascribe for that emotion may well differ (Hatfield et al., 1994). Emotional recognition in contrast is usually conscious and as a result we may well inhibit the tendency towards behavioural mirroring.

For this reason, recognising another person's emotion does not usually overtake one with that emotion. It also means that although we get a feeling of the emotional state, we experience that feeling as belonging to the other person and not oneself. Hence when I get the direct phenomenal impression of another person's feeling of anger, I need not necessarily feel that I am angry myself. My overriding feeling could be one of fear, which I experience at the same time (and as a result of) getting a sense of the anger of the other. The existence of an as-if loop would explain this capacity well, since the as-if loop could process the simulated bodily patterns of another at the same time as our ordinary bodily changes generate another set of patterns. Thus we are able to hold a feeling of another person's emotion quite separately from our own thoughts and feelings. In his novel *Enduring Love*, Ian McEwan describes a vivid example of this. A woman experiencing intense grief and jealousy confronts the main character. He says:

I felt that empty, numbing neutrality that comes when one person in the room appears to monopolise all the available emotion. (McEwan *Enduring Love* 1998: 112)

Despite the fairly metaphorical tone of ‘monopolise’ here, which implies that emotion is the kind of physical stuff that might be shared out amongst people, this description seems to accord well with our phenomenal experiences of empathy. One can have a sense of another person’s emotion whilst *simultaneously* ascribing a different feeling to oneself (in this case a numb feeling, or just no feeling at all).

### **Imaginative Empathy**

Having acquired a feeling of the other person’s bodily pattern, a more complete understanding of their emotional state will require some idea of the situation that the bodily pattern is tracking. An appreciation of the emotional feeling should intuitively suggest some likely candidates; anything that fits the core relational theme represented by that state. Alternatively, the shared social context may immediately provide such information. The other may be reacting to something I have said, or some object in the common environment, to which I can track his gaze. In this way, if I am already ‘tuned in’ to the cognitive state of my friend, then emotional responses are likely to be relevant, and interpretable as such, in just the same way as if my friend verbally communicated with me (cf. Sober and Wilson 1993). So the general strategy we are developing here is one where I constantly track the emotional state of the other, and then search for ‘appropriate’ elicitors of emotion to fit them. In this case, my own current reactions, previous experiences, or general theoretical knowledge of emotional causes will be important.

Of course, emotional expressions are not always deliberate in the way that verbal utterances are, and without this communicative intent behind them, the defeasibility of the ‘most relevant interpretation’ increases. It is quite possible to be talking with my friend about one thing, whilst he thinks of quite another emotionally eliciting situation. Moreover, deliberate emotional expressions are often less than sincere. So although the emotional state may be presented to me directly, we do not always take such information as given without the kind of supporting evidence supplied by an appropriate causal reason. If an appropriate reason is immediately available then we may well be satisfied. But if no such reason is obvious we are inclined to probe for something more personal to the other, such as some private thought or an idiosyncrasy of character (like a phobia).

At this point we begin to see another important strategy for understanding another’s emotional state. In this alternative case, I simulate pretend versions of the desires or beliefs of the other and then see what emotional feelings result in consequence. As I mentioned in the previous section, this is the idea behind Simulation Theory (e.g. Currie & Ravenscroft 2002). Here the pretend beliefs and desires are held ‘offline’ in that they do not affect my actual beliefs and desires. However I am still using my own background beliefs to help draw inferences from these simulated states. In general, the strategy of Simulation Theory will be of most relevance in situations when the other is not actually present, or if I possess information that the other is about to gain, and I wish to predict their emotional reaction. Here we begin to develop a more sophisticated empathic process.

However, it is not clear whether this strategy is likely to replace the first strategy described in more immediate interpersonal circumstances. Take for example emotion recognition tests where we are asked to match a picture of a facial expression with one of four possible descriptions. Is it the case that you get a feeling the emotion and then check each of the four descriptions to see which one matches, or do you simulate a set of beliefs and desires and then see if it results in the facial expression? There may not be a definite answer to whether the reason or the feeling comes first in the process. On the one hand, since we are often aware of the reason for our emotions prior to actually feeling them, we may be likely to simulate a reason and then check the results of that simulation against the picture. On the other hand, the direct recognition of emotional feelings is automatic.

Yet either way, a loop could be set up where once I am aware of the feeling, I refer back to the reason to gauge any further responses. As I have emphasised several times, emotions are dynamic processes that function to drive sustained behavioural responses. These responses are likely to change the subject's relation with the object of emotion, which are in turn likely to change the emotional reaction. For instance, if I flee from a predator, my feelings of fear should eventually turn to relief if my actions have been successful.

However as our understanding of the emotion of the other becomes more sophisticated, the feeling of the emotion becomes more of an end result than a starting point. Although the basic recognition of another's emotion requires a first person simulation process, the feeling we sense is immediately projected or perceived in the other person's body. In contrast fully empathic processes require the

empathiser to centrally imagine *being* the other person in a far more radical way. It may not be enough to simply imagine myself, with all my current beliefs and attitudes, in a situation similar to that of the other. Rather, I may also include in my simulation a sense of the background character traits, or long-term emotional dispositions of the other, that will subtly affect my interpretation of any emotional feeling, and more significantly affect any resulting behaviour. Thus the project of imagining ‘being in another’s shoes’ can incorporate various degrees of detail.

Overall, when understanding the emotion of another person, we can stipulate that the following series of accounts should show a gradual increase of complexity, from a basic feeling of the other’s bodily pattern to a complete empathic simulation. If this is a fair description of the various stages of understanding then we see that additional information provides a refinement of the initial feeling-based understanding, rather than a radical shift:

I sense that:

- Fred is feeling annoyed, (via facial, vocal or postural expression).
- Fred is feeling annoyed due not getting what he wants, (core relational theme).
- Fred is feeling annoyed because he missed his train, (particular object of emotion).
- Fred is feeling annoyed because he missed his train, but only by a few seconds, (focus of particular object).
- Fred is feeling annoyed because he only just missed his train and he had an important meeting to get to, (background non-psychological context).

- Fred is feeling annoyed because he only just missed his train, and he had an important meeting and because he is generally an irritable sort of person (character traits).

Yet is it possible that some additional information could completely shift our understanding of the nature of Fred's emotion? Suppose:

- Fred is a Martian who expresses profound joy with facial expressions similar to those we interpret as annoyance.

Or alternatively:

- Fred knows that I am trying to understand his emotional state and is really pretending to be annoyed in order to undermine my efforts.

It seems that we cannot rule out cases such as these and that in order to engage in empathic projects, we must assume that the other shares some common 'form of life' or is not radically deceiving us (cf. Goldie 2000: 183). However, I do not think that these possibilities should overly trouble us, since they are general epistemological worries that would undermine *any* process of understanding another person. Empathy is not a process that is likely to deliver certain judgements of the emotional states of others. It is a skill that is gradually developed throughout life, and which improves the more contact we have with the person with whom we empathise. Accordingly, any 'knowledge' we gain of the emotions of the other will be of a highly defeasible nature.

So overall, the basic capacity to recognise the emotions of other can then be supplemented imaginatively with more theory like considerations that take into account the situation to which the emotion is directed, as well as the background and character of the other person. The reverse is also possible, where I begin by considering the emotionally eliciting situation and then see what feeling results. Given a perceptual theory of emotions, empathy is a lot like any other process of imaginative perspective shifting. Only instead of trying to see from another person's eyes the most salient object in the environment, I can *feel* from the perspective of their body and character the most salient aspect of the situation to their well-being.

This general skill will be vital when explaining how people can adopt each other's viewpoint in acts of joint attention and approach the environment as a joint subject, which I describe in chapter six. The perception and dependency of one's bodily arousal on another's involved in empathy is also vital for the existence of shared emotions. However, in this chapter the ability to recognise or understand another's state has been predicated on *replicating* those states. In contrast shared emotions will involve a *coordination* of potentially different individual states.

## Chapter Three: Music and Emotion

The aim of this chapter is explain how music expresses emotions. A considerable amount of research has already been devoted to this question both historically and in recent years, and fairly exhaustive analyses of the various theories can be found in Budd (1985) and Davies (1994). So although I provide an overview of the problem and its proposed solutions here, I focus my attention on a few of the most prominent current theories. In particular those of Derek Matravers, Stephen Davies and Bruce Vermazen.

Overall, there are three main types of theory which deal with the relation between emotions and music; expression, arousal and resemblance theories. These theories principally differ according to *where* they locate the emotion expressed. That is, either in the responses of the listener, the connection the music bears to the composer, or in the music itself. My account is a development and synthesis of these theories rather than a radical rethink. My main concern is to use the theory of emotions that I outlined in chapters one and two to provide a framework that unifies the ways we encounter emotions in music with the ways we encounter emotions in more everyday contexts. In doing so, I hope to emphasise what I see as the especially intimate connection between music and emotional states.

Ultimately I argue for the kind of direct connection between music and emotion that is summarised by the slogan, ‘the music sounds the way the emotion feels’ (Pratt

1931, Budd 1995).<sup>1</sup> However, I do not agree that consciously tracking a resemblance between sound and feeling is necessary to recognising the emotional character of music, as might be implied by the above slogan. Rather, I argue that our recognitions are best characterised as imagining or having an illusion that a person is in some manner responsible for the musical emotion. As such, my claim is that our recognition of emotion in music is causally similar to our empathic activities. Given that my analysis of empathy is based on the neural simulation of bodily patterns, we might accordingly call this a simulation theory of musical expression.

Now I do not wish to commit myself to the view that when people recognise emotions in music, they are necessarily able to articulate what emotion it is that they perceive. It even seems possible to *unconsciously* recognise the expressive qualities of music, or at least respond in a manner that indicates sensitivity to its expressive qualities. So it is important to distinguish between the causal basis of recognising emotions in music and the conscious experience of recognising emotions in music. In this chapter I mostly focus on the causal story and then finish by describing the minimal form of conscious experience that recognition tends to lead to. This involves a basic sense of a person's body either attached to, or embodied by the music. In chapter four I then elaborate on the conscious experience of musical expression. This is part of a wider discussion on whether the music really has the expressive qualities that we perceive in it.

---

<sup>1</sup> Note that this indicates a connection to the phenomenal experience of emotion rather than the bodily patterns that generate that feeling. Since however, phenomenal feelings directly track bodily patterns this should not preclude the idea that music equally captures bodily patterns. The main contrast here is with resemblance theories that claim that music sounds the way emotions *appear* (e.g. Davies 1994).

## **Initial Evidence**

Before I discuss the main theories of musical expression, it is useful to get a clear grasp of the kind of thing we are talking about (as much as that is possible pre-theoretically). The ways we ascribe emotions to music might at first appear to be quite subjectively variable. Yet when general emotion labels are used (such as sad, happy, angry etc.) there is often much agreement as to which emotions are being expressed. Moreover, it is highly unlikely that one person would characterise a piece of music as joyful where another characterises it with a quite contrary emotion, such as sadness. It is more likely to be in the details that the differences between listeners will show up. The variability of responses to music is one of the issues that I discuss in chapter four. For now however, my aim is simply to point out that at some basic level of description there is sufficient unanimity in response to indicate that the musical expression of emotion is a real phenomenon.

Empirical investigations into musical expression confirm this idea. In one study by Patrick Juslin (1997), professional guitarists were instructed to play a familiar melody (*When The Saints*) according to different emotional interpretations; either happy, sad, angry, fearful or without expression. Changes to the pitches of the tune were not permitted, but the performers could use significant variations in dynamics, articulation and tempo. Despite the melodic restriction, listeners were able to accurately judge (using scales of intensity) which emotion the performer had intended to express with a success rate comparable to decoding ordinary vocal expressions of emotion. This effect was also independent of any musical training on the part of the listener. Interestingly, most listeners did not apply the non-expressive label to even the deliberately neutral performance. Rather people simply judged this

performance as low on each scale of emotion. This may be a result of a bias in the experimental design towards applying emotional labels, but could also reflect a general expressive quality inherent in the melody itself.

The above experiment supports our intuition that the average untrained listener is able to recognise distinct emotions in musical pieces. Of course, these kinds of forced choice studies push the listener into treating the music in emotional terms. In an ordinary listening situation, the associations that a listener may make could also include imagery or narrative events. However it is clear that listeners often freely locate emotional content in musical works and that at least on some occasions, composers write their music with the specific intent that it should be understood in this way.<sup>2</sup> Requiems typically demand a sad emotional character for instance. In these cases the emotional character of the music may well give meaning and unity to a work. Thus if a listener does not recognise the emotional character they may seriously miss the point.

In addition it is not as if works of music *vaguely* remind us of some emotion. Rather they often seem to strike at the very heart of the emotional experience (particularly in great works). As one is carried along by the progress of the music one also feels carried along by the development of an emotional state. This is the particular expressive advantage that music enjoys over art forms like painting or sculpture in that both music and emotions are temporally developing processes. However it is not just the broad developmental aspects that music seems to capture, but also the minute

---

<sup>2</sup> In a questionnaire of 135 expert musicians from England, Sweden and Italy, 99% agreed that music expressed emotions (cited in Juslin 2003: 291). Hence in general musicians think of music in emotional terms, whether or not they seek to exploit that capacity in their work.

nuances of the emotional state. For instance, Peretz et al. (1998) have reported that an ordinary adult listener can distinguish whether a piece of music is happy or sad within a quarter of a second. To a sensitive listener, a single chord is often sufficient for expressive effect.

Yet despite the manifest immediacy and intensity of the expressive character of music, it is not the case that a piece of music is automatically recognised by every listener as having the emotional quality that it has. There are cases when the emotional character may seem to creep up on us unconsciously. For instance, we may be watching a film in which the hero is heading into danger and find ourselves gripped with tension. We may then realise that the background music has been generating this sense of danger. Moreover, it is also possible that a listener just does not recognise *any* emotion in the music. Some listeners may be simply insensitive, but it also seems possible for listeners to switch listening styles, concentrating at one moment on purely technical features of the music and at another on its expressive character (cf. Cook 1990).

Thus this initial review points to some of the features of musical experience that a successful theory must account for. The directness and intensity with which music expresses emotions indicates that it expresses something that the listener hears as an actual occurrent emotion rather than merely a symbol or description of an emotional state. This raises a conceptual problem because emotions are the kind of things that can only belong to sentient beings, not sounds. From one perspective, an abstract set of sounds is not even remotely like the bodily evaluation of the impact of the environment upon an individual. Hence as much as possible, the successful theory of

expression should explain exactly what are the properties of music, and what are the characteristics of the listening experience, that enable this strange connection with the emotions. The theory should explain the immediacy of musical expression yet also reveal why the effect may not be automatically recognised. It must allow for the near universality of some judgements of emotional content yet also show where subjective differences may arise. And it should hopefully do this in a way that makes sense of why humans have developed the capacity to hear music as expressive at all.

### **Realist Expression Theory**

As mentioned above, my own theory of the connection between music and emotions is derived from my analyses of the principle alternatives offered; the expression, arousal and resemblance theories. In my view it is not that these different theories are mutually exclusive of one another, except in their more specific formulations. In fact, I argue that the basic resemblance between music and emotion is what can cause listeners to be aroused, and allow the composer or performer to successfully express their emotions in music. However, I do not agree that the listener necessarily recognises the emotional content of music in terms of recognising their own arousal, noting a resemblance, or having ideas about the composer.

First of all then, let us look at the expression theory: In one sense, any theory concerning the expressive content of music must be an 'expression theory', so in order to avoid any confusion, I call the theory I present at this point a 'realist' expression theory. This is because the central claim is that the emotional content of the music is there as a result of the actual emotional state of the composer. (We may also include the emotional state of the performer or the conductor as possible

alternatives.) So the realist expression theory attempts to solve the problem of who it is that owns the psychological state perceived in the music by identifying it as the state of the composer. This seems a reasonable choice if one must be made, since the composer is usually responsible for the expressive character of the work. Hence the realist expression theory states that when we hear a piece of music as emotionally expressive, we recognise the activity of the composer in conveying their emotion in the music. Now it is unclear whether the composer must *intend* to express their emotion in the music or whether it can happen by more natural or automatic means. For the moment, let us assume that the composer intends the music to express his emotion. Then, as much as it is possible for the listener to recognise the emotion, the composer has been successful in their expressive intent.<sup>3</sup> Technically, expression would be achieved in the music even if there were no listener, but by recognising that expression the listener completes an act of emotional communication.

Whilst works of music do seem to have quasi-linguistic attributes in their ability to express emotions, there are several obvious problems with the realist expression theory. Clearly works of music retain their expressive content in the physical absence of their composers, and so the realist expression theorist must mean that the composer has somehow managed to capture a *record* of their emotional state in the music. But then in order to reveal that emotional record, is it really necessary for the listener to recognise the past emotional state of the composer? This seems unlikely because there is often no way of knowing whether the composer was in fact undergoing the emotion expressed at the time of composition. We know that

---

<sup>3</sup> Cf. the expression theories of Benedetto Croce or R. G. Collingwood. I attack a general realist expression theory here rather than any specific version, since I explore both Croce's and Collingwood's theories in detail in chapter five.

composers produce their works in all kinds of ways and for all sorts of reasons. An expressive work may be motivated by the desire for fame or money rather than an inspired outpouring of emotion. In addition, it is not unknown for composers to write happy music when they are sad and sad music when they are happy.

An alternative formulation of this theory is that perhaps the composer didn't simply pour out an occurrent emotion but rather remembered one that he once had and then expressed *that* in the music. It might be as Wordsworth (1802) said of poetry, a case of emotion 'recollected in tranquillity'. In this case, it may still be necessary to recognise the expressive *intent* or *imagination* of the composer. This view has some attractions because if the composer has deliberately set out to capture an emotion in the music, he may well call to mind the way that the emotion felt and use this as a model on which to base the music. Yet it seems likely that the emotion in this case, particularly when stripped of all the details of its context situation, will be heavily idealised and transfigured by the demands of the music. An initial emotional impetus may be taken over by purely musical concerns, such as developing a theme or achieving tonal closure. In the end, it even looks like we must give up the necessity of the composer's expressive intent, because it seems possible that a composer could produce a work expressive of emotion as the result of complete serendipity. He might only have been interested in the application of a musical rule, or a work based on chance decisions, which quite unintentionally produces emotional qualities as well.

If an expressive piece of music can be created merely by the systematic application of a musical rule, then the realist expression theory is in real trouble. Moreover,

given that we are analysing musical expression from the perspective of the listener, must the listener have any idea at all about how the music was produced? Probably in order to hear it *as music* the listener should be aware that the sound is the product of human agency (though I do not have the space here to go into the definition of art). Yet perhaps to hear it as emotionally expressive the listener need not be aware that a human produced the music. Maybe the listener could regard it as a natural phenomenon in the same way as we are able to see willow trees as expressing sadness. The problem with this idea however is that we may well have to imagine some human agency behind the drooping of the willow trees in order to perceive them as sad. Once again we are brought back to the problem that if the music is heard as an actual occurrent emotion then there should be some person to whom it belongs.

An alternative is that the listener could instead treat the music as the mere *appearance* of an emotion. I explore this below when I look at the resemblance theories of Davies and Kivy. For now however, let us hold on to the notion that some person must be expressing the emotion we hear in the music. If so, we are led to another alternative, which is to treat the expression of the composer as a useful fiction. That is, the listener must *imagine* that the composer is expressing their emotion (or emotional intent) in the music in order to hear it as expressive. So even though composition is often a meticulous and drawn out process, the music may be heard *as if* it is the spontaneous outpouring of emotion. In this case we would no longer be talking about a 'realist' expression theory in the strict sense, but rather an 'imagined' expression theory. This theory seems far more plausible, and is very similar to the persona theory that I defend at the end of this chapter. However at this

stage it is worth noting that it does not seem necessary for the listener to imagine the composer or performer or anyone else *actually* responsible for producing the music. Once we give up control to listener to the extent that the imagined expressive theory allows, it is clear that the listener could imagine anyone they like spontaneously expressing the emotion in the music, even themselves.

So even if the listener *can* imagine of the music that it is the expression of the composer's (or performer's or conductor's) emotion, it is neither causally nor experientially necessary that the listener do so. In addition, we might argue that the listener's perception of the music can only be treated as fictionally the direct expression of an emotion if it *already* possesses characteristics appropriate for such a treatment (cf. Davies 1994: 180-184). So we should be interested in finding out what characteristics of the music (or the listener) could trigger this imagination. Then once we have these, we might find that it is no longer necessary to appeal to a fictional person that expresses himself. It is with this in mind that we now proceed to the arousal and resemblance theories.

### **Arousal Theory**

If the composer or performer is not the owner of the emotion expressed in the music, then perhaps it belongs to the listener instead. The listener is at least present on any occasion in which the emotional content is perceived. So maybe the listener is able to apply an emotional label to the music on the basis of his own emotional reaction to it. This is the view of arousal theory, of which one of the most sophisticated versions is offered by Derek Matravers in his book *Art and Emotion* (1998). His analysis of musical expression runs as follows:

[A] piece of music expresses an emotion  $e$  if it causes a listener to experience a feeling  $\alpha$ , where  $\alpha$  is the feeling component of the emotion it would be appropriate to feel (in the central case) when faced with a person expressing  $e$ . (Matravers 1998: 149)

Matravers qualifies his analysis with the phrase 'component of the emotion' because he holds that feelings are only part of an emotional state. He subscribes to a broadly cognitivist theory in which a propositional attitude towards some situation, in combination with a phenomenal feeling, is necessary to distinguish one emotional state from another. According to the theory of emotions I presented in chapter one, feelings (as generated by bodily patterns) are sufficient for complete emotional states, so long as they tend to track the dynamic qualities of certain situations. Either way, the problem for any account of musical expression is that abstract instrumental music does not obviously provide any situation for the expressed emotion to be directed at.

The key insight of the Matravers' theory is to make an analogy between our reactions to emotionally expressive art and emotionally expressive people. In cases where we perceive that another person is undergoing an emotion, Matravers claims that we typically mirror their emotion or feel an emotion complementary to it. Against an idealised background (in which neither the context situation, nor the identity of other will make a difference to the way we react) certain felt responses are appropriate (or more central) to the expressions of certain emotions. For instance it is appropriate to feel sadness or pity in response to another's sadness. Yet we can do this without having to take on the propositional attitude that is part of the other's emotion. So in

Matravers' view, I only respond with a feeling rather than a full emotion. But being aroused by this feeling is sufficient to cause a belief in me that the music is expressing whichever emotion that feeling is normally a response to when I encounter an expressive person.

The difference between responses to expressive music and expressive people is that an expressive person possesses their emotion independently of another's response to it. In contrast, the emotional qualities of music must be understood in terms of its capacity to arouse feelings in the listener. This capacity to arouse emotions is due to more basic properties of the music, which are not fully specified by Matravers. Yet whatever these properties are, their expressive quality is entirely dependent on the aroused response that they generate. Here it should be noted that Matravers' theory (and arousalism in general) is a realist theory of emotion in music. As I discuss in the following chapter, realist theories claim that we should treat the expressive qualities of the music like a secondary quality. So like an object's colour, the music is characterised as objectively possessing a dispositional property to cause a suitably sensitive listener, under suitably normal circumstances, to perceive it as having the emotional quality that it has (cf. Matravers 2003). But the way that the listener recognises the emotional quality of the music is in recognising their own felt response, not necessarily in being able to pinpoint which features of the music are responsible for their feeling.

Since the feelings of the listener are logically distinct from the properties of the music that cause it, the arousalist account seems vulnerable to counter-examples in which anything that arouses the right sort of feelings may be seen as emotionally

expressive. For instance, if I am perturbed by a violinist's horribly out-of-tune performance, is the music thereby expressive of pain or sorrow? Matravers rules out these cases on the grounds that the feeling is caused by a *belief about the music* rather than the music itself. So it is just like any straightforward emotional response to a situation in which (on the cognitivist view) a propositional attitude has caused a feeling. If we were to allow cases like these then we would be in the absurd position that late trains are expressive of anger because they arouse anger in us. Rather Matravers insists that, "[i]t is only feelings which are not simply components of an emotion that cause the belief that their apparent cause is expressive" (Matravers 1998: 169).

I criticise this response in depth below. But first we should note Matravers' point that the music must cause the feeling in the right sort of way. That is, the feeling and the music should be intimately bound up in the consciousness of the listener, such that the feeling aroused is only sustained by continued attention and sensitivity to the organisational features of the music (Matravers 1998: 178-179). In order to maintain this intimacy between the music and the feeling, Matravers is led to claiming that the feeling must follow the stresses and strains aroused by the particular details of the music. This drives him towards a structural arousal theory such as that of Leonard B. Meyer (1956), who argues that music arouses emotions in us due to our acquaintance with various musical conventions (such as tonality). These conventions cause us to have expectations about how the music will continue at any given point, as well as what harmonies count as more or less stable. The idea is that the composer skillfully manipulates our expectations, (particularly of returning to a point of stability) frustrating them at one moment and satisfying them the next. This then arouses in us

corresponding feelings of tension and release. The pattern of these tensions and resolutions then add up to the various different emotional states.

Whilst his discussion of conventions in music is extremely insightful, and some distinctive effects of musical expectations have been empirically confirmed,<sup>4</sup> there are a number of problems with Meyer's theory. First of all, it seems that we could still appreciate the emotional character of a work that was completely predictable to us and so did not delay any expectations. If in response Meyer appealed to more imagined tensions, then we would no longer have a direct arousal theory, and would need some account of what features of the music dispose us to imagine these tensions (perhaps a resemblance theory of the kind I introduce below). Moreover, Meyer's theory is not readily applicable to non-Western styles of music such as Javanese Gamelan, which don't tend to involve any form of structural resolution.

However the main problem with Meyer's theory is that it is grounded in an overly simplistic view of emotional states. Let us grant that expectations are always aligned with bodily patterns of tension (which I doubt). Still it seems that the two elements of tension and release are insufficient to properly describe the most basic emotions, even if we restrict emotions to their temporal profiles. For instance, there is a difference between the emotions of fear and anger that is not fully captured by profiles of tension and release. Although difficult to describe in words, the way these

---

<sup>4</sup> See for instance Sloboda & Juslin (2001a: 91-93). One particularly strong effect that seems to result from musical expectations is the phenomena of musical 'chills' or 'shivers down the spine' when listening to music. This distinctive sensation most often occurs when the music is either approaching a peak of tension, or begins to release that tension. This phenomenon has been empirically verified by Blood & Zatorre (2001), who have shown that some listeners are reliably stimulated in this manner every time they hear certain passages of music. However, listeners differ according to which passages or pieces it is that reliably cause these sensations in them.

emotions feel involve different senses of heat, visceral texture, space and action. It seems that music has many more resources at its disposal to capture these qualities.

Matravers does not demand that the feelings of the listener must stand in a completely isomorphic relation to the music (although he says that they tend to) and so he is not particularly damaged by these arguments. But without specifying exactly how the music arouses feelings, his theory still permits cases in which by closely following the music, a listener may have a feeling aroused without the music being genuinely expressive. So let us return to the example of feeling perturbed by the out-of-tune performance. Why is it necessary that this music must arouse the feeling via a *belief* about the music? The out-of-tune scraping on the violin may simply jar my feelings, arousing a sense of great unease, which in a central case would be an appropriate response to somebody expressing pain or great sorrow. Matravers cannot rule out this case by insisting that the jarring feeling is part of an emotion *about* the music, because it seems no more about the music than a case of feeling contentment when the theme returns at the end of a sonata. In one case the notes violate my intuitive sense of correct tuning. In the other the return of the theme satisfies my expectations that it would return. In either case the arousal is dependent on my sense of how the music *should be*, though the sonata example seems genuinely expressive of contentment.

The problem is that the very same jarring feeling could be aroused in two people, one of whom recognises that it's caused by the poor performance of the musician, and another who is completely unaware that the poor tuning is unintentional. Since this second listener's feeling is not a component of a belief based emotion, the music

must be genuinely expressive of great sorrow. This might not be such an unwelcome consequence, since there are similar cases in which the sad character of a piece seems to derive from the hesitant and pathetically stumbling quality of the melody. However this possibility leads us to the absurd position that the familiar listener is wrong about the (non) expressive quality of the music, or else both listeners are right and we are left with an unfortunate degree in subjectivity in expressive judgements. This would seriously undermine the realist claims of Matravers' theory. But it seems we cannot say whether the piece is truly bad rather than expressive of sorrow. Even though the familiar listener is more expert, it does not make him the more *suitable* listener without imposing unreasonable demands on what makes a suitable listener (i.e. to be already familiar with a work's true emotional content before being able to judge its emotional content). So it seems that arousal is insufficient to guarantee that the piece is genuinely expressive of emotion, because the presence of an additional (true) belief can render the work inexpressive.

Of course the most common criticisms of arousal theories 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 it is far from saying that it is *necessary* for us to be aroused.<sup>5</sup> It does not even seem to be the paradigm case of musical expression.

---

<sup>5</sup> For a review of evidence showing listeners' arousal responses to music, see Scherer & Zentner (2001). Apart from direct self-reports, arousal can be detected by physiological responses and facial expressions (2001: 374-376), the tendency to remember autobiographical episodes of the same emotional type (2001: 373), as well as compatible self-evaluations (2001: 380). Yet none of these studies prove that arousal is necessary for *all* recognitions of emotions in music. Better evidence for this comes from cases where both the capacity for arousal and the ability to recognise emotions in music is lost, as in the anecdotal example given by James that I cited in chapter one, footnote three.

Matravers claims that arguments of this kind assume too strong an idea of arousal, as if the feeling must overwhelm the listener. Instead he says that merely an incipient feeling, the barest beginnings of feeling, is sufficient for arousal. Then in many cases the natural felt response of the listener is inhibited, or the listener is jaded by overexposure to the work. In these cases the listener is not in a suitable circumstance to properly judge the expressive character of the music.

Yet it just does not seem plausible to me that the vast majority of listening experiences to music are unsuitable for proper appreciation. I agree with Davies when he says:

Simply, there are many cases in which the listener attends to the work, is fully acquainted with the conventions of the work's style, identifies the style and genre correctly, approaches the work when he is not jaded, or preoccupied, or tired, or distracted, but nevertheless in which the listener fails to feel the same emotion as that all experts, including himself, hear expressed in the work in question. (Davies 1994: 195)

In response to this sort of objection, Matravers concedes that it is possible for the critic to have a non-aroused recognition of the emotional character of the work, perhaps when trying to discern how it achieves its expressive effect. He compares this to a doctor who recognises the pain of a patient but inhibits his natural sympathetic response for the sake of professional competence. Yet as Davies rightly complains, this analogy implies that thoughtful attention to music is an *inhibitory*

factor rather than a condition for an appropriate response (Davies 1994: 197). The arousalist cannot simply explain away the frequent lack of an aroused response in listeners by insisting that there must always be some inhibitory factor without making the theory unfalsifiable.

However we should be clear about what it means to be aroused here because on the basis of the empirical evidence that I presented in chapter two it *is* necessary to be aroused by a neurally based simulation of an emotion in order to recognise that emotion in another person. Yet this arousal can occur quite unconsciously. If, as a result of our simulation, we recognise the emotion the other person is undergoing, it is usually experienced not as something we feel in ourselves, but as something belonging to the other person. If our recognition of the emotional content of music is like this, then we should similarly get a sense of the emotion belonging to the music. Since the listener's arousal is supposed to track the expressive qualities of the music, it is perfectly intelligible by Dretske's theory of representation that the perceptual state should represent the original cause, the music, as having the emotion rather than any of the intermediate stages involved. This is I think, phenomenologically accurate. Just as when we recognise emotions in people, we do not need to infer from their expressions to their 'hidden' inner feelings, but rather get a direct impression of the feeling in their face or body, so equally we get a direct impression of the emotion *in* the music.

I argue below that our recognitions of emotions in music are indeed like our recognitions of emotions in people, so I think that Matravers is on the right track when he compares the two kinds of reaction. Yet 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 just a necessary component of the process of simulation.

Ultimately, both the realist expression and arousal theories fail to account for our recognition of emotion in music because in attempting to solve the problem of who possesses the emotion perceived, they focus on the people that surround the production of music rather than the music itself. As such these theories do not provide a very informative explanation of exactly what it is about the music that encourages its treatment in emotional terms. I think however, we should admit that non-conscious arousal is a necessary component of the *causal* process of recognising emotion in music. Yet in order to justify that story we need a fuller account of what it is about the music that triggers that arousal, and how exactly the emotional content of the music shows up in our experience.

### **Resemblance**

Arousal theory claims that the music is sad because we respond in such a way. Yet maybe this gets things the wrong way round, and we should rather say that we sometimes respond with sadness because music can possess some of the same properties that sad people possess. Resemblance theories try to make sense of this perspective by showing that properties of the music *resemble* properties of the emotion.

There are several types of resemblance theory according to what it is that music is said to resemble. One possibility is that music resembles the ways in which humans

give expression to their emotions. Since 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. This is the view of Peter Kivy and Stephen Davies. Kivy and Davies both solve the problem of who possesses the emotion heard in the music by drawing a distinction between an action that expresses or reveals an emotion, and one which merely has the *appearance* of expressing an emotion. The famous example that Kivy uses is the face of a Saint Bernard dog, which looks sad though we don't believe it actually feels sad. In the same way, music need not express an actual occurrent emotion, but merely present the appearance of one. Moreover, the appearance of an emotion need not involve any propositional attitude or be directed towards any situation.

So what exactly does this resemblance to the appearances of emotions rely on? Let us first of all look at the resemblance between music and vocal utterances.<sup>6</sup> 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. Yet this would imply that since operatic recitative most accurately imitates the forms and rhythms of the speaking voice, it should be the most emotionally expressive form of music, which it clearly is not. However it is not the verbal aspects of vocal expressions that provide their emotional effects but rather their more detailed tonal qualities. It is also known 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 for instance that when speaking to infants, mothers will use sharp, staccato contours to express disapproval and slower, falling pitch contours to soothe

---

<sup>6</sup> Kivy more than Davies emphasises this possibility, taking inspiration from the 17th and 18th Century musicians who deliberately imitated the emotional patterns of the voice for expressive effect.

(cf. Storr 1992: 23). The expressive qualities of these kinds of details have been verified in empirical studies (see Scherer, Johnstone & Klasmeyer 2003 for a review). 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. Music also has the capacity to imitate other specific cues such as sighs, tremors, hoarseness, weeping or laughing, though only in a highly stylised way, and as such not particularly accurately. Davies for instance, claims that Penderecki's *Threnody for the Victims of Hiroshima* is so expressive "because it so closely resembles a prolonged, wordless scream of agony, wrung simultaneously from countless throats" (Davies 1994: 207). Yet this is the exception rather than the norm. In the vast majority of cases, music does not need to resemble the overall characteristics of a vocal utterance to have expressive effect, but merely highlight the small cues that occur within expressive speech.

Though the resemblance to vocal utterances goes some way towards explaining the expressive qualities of music, it is insufficient to account for all cases of musical expression.<sup>7</sup> 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

---

<sup>7</sup> Also it is easier to recognise the emotions of others from faces rather than voices. So why is music more expressive than painting? And why for that matter is abstract painting often more expressive than realistic portraiture? The emotional expression of art is not just a matter of picking out which emotion it expresses, but getting a vivid impression of the *character* of that emotion.

mentioned above. As a consequence Davies and Kivy also appeal to resemblance with bodily movements and posture. Melodic lines can seem graceful or heavy, they can jump or droop and chords can seem tense or gentle. Even Eduard Hanslick, the theorist who notoriously argued that music cannot express emotions, agreed that “motion is the ingredient which music has in common with emotional states and which it is able to shape creatively in a thousand shades and contrasts” (Hanslick 1986: 9-11).<sup>8</sup> The means by which music may resemble motion includes variations in volume, articulation, rhythm, texture and differences in pitch. These combined resources seem sufficient to capture virtually any movement imaginable.

However 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. Movement and position is typically presented to us kinaesthetically or visually. Although motion can also be presented to us aurally, as when a car zooms past, or when we locate the source of a sound, this is not the kind of movement that music suggests. Music does not literally move (though a decrease in volume can present the appearance of increasing distance) and the location of the performers is not usually relevant to its effect. Yet in order for music to appear to jump or droop there must be some element perceived to persist through spatial change relative to fixed spatial positions.

It seems to be a matter of Gestalt psychology that we treat melodies and rhythms as single processes rather than the mere succession of tones, in just the same way as we see a sequence of flashing lightbulbs as one continuous movement. But why is it that

---

<sup>8</sup> Hanslick’s worry was just that this is insufficient to express *definite* emotional states.

we attribute spatial positions to pitch? Kivy claims that this is due to constant associations gathered from everyday life:

The “rise” in pitch, like the raising of a physical body against gravity, requires, at least in a great many of the most familiar cases, increased energy. And the rise of pitch, both in natural organisms and machines, betokens a rise in energy level. The faster the wings beat, the shriller the sound; likewise, the more energy expended, the higher the engines whine.  
(Kivy 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 just seems 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, cited in Davies 1994: 231-232). Despite the differences here, these alternative descriptions make sense to us. They are both highly analogous to our descriptions of high and low, and indicate a universality of applying intermodal metaphors to musical pitch. Davies also argues that because it is so commonplace for us to describe temporal processes in terms of spatial movement, (such as the ‘rise’ and ‘fall’ of share prices, or empires) we need not particularly worry about the case of music (Davies 1994: 234-235). An initial natural association has simply been solidified by linguistic convention.

I think we can also provide a slightly deeper explanation for this convention. In the second chapter I described the pervasive intermodality of neural processing. In just the same way as we naturally 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 I would 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.

So we have grounds for automatically perceiving a resemblance between music and movement. However, it is a philosophical truism that anything can resemble anything else. What makes it the case that sounds heard as movement should be particularly taken to resemble human expressive gestures? Part of the reason for the resemblance can be explained in terms of the kinds of movement required to make these sounds. Levinson uses the example that the aggressiveness of the timpani in the Scherzo of Beethoven’s Ninth Symphony is partly due to recognising the way it must be struck (Levinson 2002: 144). Davies also gives a more phenomenal reason why we see movement as expressive saying: “Musical movement is invested with humanity not merely because music is created and performed by humans but because it provides a sense of unity and purpose” (Davies 1994: 229).

Finally, Kivy points to the general animating tendency of humans to imbue natural objects with human characteristics. We are prone to see faces in clouds and regard storms as angry. As he says, “far from being difficult to hear or see things as

animate, it is, apparently, difficult not to” (Kivy 1980: 59). The likely reason for this is that humans are hard-wired by evolution to be sensitive to the emotional expressions of other humans (or animals). It confers survival advantage to be able to recognise (and to communicate) when you are about to be attacked for instance. So like Matravers, both Kivy and Davies believe that the capacity to hear music as expressive is a by-product of our capacity to read the emotions of other people.

Overall I think Davies and Kivy go a long way towards explaining the expressive capacities of music, and I am prepared to accept that music may resemble both bodily movement and vocal utterances. However, I think that these resemblances are a means for music to provide a deeper resemblance to the essential inner character of an emotion, not merely its outward appearance. Part of my motivation for this is that according to the analysis I presented in the first chapter, bodily patterns are central to emotional states. Given this, it is intuitive to suppose that if music is so good at expressing emotions, then it should be because it captures the bodily patterns that centrally characterise them. So the main problem I have with Davies’ and Kivy’s appearance theories is that they do not seem to be true to the directness of the musical experience of emotion. This is unlikely to persuade however, since appearances can sometimes be especially convincing. The point that Davies and Kivy make is precisely that there need be no actual emotion behind that appearance.

However, an especially convincing appearance makes us believe that there is an emotion behind it. It seems to me that the phrase it *looks* sad or it *seems* sad, which Davies relies on as an ordinary secondary usage of the attribution of expression, more commonly reveals a lack of certainty on the part of the utterer. He is

unconvinced, as Davies says, because he does not really believe that there is an actual emotion behind the expression. A more convincing appearance however (as music provides) is simply described as sad. Recall that when we are confronted with the emotions of other people, it seems that we perceive the emotional state itself, not just its symptoms. So the problem with Davies' and Kivy's theories is that they fail to recognise that what makes sad faces *expressive* of sad emotions (even when not actually driven by emotion) is the connection they imply to the inner characteristics of emotions. Our process of simulation, whereby we recognise emotional expressions, utilises some the same neural mechanisms as are involved in actually undergoing the emotions we perceive. This causes arousal, at least on an unconscious level, and thus gives us a sense of the emotion, not merely its appearance.

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 capture. One example is the visceral sickly feeling that is so well captured by the quiet unsynchronised glissandi of a violin section. In general, the expressive effects due to the timbral qualities of instruments do not seem adequately captured by the appearance theorist. I can think of no vocal expression or physical gesture that captures the peculiar nutty quality of a bassoon for instance. Moreover there are several additional dynamic aspects to the experience of emotions that music could be taken to express. As I noted in chapter one, 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. In recognising this Malcolm Budd (1995: 207) extends the resemblance of music to the whole phenomenology of the emotional state.

Ultimately it seems just as reasonable to characterise the inner phenomenology of an emotional state as languid, restless or clumsy as it does a behavioural gesture (Pratt cited in Budd 1985: 39). 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 the listener and carry them along with the progress of an emotional state in a way quite foreign to other art forms.

There is I think, a very simple reason why music is superior to the other art forms in its expressive capacity. It is because sound is more like feeling than any other sense modality. 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. 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. Most importantly both feeling and hearing parse the world in

similar ways. They share similar qualities of abstraction and immediacy. The way that feelings structure experience is not in terms of separated objects, but in terms of the dynamics of action, points of contrast and the visceral qualities of sensation. Similarly sound (and particularly music) is understood in terms of pulse, contrast and timbre. We are much more likely to get emotionally involved in sounds if only because we don't objectify and distance ourselves from them the way we do with our visual experience. As a result, music literally *resonates* with feeling.

Overall then, like both Malcolm Budd and Carroll Pratt, I argue 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. Since I argued in chapter one that feelings track bodily patterns, and bodily patterns are the primary characteristic of emotions, this means that music strongly resembles the primary characteristic of emotions. It is no wonder then that music is the pre-eminent art form of emotional expression.

### **The Persona**

So the direct resemblance between music and feelings or bodily patterns is responsible for the expressive powers of music. However we not yet properly accounted for the *experience* of emotions in music. All we have provided so far is justification for saying that the expressive capacity is grounded in real properties of the music. As I explore in the following chapter, this allows us to distinguish aroused responses that track the music's actual expressive features from those that result from the listener's more subjective beliefs or associations. The problem is that just like arousal, even if the resemblance between music and emotion is part of the causal

story of musical expression, it does not seem necessary that this resemblance filter through to our conscious experience of music. As Aaron Ridley says:

[It] would be like offering an account of pictorial space wholly in terms of the perspectival devices contained by a picture: It might be true that we experience pictorial space in virtue of the perspectival devices that a picture contains; but the experience itself is not merely the experience of perceiving perspectival devices (which could be done without ever experiencing pictorial space). (Ridley 1995: 121, cf. Budd 1995: 205)

If the listener reflects for a moment on their experience, they may well be inclined to admit that they hear a resemblance, yet when we listen to music we don't generally have an idea of how an emotion feels plus an idea of how the music sounds and then a third idea about how the two map against each other. Somehow, we simply get an impression from how the music sounds of how the emotion feels. At the same time we do not forget that the music is music. We continue to follow its formal structure. To compare what is happening now to what has gone before and to have expectations about how it will continue. And it is unlikely that every feature of the music we perceive (at whatever scale of description) is at the same time a feature of the emotion we perceive. So how exactly does the experience of the emotion combine with the experience of the music?

In addition, my rejection of the appearance theories of Davies and Kivy leaves my account open to the fundamental problem that their theories resolved. If the music expresses the actual feeling of an emotion rather than merely its appearance, then

there must be someone to whom that emotion belongs. Arousal and the realist expression theories were rejected for unnecessarily appealing to the emotions of those involved in the production or reception of music, rather than 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? Persona theory is our answer here. It 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. We came close to this idea earlier when discussing the ‘imagined’ expression theory. The refinement of persona theory is that the person we imagine need not be any actual person; it may only be connected to that specific piece of music.

Persona theory as presented by theorists like Bruce Vermazen (1986) still treats an expressive object as the *utterance* of the imagined persona. And in this respect I believe the theory should be modified (at least in the case of music) to incorporate some additional possibilities, such as identifying the music as the *embodiment* of a persona. I still retain the central insight however, which is that the music expresses an emotion if it presents evidence for the mental state of a person. Vermazen’s complete definition, which is intended to apply to all forms of expression, is as follows:

An object expresses a mental property if and only if (subject to certain constraints) attributing that property to an utterer of the object would explain the objects having the features it has, and the property is not presupposed in the attempt to interpret. (Vermazen 1986: 207)

It should be noted that the last clause in the definition is meant to rule out cases where perceiving that there is an emotion serves as grounds for imagining a persona. Instead the grounds for the persona must be derived from the music. The 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 it is a little unclear exactly what the causal story is that Vermazen is implying. 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 likely that whenever we approach works of art we have a 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, utilising 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. This is because some of our simulation mechanisms are always active whenever we perceive people. As I mentioned in chapter two, it is what allows babies to engage in imitative actions within hours of birth. It's just that normally the behaviours we perceive others doing are not intense enough for the feelings they generate to attract our attention above the normal everyday 'chatter' of personal concerns and sensations.

Hence when we listen to music, our background belief that a person is responsible for the work primes 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 illusion of a persona. To elaborate: The simulation mechanism utilises 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 recognising 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.<sup>9</sup> They will then be disposed to verbally identify the expressive content of the music accordingly.

The second strategy is to assume that the listener does not approach the music with the background belief that it is produced by a human. 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

---

<sup>9</sup> Neuroscientists Istvan Molnar-Szakacs and Katie Overy (2006) similarly hypothesise that recognising emotions in music is grounded by the mirror neuron system (which underlies our empathic activities). However they focus on mirroring the intentional *actions* required to produce the movements we perceive in music rather than the more generalised mirroring of bodily patterns (some of which could be generated by deliberate actions).

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 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. The phenomenology is I think best compared to perceiving 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 also 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.

So far I have concentrated on the experience of the emotional feeling, but that feeling is also a sense of a person. In particular, the primary experience of emotions in music involves the sense of a person's *body* connected to that feeling. This is because emotional feelings are primarily characterised by bodily patterns (though as I mentioned above certain patterns of thinking are also distinctive characteristics of emotions). Hence the kind of persona that music suggests may not be quite the same

as the kind of persona suggested by other art forms such as sculpture or paintings. These art forms may instead provide a sense of more disembodied mental attributes. But the persona in music, if it gives us a sense of emotional feeling, is one primarily characterised by bodily attributes.

So the listener will get a sense of the solidity or the grace or the energy of that body, as is suggested by the feeling that it is currently undergoing. Depending on the nature of the music and the listening situation, the music may be heard as the *production* of this body, where the body is imagined as ‘behind’ the music in some manner. Alternatively the sound itself may be imagined *as* the body, which then possesses certain detailed emotional features (see chapter four for a complete description). In either case, there is still a person, a body, which is responsible for the expressive qualities of the music. Moreover, in cases of recognising the emotions of people, although their expressions provide a direct sense of their feelings, these features need not exhaust the characteristics of the emotion we perceive. Rather we may also *anticipate* certain actions they might make, as well as have ideas about situations, beliefs or desires that might typically accompany their emotional state. The same is true of the persona in music. We may anticipate the way the musical persona might behave and imagine the kinds of situations he is confronting.

In general, the persona is wherever the emotion is perceived to be. Thus if the emotion is perceived in the music, then so is the persona. If the emotion is perceived as *behind* the music, then so is the persona. But this is as separate as the persona’s emotion and the music ever get. Any other imagined separation between the music

and the emotional state it expresses is a distortion caused by erroneous theory (i.e. ‘it’s merely an appearance of emotion, because music *just can’t* be feelings!’).

Here we come to what I regard as a terminological dispute. It seems to me that the animating tendency that Peter Kivy appeals to when explaining why we perceive music as resembling human movement is the same thing as the process of simulating a persona that I have just described. Yet Kivy does not equate the animating tendency with the imagination of a persona, and Davies (1997) explicitly opposes the two ideas. The cause of the dispute really concerns how fleshed out this persona is. One of the central criticisms of persona theory is that it puts too great a demand on the contents of the listener’s imagination. It doesn’t seem as if we always imagine people when we listen to music. Yet first of all, this objection does not recognise the fact that the persona can be an automatic perceptual level illusion. And secondly, I agree with Jerrold Levinson (2005) that the persona (at least as it is necessarily invoked) is of the most minimal kind. Levinson suggests that it is constituted only by the emotion and its particular form of expression. My claim is slightly different in that it requires a sense of a body that possesses the emotion. Likewise I think the difference between Davies’ and my conception of the animating tendency is that I believe it gives us the sense of an inner life. This is for all the reasons I mentioned earlier about music connecting to feelings rather than mere appearances.

To reiterate, Davies believes that hearing an emotion in music is no more involved than seeing as sad looking the mask that traditionally denotes tragedy (Davies 1997: 97). But this is precisely where our interpretations differ. If the mask is taken as naturally resembling behaviour expressive of sadness rather than operating merely as

a symbol for it the way the word 'sad' does, then it does so by suggesting the sense of living person. Of course I can simply think of the word 'sad' or a schematic picture of sad face without thinking of a persona. But this would be a case of reference rather than expression. Cases of expression are concerned with the particular *character* of the mental state suggested by the particular character of the expressive object. In these cases we utilise greater imaginative resources, we imagine *what it's like* to have that mental state. We therefore invoke our background sense of personhood, and as a result the kind of perspective employed when understanding people rather than inanimate things.

In addition, Davies claims that we attribute emotions rather than other kinds of movement to music because we experience the music as displaying unity and purpose. Yet an appearance cannot express purposiveness. Since purposiveness is a mental state, a persona must be invoked to express it. Perhaps Davies had a thinner concept here, referring for instance to the purely formal way in which a melody may be perceived to be approaching closure. Yet the idea of purposiveness seems to involve a sense of driving towards something, of having a goal or desire, which must necessarily involve the sense of a persona. Only a persona could provide a sense of *intentionality* within the emotion expressed in the music, though we need not be aware of what it is that the persona is intentionally directed towards.

The main motivation for invoking a persona in music is that it explains the directness with which music expresses feeling. The thin conception also means that invoking the persona does not require the listener to look beyond the scope of the music. Vermazen seems only to worry about outward expression signaling the *idea* of an

inner state, which as Davies correctly points out, could be captured by appearance theory. Yet the real advantage of persona theory is that it makes sense of getting inside that mental state. By grounding an act of emotional simulation, the music is used to get a sense of the emotional feelings of a person.

Finally, with this particular theory in mind, there is potential for empirical verification. If our understanding of the emotional qualities of music relies on our basic empathic abilities, then we might expect that people incapable of empathy are thereby unable to understand expressive music.<sup>10</sup> In particular, people with autism are usually characterised 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. Yet in one experiment measuring their ability to recognise emotions in music, autistic subjects have in fact been able to successfully identify the emotions presented (Heaton, Hermelin & Pring 1999). The problem with this experiment however, is that the emotions the autistic subjects were asked to identify were no more complex than the emotions they were able to identify in other people. If they can recognise when a person is sad or happy then we should expect them to recognise the same in music.

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

---

<sup>10</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.

more to do with *attributing* the results of simulative processing, then we may expect that whilst autistic subjects can be 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 person, and possibly not even as caused by the music. With these rather speculative suggestions then, I am satisfied that my account is at least possible to falsify.

## Chapter Four: The Real Properties of Music

The simulation theory of musical expression outlined in chapter three explained how we perceive emotions in music. Various features of the music resemble various features of emotional states. These resemblances are then processed by the listener's simulation mechanism, resulting in at least a simulated form of emotional arousal. The feeling of this emotion is then perceived in what causes it, the music. This creates an illusion that there is a person responsible for the emotion in the music. It is also possible to trigger the simulation process by consciously imagining that a person is responsible for making the sounds we hear.

Since the music does not have its own mind it is reasonable to argue that the emotion is not really *in* the music, despite the way it seems. The music may possess certain dynamic or structural features that resemble emotional states, yet this resemblance relation is entirely dependent on being treated as such by the listener. In general, the music relies on the simulation mechanism of the listener to be expressive at all, and the simulation mechanisms of some listeners may potentially not be triggered by the music, or may function abnormally. Hence we might claim that the emotions we perceive are essentially a feature of the listener, and though they might be *caused* by the music, the emotions themselves are logically independent of the music.

However, there is a long-standing practise in metaphysics of identifying certain properties (e.g. fragility, solubility) as *dispositionally* held by objects so long as they are reliably generated in appropriate circumstances. A similar account is given of phenomenal qualities such as colour. According to many traditional accounts (e.g.

Locke) colours are secondary qualities, which are dispositional powers of objects to cause experiences of that quality in normal observers under normal conditions. So saying that an object is red amounts to saying that it reliably stimulates the impression of redness in all normal human observers under normal lighting conditions. An observer is then normal just so long as he isn't suffering from any affliction that might affect his colour perception such as jaundice or colour blindness. So whilst the object by itself is not a colour experience, a disposition to generate colour experiences is in the object, and to have colour just is to have this dispositional power to generate colour experiences.

Perhaps then a similar claim can be made about the emotional qualities of music. We could say that a disposition to express emotions is in the music, and so the music really does have an expressive quality.<sup>1</sup> However, the main problem is that unlike colours, we do not have well-established criteria for what counts as a normal observer or normal circumstances for perceiving the music's expressive qualities. Our experiences of music are far more variable than our experiences of colour, as evidenced by the different ways we talk about and react to musical works. Hence in this chapter I will try to establish what the criteria are for a dispositional account of expressive qualities in music. This will then allow me to specify in detail which expressive qualities may be said to be really had by the music, and which are simply extraneous and potentially variable elaborations made by different listeners.

---

<sup>1</sup> Dispositionalist accounts are very popular in aesthetics (e.g. Matravers 2003, Levinson 2005, Davies 1994 and Walton 1994). My account bears similarities to all of them, particularly to Matravers (2003). However I cannot rely much on any of these theorists because I disagree in some or other respect with their accounts of how music expresses emotions.

Since the overarching goal of this thesis is to show how people can share their emotions using music, it is important to delineate the ways in which people may differ in their emotional responses. It needs to be possible to reconcile these differences in some way if even the numerically distinct responses of listeners are to be non-fortuitously aligned. So as part of identifying the distinction between the real qualities of the music and the listeners' more subjective responses, I will explore what some of these common variations are and the extent to which they are encouraged by the background of the listener or the listening context. Overall I will argue that many of our subjectively variable responses to music are refinements of a more basic and universal response, and thus that many responses should at least be intelligible to other listeners (especially experts) if not actually shared by them. This then sets up the argument of chapter six in which I show how sharing a listening context with others can enable a convergence in our responses to music.

### **A Principle of Dispositional Realism**

Based on my account of musical expression, there is I think a general principle which we can use to determine whether an expressive quality is really a dispositional feature of the music or not:

The music dispositionally expresses an emotion  $x$  if and only if a *resemblance* between the music and  $x$  can be *grounded* by standard simulative processing, and gaining expertise or true beliefs about the music does not prevent that resemblance from being grounded.

By 'grounded' I mean that the resemblance between musical features and emotional features must be capable of being processed by the simulation mechanism of the listener such that it results in a sense of emotional feeling, which is then perceived as a property of the music. Regarding the specific resemblance claim; we could alternatively say that music bears an epistemic resemblance to emotionally expressive *people* in that both can trigger a simulation process in the perceiver. However, as I argued in the previous chapter, this is based on a metaphysical resemblance that music bears to the emotional state itself. The music puts you in mind of a person by virtue of its resemblance to an emotion. In other respects, music isn't much like a person at all. So I think it is less confusing to refer to the more intrinsic resemblance to the emotion itself (which is perceived as belonging to a person). Note also that music does not resemble the *experience* of the emotion, since it is not an experience. Rather we should say that the experience of the music resembles the experience of the emotion. Yet again, these experiences resemble each other because the music resembles certain features of the emotion itself.

Overall, my notion of the dispositional properties of music is both causally and psychologically reductive, and as such fairly restrictive. If for instance, one thought that the music was ironic due to its anachronistic style, this irony could not be considered an *expressive* property of the music unless certain features of the music could actually be *felt* as ironic in the manner I have indicated. Perhaps for instance, one could recall some other musical works of a certain historical period and note a resemblance between those works and the current work, as well as a contrast with more contemporary works. Concentrating on the features of these works may then trigger a simulation process, resulting in a sense of contrast or conflict in the listener

that might be characteristic of a feeling of irony. We should note however that if this process were used, the expressive quality would not technically be a feature of the ‘ironic’ work alone, but must include the features of the works it relates to as well.

The clause about the increase of expertise or true beliefs is supposed to refer to gaining knowledge about production rules, historical context, composer intentions and the like. I will explore the notion of expertise in detail below. For now, the goal is to rule out the kinds of cases that I raised against the arousal theory in the last chapter. Here it was noted that the jarring sensation caused by an out-of-tune violin performance could be equally perceived as expressive of pain or great sorrow as well as plain bad. Hence, I wish to rule on the side of the expert who knows that it is a bad performance and accordingly does not attribute the jarring feeling as an expressive quality of the music. The expert’s true belief inhibits his simulation process, or at least inhibits the *results* of his simulation process from being attributed to the music.

One important consequence of this clause is that where music relies on ‘genuine’ surprise or shock for expressive effect, it is unlikely that this quality is a real property of the music. According to my principle, the jaded listener must be considered a completely valid judge of the music. Note however that sudden contrasts in the music, even if completely predictable to the listener, should not undermine the potential to be *simulated* as sudden contrasts, *as if* they were an emotion of shock.<sup>2</sup> So although sudden contrasts may not arouse the jaded listener to the same degree as the completely naïve listener, a simulated version of shock should be aroused even in him.

---

<sup>2</sup> This is comparable to Kendall Walton’s (1990) notion that when watching films or plays that we already know the outcome to, we recreatively *make believe* that the outcome is uncertain.

As I mentioned in the previous chapter it is also clear that the resemblance that certain features of the music bear to an emotional state need not be consciously recognised as such by the listener. My point is simply that it should be possible to note those resemblances in such a way that they trigger the simulation mechanism. For instance, a listener may not perceive the subtle timbral contrasts in the music when they first listen to it. However, they could come to be sensitive towards those timbral contrasts, and as a result be caused to simulate them as expressive. Various methods for increasing sensitivity are then available, (e.g. greater attention, conceptual expertise or sensitivity to emotions in non-musical contexts) which I discuss at various points below. Meanwhile the listener might still not know *why* it is that they get such an effect from the music, or even recognise which detail of the music expresses that quality. The justification involved is more one of basic perceptual access than being able to articulate reasons for the effect.<sup>3</sup> So the listener may just be disposed to verbally ascribe the expressive quality to the work as a whole.

Finally, my principle clearly relies on the normality of the listener's simulation mechanism. Yet the normality of the simulation mechanism is fairly easy to establish by recourse to its use in recognising the emotional expressions of people. In particular, the listener should be able to recognise what emotion people *intend* to express to a degree of accuracy as is currently standard, (see Scherer, Johnstone & Klasmer 2003 for statistics). Of course, certain counter-examples can be imagined in which everybody in the world has a simulation mechanism that operates in manner

---

<sup>3</sup> This seems to me to be in line with Sibley's (1959) views on the ability to justify the presence of aesthetic properties in artworks.

abnormal by our own standards. However, due to the likely overlapping nature of the simulation mechanism with the actual arousal of human emotional states, I would be prepared to accept that if in that world expressions of fear are typically simulated as feelings of joy, then those expressions really do arouse feelings of joy. Moreover as I argued in my second chapter, those expressions may be considered as one of the constitutive parts of the bodily changes involved in generating that emotional state. Hence similarly in that counterfactual world, music that resembles the emotion of fear by our standards, actually dispositionally expresses the emotion of joy for that world.

So by the standards of metaphysical essentialism, the music can only *contingently* possess the expressive qualities that it does. This is the same under secondary quality accounts of colour however, since in another possible world, a molecularly identical surface could trigger different colour experiences due to biological differences in the perceivers of that world. However unlike colours, our sensitivity to musical expression is situated within our broader sensitivity to emotions in everyday cases. If arguments can ultimately be made that certain expressive features of humans are essentially tied to certain emotional states (which I doubt) then a stronger case could equivalently be made for music. However one of the lessons we should draw from the case of musical expression is that emotions may not be *essentially* tied to bodily changes at all, but rather tied to substance neutral dynamic patterns that may be instantiated in a variety of mediums. This thought is pursued further in chapter five.

## Harmonic Relativity

Having articulated a principle for the dispositional expressive qualities of music, we may now put this principle to the test. In particular, one aspect of music that I did not explore much in the previous chapter was harmony, which in some pieces of music is the most intensely effective source of emotional expression. This is because harmony seems particularly sensitive to cultural influence, given the different tonal systems that different cultures use. Many theorists agree that the basic effect of harmony is to provide a sense of tension or release, stability or instability. As I noted in chapter three, Meyer thinks that tension is aroused by the expectations we have about the music, and Davies thinks that the *appearance* of tense movement is presented. It seems to me that harmony functions just like timbre, and is another case of direct resemblance to bodily patterns. Yet the main point of contention here is whether the dissonance of a chord, or the ‘darker’ quality of a minor scale is due to it genuinely resembling features of darker emotions or the result of culturally arbitrary associations.

Davies claims that all harmonic dissonance is relative. A major seventh only seems dissonant to us in comparison to the other chords like perfect fifths, because fifths are more frequently used as points of stability within music. Similarly the minor third only sounds sadder to us than the major third because it is used more frequently in music that expresses sad emotions by *other* means. The reason it is used in this way is because hundreds of years ago it was heard as discordant, and hence restless. As a result it was used more often in pieces resembling the darker emotions (utilising other expressive features of music such as tempo or melodic contour). Nowadays, even though the minor third is no longer experienced as particularly discordant,

relative to the much more unusual chords that are commonly used in modern music, it still retains its negative associations due to this historical convention (Davies 1994: 241).

This implies that if someone were brought up in a culture where minor thirds were not conventionally used more frequently in music with sadder expressive content, then they would not experience the minor third as sadder than the major third. Yet we are left with the puzzle of why, historically, the minor third was *ever* experienced as more discordant than the major third. It may be a consequence of the tuning systems used prior to the adoption of equal temperament around 250 years ago. But even if this is the case, what was it about that tuning that caused the minor third to seem more dissonant?

Ultimately, it seems the expressive difference between the major and the minor third (and their various different tunings) can be objectively grounded in the different positions these intervals occupy within the natural harmonic series. To explain: The octave interval (which is exactly double the frequency of the root) is cross-culturally experienced as the 'same' note though higher in pitch. The octave is also the first natural harmonic (followed by the fifth above that, then the octave again, then the major third, fifth, minor 7th, octave etc.). So this explains why the octave is heard as a particularly stable interval; the overall frequencies contained in the interval lie in simple ratio to each other. Then as we move along the harmonic series, the ratios between the intervals involved become more complex, and so more discordant or unstable sounding. Thus the second harmonic is an interval of a fifth above the octave, and so the fifth is experienced as the next most stable interval, then the fourth

and so on. If this is the case, then the major third may be experienced as more stable than the minor third because it appears sooner in the natural harmonic series (between the third and fourth harmonics as opposed to between the fourth and fifth).<sup>4</sup> So overall, the tension in an interval can be aligned with the complexity of its frequency, in just the same way as the sound of an oboe playing an A has greater expressive tension than the same note played as a pure sine wave.

We may equally be able to find objective grounds for the expressive difference between the major and minor *scales*. The harmonic minor scale contains three semitone steps and an augmented 2nd (effectively another minor 3rd) where the major scale has one less semitone and no step bigger than a whole tone. As such, it seems principally because of its unevenness that the minor key is taken to be darker than the major. The minor key is more angular and more complex than the major key. This, combined with the slightly more dissonant minor 3rd makes the minor scale more suited to express the darker emotions.

Introspectively it is hard to be sure whether one's sense of a tonality can be disassociated from all the pieces of music that one has heard in that tonality. So it may not be possible to conclusively demonstrate that the expressive difference between the major and minor tonalities is grounded in the properties I have described. Yet music psychologists Marcel Zentner and Jerome Kagan report that 4 month old infants preferentially look towards the source of sounds and are less motorically active when the melodies are composed of parallel major thirds (consonant) as

---

<sup>4</sup> Moreover, this major third relates to the root tone (e.g. if the fundamental is a C, the fourth harmonic is an E natural), where the minor third only relates to another, far less prominent, harmonic (e.g. G to E). To find an interval a minor third above the root (e.g. Eb to C), we must go all the way up to the *nineteenth* harmonic.

opposed to parallel minor seconds (dissonant). They also fretted and turned away from the dissonant music more (Zentner & Kagan (1996, 1998) cited in Scherer and Zentner (2003): 367). This suggests an innate preference for more consonant sounds.

Whether the same could be said for the more subtle difference between major and minor tonalities is less clear. One empirical study found that even three year olds were able to consistently assign either positive (happy or neutral) or negative (sad or angry) schematic faces to pieces according to the major or minor harmonic modes (Kastner & Crowder 1990, cited in Dowling 1999). However, a similar study found that children could not assign emotion (based on only two choices of schematic faces) according to tonality until the age of eight (Geradi and Gerkin 1995 cited in Dowling 1999). Clearly more evidence is needed here. Yet it certainly seems plausible that a natural tendency to find certain harmonies slightly more dissonant may then have been intensified by historical musical tradition, until we reach the situation we find ourselves in today in which the minor mode has developed an almost symbolic association with darker emotions. Thus people may find it easier to recognise such qualities when absorbed in this tradition.

I do not think that the major and minor modes would *inevitably* be associated with generally lighter or darker emotions. It is possible to write music in a major key that sounds strident and aggressive. Recall the experiment I discussed in chapter three where guitar players were able to express fear and anger with *When the Saints*, a tune in a major key! Moreover, it seems clear that if one had only ever heard pieces in the major tonality, then some of those pieces would seem darker, sadder or more aggressive than others. Not just because of other features such as dynamics, but

because the particular harmonies *within* the scale that were favoured in that piece would help to give it that character.

So although I agree with Davies that the expressive effect of any chord or interval is only generated relative to other chords and intervals, there seems to be a plausible natural reason why one is experienced as more dissonant than another. However, this is still enough to permit differences in listener's responses to music according to their different familiarity with the harmonies used within that music. To a listener who has only heard medieval church music, a tritone (the so called 'devil's interval') will seem more intensely dissonant than it will to a listener acquainted with more contemporary works. Moreover, the same effect operates not just according to the listener's acquaintance with the wider cultural context, but also within single works. A tritone will seem more dissonant in a piece that otherwise uses only major third harmonies than one in which far stronger dissonances are frequently employed.

Given that dissonance is aligned with darker emotions, what then should we say is the *real* expressive quality of the work when the same interval can be experienced in different ways? The problem is less troubling in the within-piece case, since we could stipulate that a normal listener is one who has had their parameters for dissonance set by that particular piece. We could accordingly stipulate that all works should be treated on their own terms in this way. However, this would undermine the idea that some works *as a whole* are more intensely sad or happy or angry than other works. So we are driven to the opposite view; that the normal listener's parameters should be set by *all* other works of music. Of course, this puts implausible demands on the experience of the normal listener. Yet finding some balance within these two

extremes such as a reasonable acquaintance with works within a particular genre, culture or historical epoch seems fairly arbitrary.

This problem could be quite deep, since the same argument can also be made for relative ranges of pitch, dynamics, timbre, tempo, and complexity. Resemblance to feelings across all the expressive parameters will be sensitive to the ranges of those parameters that the listener is accustomed to. This kind of thought suggests that perhaps *all* music is in fact expressive and that some historical styles which were expressive to their contemporary audiences, have just been rendered comparatively neutral by our current forms of expression. For example, despite its name, the ‘Sturm und Drang’ (storm and thunder) style of some of Haydn’s symphonies no longer strikes the modern listener as particularly intense.

If we are looking for a standard for the expertise of the listener that will set the absolute expressive value of a work or features within the work, then we will never be satisfied. We should rather settle with a claim that works and their detailed features can only be *more or less* sad or happy or angry. These qualities will gradually shift according to the experience of the listener, though their expressive ‘location’ relative to other works will remain constant. Furthermore, the overall ranges for these qualities will increase as more and more works are composed (though the intensity of a particular feature could *increase* if the majority of works slowly become more conservative). If we accept this relativity then, we really only need the most minimal set of criteria for the cultural familiarity of the normal listener, i.e. that he should simply be aware of other pieces of expressive music and thus be able hear works as more or less sad or happy accordingly. This is not to deny that

some listeners can be more sensitive to the expressive qualities of a work or to music in general. This sensitivity is a form of expertise. Yet as I argue below, sensitivity is more about the ability to discern features in the music than to be able to judge the 'correct' intensity of emotional expression.

Finally, we should also note that the same relativity can be discerned in our everyday emotional episodes. To someone who has suffered incredible loss, like the murder of their children, the loss of a shoe will be accordingly less intense than for someone who has never lost anything before. Similarly, the objectively same set of bodily changes may be felt more keenly by a child than by an adult. Hence I do not think we should worry unduly about the relativity of the expressive qualities in music.

### **Expertise**

The case of harmony provides a guide to what criteria we should demand of normal listeners when recognising emotions in music. Though the sensitivity of listeners might vary from case to case, as long as that sensitivity is grounded in the resemblance the music bears to emotional states, their recognitions are valid. At the same time, we can also *rule out* certain expressive effects on the grounds that they do not target such a resemblance.

For instance, it is commonly recognised that listeners may associate a piece of music with a particular feeling based purely on associations with the context in which they first heard that piece, (also known as the 'darling, they're playing our tune' effect). For example, Mahler is said to have had a strong personal association between anguished emotions and brass bands as a result of a childhood experience in which

he fled into the street from an argument between his parents at the very moment when a brass band was passing. This personal association is likely to have deeply affected his compositions. So some of the brass figures in Mahler's symphonies, even if they resemble a joyful emotion, may have had for him an anguished undertone that is simply not discernible by most listeners.<sup>5</sup> For this reason, we should not accept such associations as the basis for judging the real expressive quality of the music, even though such an association might have been Mahler's specific intent. Similarly, expert knowledge about the composer's intent is irrelevant to judging the expressive qualities of the music if it only points to associations of this kind.

We can also rule out expressive effects due to arbitrary associations with words, images or situations. If for instance, an opera contains a motif that appears several times whenever the villain appears, we might come to associate repetitions of that motif with an aggressive mood. Yet if that motif does not independently resemble an aggressive emotion, then we should not say that it is genuinely expressive of aggression. Of course, it is likely that a composer would only choose to make such an association if it was already appropriate to do so based on the motif's intrinsic expressive qualities. Several cases of historical associations are equally validated in this way. For example, there is a historical association between a falling minor third and a troubled sigh. Similarly, trumpet fanfares are associated with royalty and thus with a sense of power or pride. Yet the natural resemblance that trumpet fanfares have with bold and powerful movement makes it appropriate to make such an

---

<sup>5</sup> Some theorists have argued that the entire expressive content of music is the result of private associations like this (e.g. Descartes, Hanslick). Yet the analysis of chapter three has shown that the basic resemblance between sounds and feelings is sufficient for music to be expressive without associations of this kind. Moreover, explaining the expressive qualities of music on this basis would not explain the wide levels of agreement we find on the general expressive character of musical works.

association. Hence an association such as this may serve to *underline* an expressive effect due to resemblance. Accordingly it might help to draw the attention of the listener to that quality where it might otherwise have gone unnoticed (cf. Kivy 1980: chapter 8 and Davies 1994: 242). Yet such associations cannot create a genuine expressive effect where none already existed.

Another potential source of unjustified responses is our personal preferences, i.e. finding the music good or bad or indifferent. Loving a piece of music could cause us to judge its expressive quality as particularly joyful. Moreover, it may be difficult to properly divorce one's recognition of emotion in the music from one's evaluation of its quality, since presumably, judging that the music is good or bad can significantly affect one's level of emotional arousal. We might also value a work just because it expresses an emotion so well. Yet the simulation capacity of the listener should be capable of operating independently of whether the listener likes the music or not. As such it may be possible for some expert critics to disassociate their judgements of expressive properties from their personal tastes. However it would be far more reliable to appeal to what expressive qualities *different* experts could agree on, where they nevertheless have differing judgements about how good the music is.

In general the principle of dispositional realism that I articulated above is intended to preserve the idea that one can become an expert in the expressive properties of music. However, this expertise is not about making inferences based on one's knowledge of the music or its cultural context, but more about training one's perceptual faculty. In particular, one of the main ways that one can gain expertise about a piece of music is simply by repeated listening. Familiarity with an individual work often allows the

listener to pick out details that are not immediately apparent to the first time listener. These details may include counter-themes or subtle shifts in timbre or instrumentation. Familiarity also enables a listener to anticipate and appreciate more long-term structural qualities of the work, both as they extend before and after the current musical moment. As I argue below, some of these long-term features may allow the expression of more complex emotional states. In general it is quite possible for the additional details perceived within a piece of music to alter its expressive character. For example, a piece of music with a passionate theme may be heard on closer inspection to be articulated in a rather mechanical fashion. Noticing this detail will lead one to simulate the music as having a more subtle or ambiguous expressive quality.

The appreciation of details is also affected by the attention level of the listener. On some occasions, the listener's attention may wander in and out of the music so much that the expressive content of entire sections is missed. This may be the result of performing other activities whilst listening to the music, some complementary to it such as dancing, others not, such as cooking. Or it may be that the music triggers a private reverie that distracts the listener for a time. Completely undistracted and fully attentive listening is a fairly rare thing. My principle allows these distortions to be overcome because gaining expertise as a result of repeated listenings will tend towards attention to all the details of the music. It is also clear that repeated listenings help to rule out distortions caused by temporary differences in the listener's background state, such as being drunk or preoccupied by a private emotion. These conditions can cause the expressive qualities of the music to be mixed or blended with one's own current emotional state, distorting its real expressive quality.

Yet even given that the listener is in a sober and emotionally tranquil state, undisturbed by private associations and paying close attention to the music, significant differences may still result due to the ways that listeners organise the music perceptually, that is, how they prioritise certain features above others. Knowledge of the means by which a piece was produced, or theoretical awareness of resemblance effects, can allow the expert listener to more distinctly pick out the expressive features of a work. But more significantly, different styles of music recommend different ways in which the music is to be perceptually organised. For example, a listener to a typical Classical era work is likely to concentrate on the theme and the way that it develops. In contrast, a listener to a Serialist composition may concentrate on the textural and atmospheric qualities of the work. Again these different forms of attention may reflect the expertise of the listener concerning their familiarity with the genre and the conventional ways it is appreciated.

All of these different attentive attitudes could affect whether or not an emotion is recognised and the intensity with which the emotion is expressed. However I would not expect this to result in different basic *types* of emotions being recognised. Rather I would expect a refinement effect where listeners tend to agree on the basic expressive quality of the music but disagree on its more subtle, complex, or long-term features. Now it is possible that non-expert listeners could validly use their simulation mechanisms to recognise *contrary* expressive properties in the music as a result of attending to different detailed features, and then attribute those properties to the work as a whole. Yet an expert should be able to reconcile or synthesise these contradictions by being able to specify more exactly which features of the music

display which emotion. Thus there is still a definite sense that even very fine-grained expressive qualities can be really held by the music.

This claim is backed up by empirical evidence of the wide levels of agreement about the expressive content of music that I cited in the previous chapter. There is also evidence of cross-cultural agreement on basic expressive properties. In one experiment by Balkwill and Thompson (1999) for instance, Western listeners were asked to identify the emotional character of four Hindustānī ragas, which had been classed by Indian experts as expressing either joy, sadness, anger or peace. It was found that Western listeners tended to make the same judgements as the experts, though they performed slightly worse for angry and peaceful ragas. A follow up experiment showed that Japanese listeners were equally able to discern intended expressions of joy, anger, and sadness in Western, Hindustānī and Japanese folk music (Balkwill, Thompson & Matsunaga 2006). A similar experiment by Adachi, Trehub & Abe (2004) found that both Canadian and Japanese listeners were able to accurately judge whether songs by Canadian children were either happy or sad (though Canadians tended to perform slightly better than Japanese listeners for expressions of happiness).

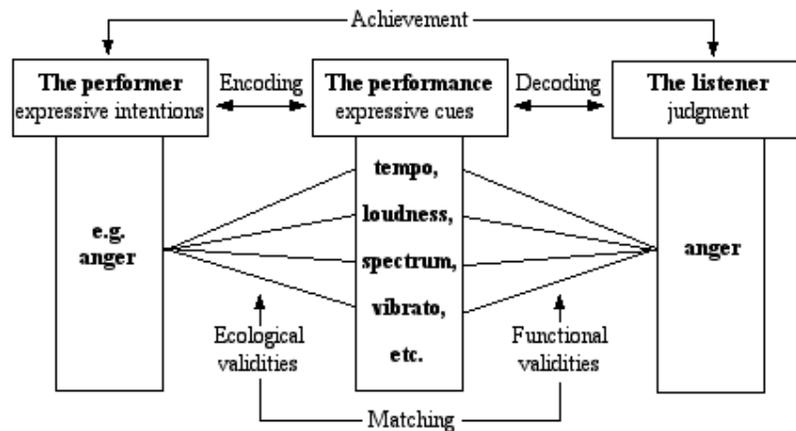
According to these results, expressive cues in music appear to transcend cultural differences of tonality, genre and cultural associations.<sup>6</sup> However these studies are limited by the extent to which people all over the world are exposed to the same influences through mass media film and popular music. In addition, these studies

---

<sup>6</sup> Stephen Davies similarly claims that Western listeners to Javanese Gamelan music (which employs a radically different tuning system to our own) would be unlikely to mistake their battle pieces for funeral music (Davies 1994: 244).

only look at very basic emotion labels. They are not sensitive to detailed changes in expressive content throughout the music which may reveal more disagreement. Some research has been attempted where arousal responses have been measured continuously throughout the music (Vaitl et al. 1993, cited in Scherer & Zentner 2003: 377). This study found that physiological responses did vary according to particular musical features such as melody and rhythm, though interestingly there was a weak correspondence between physiological responses and arousal ratings made by the listeners. Clearly more studies of this kind will be extremely useful to verify the particular role that various features of the music have. Our everyday recognitions of emotions in people are sensitive to quite subtle changes in expression. The same should be expected of musical expression.

In general however, music psychologists have been able to note several variables in music that reliably indicate the real expressive quality of the music; verifiable by using these features to accurately judge the expressive *intent* of the musician as well as what emotions listeners will recognise. These include some of the features that we have conceptually recognised as affording resemblance to emotions such as dynamics, tempo, timbre and articulation. Unfortunately larger scale details such as overall rhythm, melodic contour and tonal progression are too piece specific to draw many general conclusions. Verifying the effects of these features will require the kind of continuous measurement that I mentioned above. At any rate, Juslin organises some of the basic features that indicate emotional states into a lens model. In this model, the music is conceptualised as a filter comprising various expressive features through which the expressive communication between the listener and performer is mediated.



A Brunswickian lens model for communication of emotion in music performance

(Juslin 2001: 324)

Juslin notes that as in interpersonal expression, since several variables may equally indicate the presence of the same emotion, there is some redundancy as to whether a specific cue is required to communicate a specific emotion. Rather each cue acts as a *probabilistic* indicator of expressive intent such that where several cues are present it more reliably indicates the presence of a particular emotion. Hence although more sensitive listeners may be able to pick up on more expressive details, and thus make more *reliable* judgements, fairly non-sensitive listeners should also be capable of identifying the expressive content of a work.

This probabilistic model is helpful in articulating an account of the resemblance between music and emotions. That is, we can locate specific features of the music that causally contribute to which emotion we perceive in it. However, it will *not* fully capture what is going on in the music, even at the level of the cues identified. This is because these various features must still be blended together to have their particular expressive effect. For instance, being slow may often be reliable indicator of a sad

expressive quality, yet sometimes it is not (consider a slow high whining siren for instance). The slowness of a work may not even consistently make it *more* sad, it may sometimes contribute to a fearful quality instead. Rather in order to express sadness, we must combine slowness with other features such as pitch, harmony or a falling melodic contour. This is why each cue can only be a probabilistic indicator, and why in general we identify the *music* as having the expressive quality, rather than simply its component features. In addition, the presence of several interacting cues can indicate the presence of more specific or subtle emotional states, i.e. not just anger, but anger of a greater intensity or anger mixed with a slight sense of melancholy. It is this kind of subtlety that is important for the composer or listener in determining the special aesthetic quality of the work.

In general a more accurate idea of the expressive qualities of music would better focus less on basic emotional labels and more on dynamic emotional variables such as degrees of activation, 'darkness', or power (cf. Zentner & Scherer 2003: 381). These are the kinds of variables that can be continuously measured throughout the music. They are what will reveal fine-tuning in listeners' responses, and hopefully a more accurate indication of expertise. They also allow us to identify more consistent effects of certain features. For instance, being loud generally contributes to the work's expression of power. Hence more empirical evidence that looks at continuous responses is required to determine the specific features within musical works that underlie its overall disposition to express emotions.

## **Differences between Music and Real Emotions**

Having identified what the real expressive powers of music amount to, we are now in a position to judge exactly how much music is like emotions. Our argument has been that when we listen to music, some of the same epistemic conditions are satisfied as enable us to recognise emotions in people. The reason this can be done is because music bears some metaphysical similarities to emotional states. However, these similarities need only *necessarily* be of the most minimal kind to enable the listener to pick out an emotional state.

Yet in the previous chapter I noted that music, in terms of intensity and immediacy, is generally superior to other art forms in expressing emotions. It is even possible for musical expressions of emotion to affect us more deeply than expressions of emotions in people. I explained that this is because music bears considerable sensational and structural similarities to feelings and the bodily patterns that generate them. Thus music captures the most central component of emotional states. I also argued in chapter one that what makes a certain bodily pattern appropriate for a certain emotion is partly the behavioural responses that it underlies. Music can resemble behavioural gestures as much as inner changes. So the various bodily aspects of emotions are vividly captured by music. Moreover, these movements and feelings are not merely conventionally interpreted as *emotional* movements and feelings. They have a complexity, duration and intensity that are unique to emotional states.

However, there is more to emotional states than bodily patterns and the behavioural movements that accompany them. In particular, the primary function of emotions is

to represent general relations between subject and world (the core relational theme). Similarly, our behavioural responses only make sense in so far as they react to and manipulate those themes. By generating the illusion of a persona, music can generate a sense of the intentionality of an emotional state. Yet a more complete resemblance to emotions should capture the situations that the bodily patterns are tracking. This is precisely the aspect of emotions which, it is widely believed, pure instrumental music is incapable of expressing.

It is this consideration that grounded Eduard Hanslick's famous attack on the expressive powers of music. Since Hanslick held that emotions are distinguished by their situations and music cannot capture these situations, it followed that music could not express distinct emotional states. Hanslick's most vivid example was to show that an aria from Gluck's opera *Orfeo ed Euridice* seemed to appropriately match words of passionate love as well as words of anguish. Yet we might complain that Hanslick can only do this by extracting the aria from its full musical context. Similarly, we could extract the most heart-rending line from a tragedy and place it in a context that makes it amusing instead. So what Hanslick really needs to show is that whilst preserving the very same musical context, the aria could be expressive of both joy and sorrow.

Yet even if this was possible, Hanslick's example has been criticised by Kivy (1990) because in both cases the emotion is still a passionate one. Hence the expressive character of the music is definitely passionate even if that passion is of an indeterminate kind (cf. my comments above on experts reconciling contrary interpretations). Sometimes ambiguity is a valued aspect of artistic expression.

Meanwhile it is still the case that *other* pieces of music can more unequivocally express joy or sorrow.

Despite the problems with Hanslick's example, we may still agree that music is lacking a feature that often helps to distinguish emotional states. Of course, my argument is that we recognise emotions in music the way we recognise emotions in *other* people, and we equally do not need to know what the object of another's emotion is in order to tell which emotion they are undergoing. Yet part of the project of fully empathising with another is to try and get inside their mental state and appreciate its details from a first person perspective. In short, we often try to grasp the *reasons* for their emotion. As I mentioned in chapter two, the everyday situations in which we empathise with others are often full of contextual clues about what the other's emotion could be about. Abstract music in contrast just doesn't seem to supply those contextual reasons.

So if we want to specify how much musical works are like everyday emotional states, then perhaps we should claim that music only expresses *moods* rather than regular emotions. I claimed in chapter one that moods are about everything rather than nothing. So as long as music invokes a general sense of aboutness, it could capture such a mood. However, I think we can articulate a better defence of the emotional accuracy of music than this. Recall my claim in chapter one that the contents of emotions are certain *formal* or general qualities of situations (cf. Prinz 2004: 62-63, 185). The particular object of the emotion (say the loss of a child) is represented as having emotional relevance in virtue of instantiating a general core relational theme (loss of something valued). Otherwise emotions are blind to colours or shapes or

even the identity of objects in so far as they do not contribute to this core relational theme. Similarly the visual perception of colour only need capture surface qualities, not the identity of the object. So I think it is unreasonable to demand that music capture the complete details of a situation that might accompany an emotional state. Many details of that situation are just not part of the emotion's representational content. We can then admit that music typically lacks the particular object of the emotion. But what really matters to the content of the emotion is the general theme which the particular object instantiates.

Now as I discuss below, certain 'higher' emotions such as jealousy are more essentially tied to a certain sort of conceptual content. For the moment however, even if we stick to basic emotions, we must still account for how music captures their general content. Yet in what sense is this general content anything more than the dynamics represented by the bodily pattern? I argued in chapter one that emotions represent the self's relation to the world, so we might complain that the music lacks a self and a world. However music captures a subject by invoking a sense of a persona. Moreover, as Malcolm Budd says, it is not typically the case that every single detail of the music is responsible for determining the character of the emotion (and thus the persona) expressed (Budd 1995: 145). So it is quite possible for *other* details of the music to even provide a 'world' with which the persona interacts.<sup>7</sup> Music can provide a sense of a persona conquering that world (e.g. Beethoven's *Eroica* symphony), or one overwhelmed by the forces that surround it (e.g. Birtwistle's *Melancholia I*). So there is no reason to suppose that music cannot properly capture at least some core relational themes.

---

<sup>7</sup> Alternatively, the world of the persona could simply be the actual world. I explore this possibility when considering how a performer can identify with the emotion they express.

Another apparent consequence of lacking an object situation is that musical emotions do not capture the valency of emotions, i.e. the broad sense that emotions are positive or negative or stimulate approach or avoidance behaviours. This is suggested by the fact that we often seek out and enjoy sad or angry music, and even allow ourselves to be fully aroused by such emotions. But why would we do this if these emotions are normally unpleasant to experience? However it is clear that the very same problem arises for our enjoyment of sad books or films, which do at least provide imaginary situations. Hence this lack of valence is more likely due to the *simulated* nature of the emotion expressed rather than anything necessarily concerning the abstract nature of music. Moreover even if we agree that music can express generalised moods, the problem still arises because normally we are just as averse to sad moods as we are to sad regular emotions.

Overall there may be many different reasons why we enjoy sad art works (see especially Levinson 1997). Yet I think for the case of music at least, the simplest answer is that to feel tears welling up, to feel heavy and drooping, even to feel a pang of the heart, is not essentially a *bad* feeling. Remove the need to actually *do* anything about it, and there's no reason why the feeling of sadness can't be quite pleasant, especially when music dignifies it with formal elegance.<sup>8</sup> This does not seem like a pathological or masochistic attitude to take. Though, if one were inclined to repeatedly arouse oneself with sad music we might suspect some deeper personal issue at work. So the valence of an emotion does not seem to be an *essential* part of

---

<sup>8</sup> Kendall Walton (1978) makes a similar argument; that it is not the feeling of the emotion that is unpleasant but actually being in that state of affairs. Though we might jump out of our seats (an avoidance reaction) when watching a scary film, this is because we make-believe that we are actually in the situation.

its phenomenology in every case (though in some cases, to *suffer* or *enjoy* the feeling may be a pervasive part of its character). Of course, if we recognise that a piece of music sounds sad, we can infer that sadness is normally ‘a bad thing’. But ultimately, it is how much we like the music that determines whether we avoid it or seek it out, and as I mentioned in the previous section, preferences like these are not strictly relevant to the music’s real expressive properties.

So although it is true that music fails to capture the particular situations that typically cause emotional states, I do not think that this detracts too much from the resemblance that music bears to emotions, even the meaningful content of those emotions. The main difference is the simulated nature of the emotion expressed. And as we see in the next chapter, there are cases when even this is overcome, as when the music constitutes part of the actual emotional reaction of the performer. Also, given the lack of a specific emotionally inducing situation, it might be fair to say that the music does not express a *particular* instance of the emotion in the sense that the happiness I feel when I meet my friend is a particular emotion. Yet it is particular in the sense that it’s happening right here and now and typically to a specific imagined person (see below). So I don’t think it’s right to say that music captures happiness *itself* (e.g. Budd 1995 and Schopenhauer 1907). It would be more accurate to say that certain extraneous features of the causes and consequences of the emotion have been stripped away.

However as I mentioned above, the fact that music does lack particular object situations to which the emotion is directed does limit *which* emotions music is able to capture. In particular, it is often claimed that music is unable to express the more

complex social emotions such as embarrassment, jealousy, envy, hope and guilt. I argue below how with the aid of an imagined persona, a sequence of several feelings within a piece of music may give the distinctive sense of a more complex *blended* emotion. Yet no such technique seems available to express emotions such as jealousy or guilt.

Davies argues that the distinctiveness of these higher emotions relies on the particular sorts of thoughts or situations that cause them. Similarly, he argues that thoughts are often required to distinguish between emotions within the same general type, such as grief, despondency, dejection, gloom, moping and broken-heartedness within the category of sadness (Davies 1994: 226). These forms of sadness are distinguished by what they are directed towards. Without the context of the death of a loved one for instance, the expression of grief is indistinguishable from one of very intense and prolonged sadness.

In order to fully understand why music cannot express these states, we need to look more closely at our theory of emotions. According to Prinz's perceptual theory, the same or very similar feelings, what he calls the 'nominal' content, may track different kinds of situations. For instance, the feeling of jealousy may only be distinct from the broader feeling of anger because it is directed at romantic rivals. And it is only because some cultures hypercognize the general qualities of certain emotion inducing situations that such emotional states are distinguished at all. In these cases then, the more specific object of that emotion is necessary to distinguish that hypercognized emotional state. Since music cannot provide that specific object (at least without additional imaginative input from the listener) it cannot express

those states. If however, the bodily pattern associated with jealousy only accompanied situations of that induce jealousy rather than general states of anger, then that bodily pattern, expressible in music, would be sufficient to distinguish that state.

Levinson suggests a slightly different possibility however; that the music may provide just enough information to 'hook into' a complex emotional state without capturing its key distinguishing feature (Levinson 1990: 344-346). By way of example, he refers to a passage in Mendelssohn's *Hebrides Overture* which expresses happiness after a relatively darker period, and which in particular resembles a reaching gesture. This Levinson claims, is enough to pick out hope. Yet on carefully listening to this passage several times, even whilst guided by Levinson's description, I simply don't get the sense of anticipation or future directedness which would enable me to pick out hope. It seems at least as reasonable to suppose that it expresses a sense of relief from the earlier darker period, with a particular swelling of joy at the point Levinson believes is expressive of hope.

Now it may be that the feelings of hope have a distinct nuance all of their own, which are never the same as feelings of mild happiness. I must confess however, that I don't have a clear sense of what hope feels like. I am not sensitive enough to this particular feeling to distinguish it from happiness, though Levinson may be. Moreover, it may be that hope need not involve a particular thought. It may be possible to have a hopeful mood wherein one's life in general appears to have a hopeful quality. Without even a general idea of the future however, I doubt that it could be distinguished from a mere sense of well-being.

Yet I think it conceivable that a piece *could* suggest thoughts of the future to a listener who was very familiar with that piece. For instance, Sibelius' *5th Symphony* delays the full exposition of its main joyful theme until the very end. So in a similar case, when listening to earlier passages that seem close to this theme, perhaps involving reaching gestures similar to the *Hebrides*, the listener may be led to anticipate the thematic climax. They could then say that the current passage expresses a feeling of anticipation for this climax, it seems to reach for it. If it then also seems happy after a period of sadness (perhaps when musically speaking, the climax was becoming ever more remote) it could therefore be suggestive of the emotion of hope.

Levinson himself implies such a possibility, but only demands an awareness of earlier passages to express any complex emotion, and so such a climax (when it is actually heard) could only confirm an earlier impression of hope (Levinson 1990: 370). However, such an idea of a later part of the piece seems essential to me in order to convey the necessary sense of future directedness. Of course, it is unreasonable to confidently judge the capacity of such a piece to express hope without actually hearing it. Also the problem with this case is that it requires the listener to perceptually organise very long term characteristics of the music in a very specific way. Nevertheless, under very limited circumstances, it seems possible that music might adequately express some of the higher emotions.

### **Imaginative additions**

Overall music has genuine dispositional properties to express virtually every aspect of emotional states, though it is limited in regards to *which* emotions it expresses as a result of not expressing some more specific contents. Of course, there is nothing stopping the listener from imaginatively providing those contents for himself. It is just that additional imaginative projects may detract from the real expressive content of the music. Schopenhauer recognises this point when he says:

Certainly we have a tendency to realise [the emotions] while we listen, and to clothe them in imagination with flesh and bones, and to see in them scenes of life and nature on every hand. Yet taken generally, this is not required for their comprehension or enjoyment, but rather imparts to them a foreign and arbitrary addition: therefore it is better to apprehend them in their immediacy and unity. (Schopenhauer 1907 Vol. 3: 235)

Yet even if we restrict our imagination to the emotions expressed, the ways in which people conceptualise emotions can differ. Of course, the whole thrust of these last two chapters has been to show how in all cases the experience is based on a feeling of bodily patterns. Yet there is nothing to prevent listeners from elaborating that experience to themselves in different ways. In particular, the imaginative addition that I want to explore here is the different ways that listeners can ‘flesh out’ the persona they hear in the music. It is important to explore these different imaginative possibilities because in these cases the imagination is still very much caught up with, or guided by the expressive qualities of the music. It seems that if the listener must imagine a person, then they often imagine the persona in some quite definite form.

There then seem to be various contextual reasons why a listener might imagine one form of persona rather than another. However, it should be noted my predictions of these contextual factors are rather speculative, though they should be susceptible to empirical investigation.

The two main kinds of persona that I briefly indicated in chapter three concerned whether the persona was imagined as somehow ‘behind’ the music or embodied by the music itself. In the first scenario listeners imagine that some person is producing the music, either in a conventional manner, such as by playing an instrument or singing, or simply by making the music emanate from their body in some magical fashion. In this case then, the persona is imagined to undergo an emotional state that is directly responsible for the sounds that the persona produces, and in which the quality of the sound reflects the character of their feeling. This first kind of persona seems most likely to be imagined when it is easier for the listener to perceive how the music may be produced by a person, for example, if it is obviously a piece for solo instrument. Moreover, if the listener has background experiences of the production of that kind of music, either because he is a musician, or because he is currently observing musicians in the very act of producing the work, then this form of imagined persona is more apt.

In the second scenario the music is imagined to *be* a person that exists in musical form. Saam Trivedi describes what I take to be this possibility when he says we imagine the music as “the very being that is animate and whose emotions are being expressed musically” (2001: 414).<sup>9</sup> Malcolm Budd similarly describes cases where

---

<sup>9</sup> Though Trivedi claims that this does not count as a persona because he adopts a rather restrictive view of a persona as something automatically distinct from the music itself.

you imagine the music to be an instance of the feeling (1995: 148). Both descriptions pick out a persona who possesses the feeling, and though this persona is embodied by the music, it is still imagined to have roughly human characteristics, or at least to move in human-like ways. Yet the main feature of this persona is that by hearing the expressive music we imagine that we directly perceive its emotional feelings, as if we were hearing inside its body.

It seems to me that this kind of imagined persona requires a little more imaginative work on the part of the listener, given that it is quite unlike the creatures we perceive in everyday life. Yet this persona also seems more likely when the music is one of large-scale orchestral forces, i.e. the sound has the kind of solidity and continuity that makes it more appropriate to hear it as a single living thing. Note also that the relation of the persona's emotion to its musical body could vary according to which features in the music are perceived as directly bearing on quality of the emotion expressed. For instance, if the emotional quality of the work only seems to be expressed by the particular emphasis or attack on the notes, then the other features of the work may be imagined as the overall body that bears those particular emotional signs, analogous to a facial expression on an otherwise unexpressive body. On the other hand, if a large number of the work's features all seem to contribute to its emotional quality then the whole work may be imagined as a body leaping about in an expressive manner. In general the two main forms of persona I have identified could be mixed such that some features of the music *embody* the persona, and other features are imagined as *produced* by the persona.

A third possibility that Budd suggests is that the listener may imagine undergoing the emotion himself. This is also a case of a persona, since the listener utilises a *self-conscious* idea of his own person or at least his own body, to make sense of the emotion in the music. The particular way in which the music's expressive qualities are identified (as an expression or instance of the feeling) can then correspond to either or both of the two major types of persona I outlined above.

Identifying oneself as the bearer of the music's expressive content is a particularly interesting variety of musical experience because the listener allows himself to become deeply personally absorbed in the music. In these cases the listener may also imagine some scenario or think about some specific situation in his own life that the emotion represents. It seems that this form of imaginative engagement is most likely when performing because the musician should have a sense of control over the features of the music that enables a corresponding sense of personal ownership over the expressive content. However, we see an interesting contrasting view in an interviewee in Evan Eisenberg's book *The Recording Angel*:

When I play a record... it's as though someone else were expressing my feelings. When I play piano, it's as though I were expressing someone else's feelings. (quoted in Eisenberg 1988: 132)

Depending on the context, the listener imagines two different sorts of persona (either himself or another). Perhaps when playing a piece composed by someone else, the listener feels like they are controlled by that composer, or channeling the composer's emotional state. In contrast when they put on a record, they are more able to identify

emotionally with the music because the composer or performer seems more distant, though they still have a sense that another person mediates the emotion. Identification with recordings is also plausible because listener can become extremely familiar with the content of the music and so is able to anticipate its every nuance. In all these cases there are variations according to the sense of control and so in general I would expect the higher the degree of control, the more likely the listener will personally identify with the emotion expressed.<sup>10</sup> Accordingly in the next chapter I analyse the case of improvising musicians as paradigm examples of identifying with the music they produce.

Finally, it is worth mentioning a quite common variety of musical experience in which the listener imagines a visual scene, as if the music were accompanying a film. Now the abstract nature of music lends itself to more than emotional interpretations. Sometimes listeners even imagine conceptual arguments to fit the music. Yet the reason I mention this particular imaginative addition is because it seems possible for that visual scene to be taken as an *emotional* interpretation of the music. That is, the listener has an implicit sense of a person (probably themselves) *confronting* that particular scene and having an emotional response as a result.

This imaginative project seems most likely when the emotional personality of the listener is such that when they have emotions, they tend to focus more exclusively on the object or event causing that emotion than on the way their body feels (cf. Laird &

---

<sup>10</sup> Whether or not the listener identifies with the subject of the emotional music may potentially affect how we label the emotion perceived. One interesting possibility is that when listening to the kind of music we typically call scary, taking a first person perspective may make it appear threatening (angry), where a third person perspective may render it threatened (i.e. fearful). It is still the same feeling that is being expressed, it's just that the perspective on that feeling differs. Perhaps this could explain the occasional confusions between these two forms of musical expression that have been found in empirical studies (e.g. Terwogt & Van Grinsven 1991 cited in Dowling 1999).

Bresler 1992). As such they may have developed very strong associations between certain emotions and certain scenes. Clearly listeners who engage in this sort of imaginative project will have quite different experiences. Yet if that visualisation is based upon the proper emotional qualities of the music, then although the potential variety of scenes that the listener may imagine is vast, it should be possible for other listeners to accept the appropriateness of that scene for the emotion in question.

Having now identified the various forms of the persona in music, there are then several ways in which the listener's more developed imaginative construction of the persona can differ. Perhaps most dramatically, the listener might identify the music with the emotions of *several* personae rather than just one, either all simultaneously, or sequentially one after another. This is a possibility that Davies raises when he complains that a persona cannot provide unity to a work of music that has none already (Davies 1997: 106 cf. Vermazen 1986: 202). The persona theorist should admit this at least. The only necessary role of the persona is to allow the expression of mental states at all.

Yet if by its formal or thematic unity, the music suggests that a single persona stands behind several emotional transformations, it seems possible for music to express more complex emotions. For instance Ravel's *Piano Concerto for the Left Hand* expresses feelings of sorrow transforming into feelings of anger and then joy. When the listener imagines a single person undergoing all these successive states the music may express the more complex blended emotional state of defiance. By developing my imagination of the persona I am then able to fill out this emotion in a variety of ways. Since I know that Paul Wittgenstein, who lost his right arm in the First World

War, commissioned the piece, I can imagine how awful it must have been for a concert pianist to lose his arm. So I ascribe the sorrow and anger in the music to the sorrow and struggle of Wittgenstein when trying to overcome his disability. Then, since he is still able to overcome his disability and perform a great piece of music, the following joyful feelings express his feeling of joyful defiance. Joy succeeding sorry becomes joy *about* overcoming sorrow.

Similarly, when the listener is well aware that a particular person is responsible for the music, their understanding of that person's *character* may lead them to interpret the emotions they perceive in the same way as we do when empathising with people we already know (see chapter two). For instance, knowing that Beethoven composed the music, the listener may think 'Beethoven was a bad tempered but passionate man who overcame his loss of hearing'. This information then feeds back into the sense of the emotion expressed, and a sense of joy becomes one of heroism.

Of course, it is by no means necessary to imagine such elaborate programme notes attached to the music. Yet once the persona conception is invoked, the imagination of the listener may develop it however they wish, and as a result experience the expressive content of the music with all kinds of additional nuances. Moreover, in any given listening situation, which of the above elaborations are employed will largely depend on the listener's background attitudes and beliefs about musical experience. Yet as in the case of imagined situations, we would expect the personae that listeners imagine to be appropriate to the expressive content of the music. For instance, one would not typically imagine a punk rocker embodying Bach's *Goldberg variations*. In addition, it seems that certain aspects of the listening

environment will push that elaboration in one particular direction rather than another. As such, alignments in the listeners' background attitudes and listening situation may well result in a corresponding alignment in their imaginative engagement with the music. Unfortunately, there do not appear to be any empirical studies that try to confirm this possibility. Yet I would predict some significant affects in the responses of listeners both at the level of immediate arousal and their imaginative projects when listening in groups as opposed to individually.

### **Music and Colour**

Overall I have built a case for saying that expressive qualities really are in the music. The emotional experience that these expressive qualities generate in us is of a simulated or fictional kind. Yet this is mostly because the perceived emotion is not representing a real situation. In the next chapter I present cases where despite utilising the simulative capacity of the performer, the music can qualify as partly constituting a real emotional state.

In general we can legitimately say the music possesses the expressive quality in the same way as objects possess colours (where colours are construed as secondary qualities). The music possesses a dispositional property to express to suitable listeners the intrinsic nature of an emotional state. There are of course differences in the ways that listeners understand what it is they are hearing, and the kinds of imaginative activities they engage in. Yet ultimately, these differences should be grounded in the resemblance that music bears to emotions. Hence the different responses of listeners are likely to be refinements of more universal responses rather than radically different. Our responses should at least be intelligible to the expert

listener, if not also to each other. This allowance for expertise does not make emotions in music less dispositional than colour. We might equally say that some people are more exquisitely sensitive to colour variations than others.

As I noted above however, one important difference between colour perception and emotional perception in music is that our capacity for the latter has in all probability evolved for a different purpose. That is, our recognitions of emotions in music are situated within a context of recognising emotions in interpersonal contexts. Now given that the mechanisms by which music expresses emotions are largely sub-personal and innate, it is conceivable that a piece of music could express an emotion to a listener that he had never before experienced, or at least a more specific variety of that emotion. Yet in general our experiences of emotions in everyday contexts are liable to play a significant role in developing sensitivity to the expressive qualities of music. This is particularly relevant to musicians, who it is typically claimed must mature emotionally before they can *express* emotions very convincingly (a musical performance can equally be ‘beyond one’s years’ in more than just technical respects).

In this way the background experiences or personality of the listener may cause them to be more susceptible to certain emotions that should allow them to recognise its occurrence in music more readily. Similarly, different listeners may be used to dealing with emotions and the way they are expressed in different ways. For instance, men might be more inclined to suppress their sensitivity to emotions that indicate weakness. Furthermore, even if listeners do recognise the very same emotion, the different affects that such recognition can have on different listeners could be

profound. One listener may be completely jaded, and regard the expression as sentimental or silly where another is moved to tears. In my own experience, there have been some occasions in which a piece of music has seemed to capture the entire human condition. I suspect however that most other listeners to those pieces would not have experienced such an intense reaction.

Music seems to go deeper into the human experience of feeling than any other art form, and can drag these experiences out into the open for all to see. I mentioned above how we may be more sensitive to emotions in music than we are to emotions in people. It seems plausible to suppose that when an emotion is so obviously attached to definite person with their own concerns, (that may conflict with our own) our sense of being separate from that person is sometimes intensified. Yet when presented in abstract there is no such barrier to embracing the emotion expressed as one's own. Similarly when there's no specific object or state of affairs that the emotion is directed at, the imagination is free to associate the expressive content with whatever situation one likes, or if contagiously aroused by the music, to introspect on one's own physical feelings and to allow them to be shaped and transformed by the way the music progresses.

Yet despite the abstract nature of emotions in music, the emotions expressed are not *sui generis*. If music expressed completely distinctive emotions then it would be quite mysterious why we describe it using everyday emotional terms, or why music arouses and absorbs us to the extent that it does (cf. Davies 1994: 203). The main difference between music and everyday emotions is that music *idealises* the emotional experience. As Susanne Langer says, "sheer self expression requires no

artistic form” (Langer 1951: 184). But in its role as an art form, music extends, reflects and transforms common everyday feelings into sometimes intensely poignant, sometimes heroic attitudes. This is a difference of exaggeration or complexity rather than a difference in kind of emotion expressed. Music makes us concentrate more on feelings than we might otherwise do so. As such the capacity of music to allow the deep appreciation of emotions, as well as give them a sense of order and completion is one of the most valuable experiences that art can provide.

## Chapter Five: Expression and Extended Cognition

The 'realist' expression theory was rejected in chapter three because it was not necessary for the music to be causally related to a composer's emotional state for that music to be expressive. However, it was later granted that on a phenomenal level the listener treats the music *as if* it were the expression (or embodiment) of a person's emotion. In this chapter, the expressive content of music is considered from the standpoint of the composer or performer. I claim that whilst the composer or performer need not express his own emotional state, it is still a value that may be effectively pursued, and for which music is highly suitable. Moreover I will argue for the possibility of an extremely intimate connection between the emotional content of the music and the emotional state of the person who produces that music. Under certain specified conditions, the music may not just influence, but also partially constitute the musician's emotional state.

The most significant of these conditions requires that the musician be able to control the content of the music in an immediate and detailed way. Allowing for this level of control leads me to focus on jazz improvisation. Having focused on this style of music I then analyse some of the explanations for how it is that jazz improvisation is accomplished and in what ways this contributes to the value of the genre. In pursuit of a deeper explanation for what I call the 'individuality' of jazz improvisation, I then explore the expression theories of Benedetto Croce and R. G. Collingwood, in particular the idea that the artist clarifies his own emotional state by expressing it. Whilst this central insight is preserved, I emphasise the essential role of the artistic medium (in this case the musical instrument or the sounds produced) to that state.

This leads me to the theory of extended cognition (also known as vehicle externalism), which argues that the physical constitution of some mental states extends beyond the brain of the subject. After showing how this theory applies to improvised performance, I then justify two claims: Firstly, that playing the instrument cognitively extends the musician's creation of the music. Then secondly, and more significantly, that playing the music cognitively extends the musician's emotion.

The first claim is argued to be true of all cases of jazz improvisation. The second claim is only true of a subset of cases, yet it captures one of the central goals of jazz performance. Given this, the second claim then breaks down into a stronger and a weaker version. The weaker version states that the music *elaborates* the emotional state of the musician, playing the same role as his bodily changes in generating and maintaining the bodily pattern. The stronger claim is that the music *replaces* the role of bodily changes in generating the bodily pattern.<sup>1</sup> And thus that the musician's experience of the music more fully constitutes his experience of his emotion.

### **The Musician's Emotion**

One of the key problems with the realist expression theory was that it left unanswered the question of how it is that feeling a certain way whilst producing a work is supposed to convey these characteristics to the work itself. As a typically mediated and highly contrived process, music composition is unlike painting where potentially a violent stroke of the brush could be recognised as such by the viewer.

---

<sup>1</sup> Note that since conscious attention to feelings is not necessary for emotional states, the experience of an emotion is distinct from the emotion itself (see chapter one). Accordingly, I do not argue that the *experience* of the emotion is partially constituted by the music.

Even in the more direct case of performance, it is not the case that the emotion of the musician is necessarily transmitted to the work. In order to explicate this, it is worth distinguishing several levels at which the emotional state of the musician may affect the character of the music:

First of all, the musician may be in an emotional state prior to performance. Performing whilst in that state then affects the character of the music. For instance, being anxious causes the musician to make mistakes. A second possibility is that the emotional state of the musician affects the music in such a way that the character of the emotion is reflected in the character of the music. For instance, the musician feels anxious, and playing the music in an anxious way causes the music to have an anxious character (perhaps a tremor in the hands leads to a tremor in the sound). A third possibility is that the musician is in relatively neutral state prior to performance but is fully aroused by the character of the music he is playing (as a result of not inhibiting the simulative recognition of the music's content). It may then follow that being infected by the character of the music affects the way the musician plays that music. Finally, a fourth possibility further allows that this changed performance style transmits the quality of the musician's emotion back to the music. Hence a feedback loop could be generated whereby the music arouses the state of the musician, affecting the way he performs, which then affects the character of the music, potentially further arousing the musician and so on.

In all cases, we should note that the emotion of the musician need not enhance the expressive value of the music. For example, taking the fourth possibility mentioned above; suppose that I play a piece of music expressive of anger and become aroused

by anger as a result. This then causes me to play more recklessly, striking the keys of my instrument more forcefully and quickening the tempo. As a result of this the music expresses anger even more convincingly, but at the same time, I hit lots of wrong notes, my timing is disrupted and soon the entire performance loses coherence. From one perspective this may be an interesting performance, but more likely it will be regarded as a distortion of the work, perhaps spoiling the planned development or resolution of the emotion conveyed by the music. Thus there is a difference between a musician's aroused state affecting the performance and the musician using that arousal in a controlled way to enhance the performance. It is important that the goal of expressing one's emotion does not conflict too much with the more intrinsic goal of conveying what is indicated in the score.

Due to the high complexity of much classical music (and its corresponding difficulty) a cool, calm approach is usually demanded to ensure a successful performance of the work. As such, musicians are required to inhibit their emotional responses to the music. Antonio Damasio relates an experiment in which he monitored the physiological activity of pianist Maria João Pires as she listened to a familiar piece of piano music and found that she could either allow or restrict "the flow of emotion to her body" at will, reducing her physiological responses, including lowering her heart beat and flattening her skin conductivity (Damasio 2000: 50). Since this experiment has not been performed on other musicians, we cannot say how common this ability is. Though we can admit that the ability to inhibit one's own feelings is at least encouraged by classical performance practise.<sup>2</sup> Yet encouraging emotional inhibition

---

<sup>2</sup> It may be that learning to perform enables musicians to develop greater control over their emotional responses or that people who already possess greater control over their emotions make better musicians. Interestingly, musicians tend to be 'bold introverts', that is, they show several introvert tendencies (such as detachment and self-sufficiency) but can also be gregarious and expressive, so

sets up a tension within the classical performance situation. In order to ensure that the emotion expressed by the music is effectively conveyed, the musician must utilise their understanding of the way that emotion feels. This is particularly true of works where an expressive label (such as 'agitato') has simply been added to the score without specifying in detail how this is to be achieved. Understanding the emotion in the music will require the performer to use their own emotion generating mechanisms (just as in the case of the ordinary listener). Hence the classical performer must strike a balance between sensitivity to the emotional state they are trying to represent, whilst resisting getting carried away by it.

Of course, it is possible that with exceptional control over the technicalities of performance, a certain amount of attention could be freed up to allow oneself to be more fully aroused. Yet ultimately, given the concern within the classical genre for fidelity to the score, the performer does not have much freedom to allow their emotion to affect the character of the work. Even quite subtle changes to the timing and dynamics of a piece can be detrimental to its coherence. Moreover, beginning a performance in an emotional state is unlikely to be beneficial since there is no reason to suppose that one's state will appropriately match the emotion to be expressed in the work. On some occasions the performer may be able to decide which work from their repertoire would best match their state. Yet given a limited repertoire this is unlikely to be a very fine-grained choice and will allow only generic matching of expressive music to emotional state. So one would need more radical choice over the content of the music in order to exploit one's emotion appropriately.

---

long as it's on their own terms i.e. using musical performance (Kemp 1997: 27-28).

In contrast to classical performance, improvisatory performance allows the musician far greater control over the character of the work. Jazz in particular is much more flexible about what counts as right or wrong notes, emphasising spontaneity and communication over technical accuracy. As such, this genre should permit greater levels of emotional arousal in the players and a more immediate connection between the state of the musician and the nature of the work. There are still limits of course. Jazz improvisation is not the pure spontaneous production of music that it is sometimes taken to be. In most cases there is some kind of pre-planned reference material such as a written chord structure, melodic theme or background tonality that the musician uses to guide his performance. This reference material can then be more or less detailed as well as more or less possible to ignore. At one end of the scale, the performer has to follow a pre-written part and can only add his own emphasis or elaboration to it. At the other end of the scale, there is no referent material at all and the performer is not compelled to follow any kind of theme or tonality whatsoever (cf. Smith & Dean 1997: 32).

For the sake of simplicity I am only interested in solo performance here. But one of the main reasons reference material is used at all is to enable several musicians to collaborate effectively. Again, the sounds produced by other musicians can be more or less embraced or ignored. Sometimes the soloist engages with the other performers by imitating or developing what they play in a quasi-conversational way. At other times, the other musicians are only there to provide a minimal 'pad' from which the performer can launch his own flights of imagination. In chapter seven I propose an ideal balance between sensitivity towards other musicians whilst still allowing individual expressive immediacy. For now though I take as my paradigm a

relatively unconstrained solo improvisation style which whilst sticking to the stylistic forms of the jazz genre, imposes no time constraints or adherence to a particular chord structure. This should be of ample sufficiency to let the performer's emotion significantly affect the character of the music in a way that properly corresponds to the character of his feeling. It is thus closer to Modal or Free Jazz than Bebop.

It's not the case that I am reliant on this specific kind of improvisational style or even on the jazz genre generally in order to justify my claims. It's likely that other forms of improvisational music may equally permit the kinds of expressive potential I argue for here.<sup>3</sup> For example Ravi Shankar claims that at the creative peak of improvisation he loses all awareness of the strings of his sitar. He feels so directly fused with his instrument that creating music seems "more a release of inner energy than an activity requiring conscious effort" (Slawek 1998: 338). Yet I restrict my attention to jazz improvisation, partly due to my familiarity with this genre as a performer and listener, and partly because historically it has emphasised the personal expression of the musician.

### **Musicological Explanations**

So I am interested in the extent to which jazz improvisation enables the musician to become emotionally engaged with the music. In addition, I claim that a high level of emotional engagement is not simply a possibility in jazz but one of the main goals of improvisational practise. It helps us to answer two very general questions we might

---

<sup>3</sup> It may also be argued that the presence of constraints can actually aid expressive engagement because it leaves the performer one less thing to worry about deciding (cf. Pressing 1998: 52). Moreover, in Free Jazz, it is possible that the extreme lack of structure can itself be constraining if the musicians constantly try to play sounds as different as possible from anything else they have played before.

have about jazz improvisation, which I distinguish as the ‘how’ and the ‘why’ questions. The ‘how’ questions concern how it is that the musician is able to generate the music at all. Given all the musical possibilities that are available, how does the musician choose what notes to play? The ‘why’ questions then concern the aesthetic value of improvisation. That is, what makes this style of performance valuable for the performer as well as the audience? What distinguishes good from bad improvised performances?

To some extent the ‘how’ and the ‘why’ questions are related in that the means by which an improvisation is generated should reflect what makes it worthwhile. At the least, the very fact that it is improvised should have a significant role to play in determining its special aesthetic value. Conversely, a consideration of why improvisation is done should give us some clues as to the manner in which the musician goes or should go about doing it. So it is my hope that my claims about the extended cognition of emotion in music help explain how improvised performances are achieved and also give us insight into why improvisation can be such an aesthetically valuable activity, especially from the standpoint of the musician.

However we can go some way towards saying how and why improvisation is done without having to appeal to the notion of extended cognition. We see that it is only when we pursue the questions to a fairly deep level of explanation that the justification for a stronger cognitive interaction between the musician and the music becomes apparent. Hence I first explore some of the musicological explanations of jazz improvisation before arguing for my extended cognition account.

To begin by addressing the how question, we can appeal to the familiarity of the musician with the conventions of the jazz idiom, his absorption in its stylistic language. For instance, the repetitive practise of various jazz ‘licks’ can provide the musician with a great deal of musical material that can be strung together to form a jazz solo. Yet the question then becomes, given various choices of appropriate licks, how does the musician choose which one to use at any given moment? We may respond that to a large extent, certain figures will simply fall more naturally under the fingers of the musician as he plays, particularly when he is playing very fast. Given that his hands are already in some position on the instrument, muscle memory can take over and find whatever acceptable figure happens to lie closest to that current position.<sup>4</sup> However the performer is under no obligation to go for the easiest choice in any situation. There is nothing to stop him from being musically adventurous and leaping to a very distant hand position if he so desires.

More importantly, improvisation is at its worst when the performer merely strings together ‘correct’ sounding notes without any sense of musically meaningful relations between these notes. Hence many jazz musicians emphasise the idea of telling a story, or having a conversation or argument with themselves through their playing. These analogies describe the way that the musician must be constantly looking back to what he has already played, and then repeating, contrasting or developing these figures in various ways such as by adding notes, translating the whole figure up or down, or stretching the intervals and rhythms involved. In addition the best players should be able to simultaneously develop a larger scale shape to the solo, such as building up in intensity or dissonance and then resolving it at the last moment.

---

<sup>4</sup> This observation will become relevant to my extended cognition account later on.

So we have an initial answer to the question of how the improvising musician is able to generate music, as well as some clues about the relative quality of different performances. Of course, the question now becomes; where do these larger scale structures come from? As the musician develops his material in performance, what guides him to take it in one direction rather than another? Are there any principled ways in which jazz solos are supposed to be structured?

One possibility is that the musician absorbs certain large-scale structural principles from his previous experiences of listening to jazz music, which then become organised around mental prototypes of what counts as a great jazz solo. There is some initial motivation for this idea in that many musicians describe their introduction to jazz music in terms of being inspired by some great player that they then sought to emulate. At the same time, we can suppose that this emulation takes practical form in terms of developing an idea of what their musical heroes' solos tended to sound like, which they then try to capture in their own playing. This emulation need not be pre-planned, or even particularly obvious to the musician. They may simply have an intuitive idea about what a jazz solo should sound like and try to match this intuition. We can also see how such emulation might be reinforced by rehearsal as well as how in the course of performance, the musician may find himself reminded of some great player's sound, and then allow this memory to determine how he continues.

This then is one of the ways that the historical continuity of jazz music may be established as an oral tradition passed from one generation to the next. However, I

don't think we should be satisfied with this answer because as much as jazz musicians may want to maintain the spirit of the jazz sound inherited from their musical ancestors, they are also routinely criticised if they fail to demonstrate their own individual sound. For instance, Jazz analyst Ajay Heble sums up a variety of perspectives on jazz improvisation claiming:

What all of these critics, despite their varied emphases and points of focus, have in common is their insistence that improvisation is a powerful ally in struggles for self-expression, self-determination, and self-representation.

(Heble 2000: 93)

It is taken as a sign of musical maturity in jazz circles that the musician transcends their musical influences in finding a more distinctive voice (cf. Berliner 1994: 256, 274). Moreover, jazz musicians often claim to enjoy taking risks or surprising themselves in their performances. Yet if all they are seeking to do is replicate some pre-conceived ideal, why bother with such risks, or even to improvise at all? I don't think we can claim that jazz musicians improvise just because that's what their predecessors did. In order to justify improvisation as a significant musical practise we should look to values that would motivate the musical heroes as much as the modern players. Otherwise we can expect the virtues of improvisation to become progressively de-emphasised in jazz the way it has in classical performance. Thus, it seems to me that improvisation is not just a means of demonstrating virtuosity or as establishing oneself as part of a traditional community, but an important way in which this genre embodies the artistic goals of creativity, individuality and social engagement. It is the constant experimentation by good jazz performers, combined

with, and ultimately intensifying the personalisation of the jazz sound, that enables this genre to develop. Hence any account of jazz improvisation should try to show how it achieves these values.

So we end up with the claim that improvisation is a way for musicians to demonstrate their individuality. However, if we look more closely into what it is that constitutes this individuality, i.e. having a distinctive kind of sound, we might again find it difficult to explain why this is best achieved by improvisation. What makes one player audibly distinguishable from another may be a certain expressive tone or style of emphasis, use of particular motifs, favoured intervals and harmonic voicings, or the melodic ways they structure their solos. But why can't the jazz musician simply sit down before the performance and plan out all these distinctive features? It seems that it would be easier to create an unusual and unique musical performance if you had a chance to really think about it, allowing you to conscientiously avoid sounding like anyone else you can think of. This certainly seems to be the preferred method of the classical composer anyway.

Perhaps it is because individuality must be tempered by coordination with the other musicians involved, and so improvisation is geared towards the flexibility to deal with the contingencies of the performance situation. Yet while this is certainly an important consideration, it seems equally possible that *all* the musicians involved could collaboratively plan what they were going to play beforehand. They could then all get to be unique and distinctive without worrying about treading on each others' toes too much. So we are still left with the question of how improvisation *as improvisation* contributes to the sense of individuality.

I think the answer here is that the particular sense of individuality the jazz musician is driving towards is one characterised by ideals of personal freedom, the naturalness of playing what ‘feels right’ at that particular time, and the kind of sincerity achieved by committing to the moment of performance. Equally, the audience can appreciate that the improvisation is something unique to this particular situation that helps to generate, as well as reflect the mood of this moment in time. It is for this reason that I think the structure of an improvised solo is one ideally derived from the way the musician feels at that moment and their sense of the actual musical situation at hand. Hence we can appreciate what jazz trumpeter Lonnie Hillyer means when he says:

All I am looking for is to get what I feel through the horn. When I was younger it was like trying to get what Dizzy felt through the horn, but, eventually, it’s got to come to be about you. Emotions are what it’s all about. (quoted in Berliner 1994: 261)

So one of the ideals of jazz improvisation is the immediate emotional expression of the musician. However this is not to say that musicians will simply regurgitate whatever they happen to be feeling at that exact moment in musical form. If this was all there was to it, they could just as easily shout or groan with the same expressive success. Rather I claim that the music is a means for the musician to self-reflectively *construct* their feelings, to control and transform their felt states into something more artistically profound. It is with this in mind that we now turn to the Expression theories proposed by Benedetto Croce and R. G. Collingwood, who both developed this very idea.

### **Croce and Collingwood**

Croce and Collingwood both argue that the central purpose of art is to express the emotions of the artist. In particular, the music becomes something that the musician uses to think about himself, to explore and articulate his emotional state. Yet unlike Croce and Collingwood, I do not claim that emotional expression is the central purpose of art. I simply argue that it is one of the most important and valuable goals that the improvising musician can seek.

Croce derives his idea of expression from a more general theory of mind. He claims that before we can conceive of objects within general categories (such as whether something is real or unreal, or using Kantian categories like space and time) we must first grasp them as individual entities. This is what he calls intuition. He then goes on to identify intuition with representation and expression. The reason for this is that understanding an object as an individual involves an active process of imposing order upon the raw impressions that we receive by fusing them together into coherent form. And we do not properly understand an object as an individual until, by imposing structure, we are able to articulate its particular characteristics. So for example Croce says, “How can we truly intuit a geometrical figure if we do not have so clear an idea of it as to be able to draw it right away on paper or on a blackboard?” (Croce 1992: 9).

Thus Croce thinks of expression as our primary act when understanding the world. Only after we have imposed some structure upon our impressions are we then able to distinguish objects from each other and form more generalised concepts. The same

process applies to emotional states. We begin with the raw impressions of our feelings and by expression form them into distinct perceptions. Generalising these states under labels like 'sad' or 'happy' is then part of the later conceptualising stage.

Similarly to Croce, Collingwood draws a connection between expression and an immediate kind of understanding:

Expressing [one's emotion], we saw, has something to do with becoming conscious of it; therefore, if being fully conscious of it means being conscious of all its peculiarities, fully expressing it means being conscious of all its peculiarities. (Collingwood 1958: 113)

It is on this consideration that Collingwood bases his well-known distinction between art and craft. Because expression is a matter of becoming conscious of one's emotion, the artist can have no detailed idea of what emotion he is expressing until he actually completes it. In contrast crafts (such as carpentry) employ a much clearer means-end distinction, where the craftsman has a definite goal in mind of what he wants to produce prior to executing it.<sup>5</sup> If we argue that sometimes artists also have a clear idea of what they want to produce prior to executing it, Collingwood will argue that by this point the process of expression (and hence the main goal of the artist) has already taken place. Actually producing the artwork is just something that enables other people to access the expressive idea of the artist. This is what leads us to the familiar claim that Collingwood's theory is idealist. For instance he says, "a piece of

---

<sup>5</sup> Though contra Collingwood, we may argue that not all crafts are completely specified prior to execution. For instance, a cook may gradually refine a meal as he cooks; adding ingredients until he perceives some exquisite balance of tastes has been achieved. Does this make cooking an art?

music is not something audible, but something which may exist solely in the musician's head" (1958: 151). Here he seems to be stating that there is only a contingent relation between the *true* artwork, the artist's expression, and the physical work before us. In other words, the existence of the artwork is sufficiently realised if the artist merely articulates it to himself mentally. Croce appears to take a similar stance:

The aesthetic stage is completely over and done with when impressions have been worked up into expression... the work of art (the aesthetic work) is always internal; and what is called the external work is no longer the work of art. (Croce 1992: 56-57)

Like any other object, the physical work is only so much raw material until the mind is able to organise the impressions it gets from it into an expression. Hence the physical object produced by the artist is just the stimulus for the construction of the true aesthetic object in the mind of the artist (or hopefully audience). So even if we argue that artists do not typically have a fully articulated idea of their works prior to physically producing them, Croce can say that these artists;

... make brushstrokes, not to externalise their expressions which do not then exist, but as if to try out and to have a simple point of support for their internal meditations and contemplations. (Croce 1992: 114)

This idea of support forestalls an objection based on the implausibility that mental images can be as distinctly articulated as most complex works of art. The physical

production of a complex work need only stimulate the raw impressions necessary to remind the artist of the complete expression. The problem is, if the external artwork now functions only as a reminder, it seems to be reduced to a mere heuristic device to get at the expressive idea, i.e. that potentially the artist could just as easily use some other means, or none at all.

Separating the true artwork from the actual physical production is counter-intuitive since, as Richard Wollheim argues, the difficulties (and unpredictability) that accompany the handling of the physical medium are often a vital part of what makes the work expressive (Wollheim 1980: 42). Moreover, any ideas that the artist has prior to producing the object are ideas about how the medium may be manipulated. They are ideas *of* the medium and presuppose its actual physical characteristics. So for instance, when a jazz musician expresses their emotional state in performance, they must think of that emotion in distinctly musical terms. It is the qualities and relations obtained by the actual physical sounds (sometimes accidentally) that constitute his emotional language. Similarly Aaron Ridley, who attempts a defence of Collingwood's theory, immediately rejects the ideal theory using the example that "the peculiar fluidity and grace and power of Bernini's *Ecstasy of St. Theresa* are inseparable from the fact that it has been almost caressed out of a piece of stone" (Ridley 1997: 264). Similar arguments could be generated for musical performances where the stretching of a particular musical instrument's capacities is part of what it is for that music to express an emotion like boundless joy.

Though he only analyses Collingwood's theory, Ridley's defence could equally apply to most of Croce's claims. Instead of trying to justify an idealist interpretation

of Collingwood, Ridley argues that Collingwood was only concerned to highlight the importance of active imaginative engagement when confronting artworks. Ridley appeals to Collingwood's analogy of listening to a lecture (Collingwood 1958: 140-141). Just as a lecture may seem like mere noises to a person who cannot understand English, so the very same string of sounds may be heard as the exposition of theories to one who hears with understanding. This latter person is imaginatively engaging with the *meaning* of the lecture via its sounds. Ridley then argues that this distinction does not imply the existence of two possible objects, one a set of noises and another a set of ideas, but simply that one thing may be experienced in two different ways thus:

[A]n imaginative experience (of music, say) need not be an experience of something imaginary (in someone's head), but might rather be a *way* of experiencing something real (such as music). (Ridley 1997: 265 emphasis in original)

This then would make Collingwood's claims compatible with the kind of dispositional theory of expression that I raised in the last chapter, where an object expresses *x* in terms of its dispositional property to cause the idea of *x* in an appropriate viewers in appropriate circumstances. So when Collingwood says that the music may exist solely in the musician's head, all he really means is that the musician may be the only person who properly understands the work. It does *not* follow that the external work does not actually possess properties expressive of that meaning, and thus that it is inessential for accessing the expressive idea.

Even if we understand Collingwood (and Croce) in this way, the issue that interests me here is whether a separation is still made between the production of the actual physical artwork and the emotion of the artist. Even if the artist must produce the work in order to become conscious of his emotion, this does not entail that it does not exist prior to him becoming conscious of it. Yet sometimes Collingwood suggests a very close connection between the artistic activity and the formation of the emotion:

[The painting] is produced by an activity which is somehow or other bound up with the development of that [aesthetic] experience itself. The two activities are not identical; he distinguishes them by the names 'painting' and 'seeing' respectively; but they are connected in such a way that, he assures us, each is conditional upon the other... There is no question of 'externalising' an inward experience which is complete in itself and by itself. There are two experiences, an inward or imaginative one called seeing and an outward or bodily one called painting, which in the painter's life are inseparable, and form one single indivisible experience, an experience which may be described as painting imaginatively. (Collingwood 1958: 304)

So Collingwood explicitly says that the act of painting and the aesthetic experience are *not* identical, but are so closely related that they are experienced *as if* they were indivisible. He is describing an interdependence between the two processes, at least in certain cases. This suggests then that the act of expression enables the formation of the emotional state. We also see this view in Dewey, who explicitly rejects the

idea that what is expressed is just an externalisation of a pre-existing emotional state. Rather it is physical interaction with a resistant medium that allows the emotion to be formed:

The thing expressed is wrung from the producer by the pressure exercised by objective things upon the natural impulses and tendencies - so far is expression from being the direct and immaculate issue of the latter. The third point follows. The act of expression that constitutes a work of art is a construction in time, not an instantaneous emission... It means that the expression of the self in and through a medium, constituting the work of art is itself a prolonged interaction of something issuing from the self with objective conditions, *a process in which both of them acquire a form and order they did not at first possess*. (Dewey 1971: 78-79, my emphasis)<sup>6</sup>

So we are developing a view in which the physical production of the work allows the emotion to develop. Yet it is unlikely that Collingwood would *fully* endorse this notion because of his commitment to the way that art can enable the sharing of emotions between people.

To explain: Collingwood emphasises that since to express an emotion is simply to understand it as a unique thing, the audience equally expresses an emotion of their own when they recognise it in the artist's work (e.g. 1958: 315). As in the case of the lecture, the students who listen with understanding in some sense make the thoughts

---

<sup>6</sup> In many ways Dewey anticipates the idea of extended cognition in his account of expression. For instance he claims: "The epidermis is only in the most superficial way an indication of where an organism ends and its environment begins" (Dewey 1971: 73).

of the lecturer their own (cf. Ridley 1997: 268). Hence the only real difference between the artist and the audience is that the artist got there first. As Collingwood says, “he is singular in his ability to take the initiative in expressing what all feel, and all can express” (Collingwood 1958: 119).

Since according to Collingwood, someone is able to express the very same state of emotion as the artist simply by understanding the work in the right way, the formation of that emotion must be separable from *actually* producing the object. This would be true even though experiencing the work may be the only way to access that emotion. So even if we suppose that the audience must imagine the construction of the artwork in order to properly access the expressive state the artist, this only reveals that the *idea* of producing the work is what is important, not its actual production.

Yet Collingwood seems unduly optimistic that the audience will discover the same expressive state in the work as the artist, just so long as the artist is skilful enough in externalising his expressions (1958: 149-150). As long as it is a matter of reconstructing the experience of the artist from our impressions of the physical work, there is no guarantee that others will perceive the same thing as the artist or each other. Every imaginative reconstruction of the work is in some sense a re-interpretation. And there will be inevitable distortions that accompany the filtering of a work through another’s experiential background. The advantage of identifying the artwork with the physical object is that at least there is some definite thing upon which our interpretations can converge.

Moreover, artworks like improvised jazz performances are much more tied to their actual singular occurrence than other notational art forms like poetry or scored music. The improviser doesn't get a second chance to replace what he has already played with something more appropriate. Rather, the whole performance may be viewed as the struggle of the musician to articulate his state, forming and reacting against the unalterable characteristics of that very event. This is what makes the performance *his expression* in a way that the rest of us can only appreciate rather than replicate. Ultimately it is the deep individuality of this kind of expression that really undermines any claim to idealism, at least so far as these kinds of artworks are concerned. This is because what guarantees that a work captures an individual emotional state is the unique physicality of its expression; the fact that it is embodied in a particular physical event by a particular person in a particular set of circumstances.

The main reason Croce and Collingwood come across as idealists is because their concern is to say the real work of art is a kind of mental state or process. Under most interpretations of mind this would automatically imply that the public object could not be the real work of art, and so conflict with our basic intuitions about what artworks are. Yet I think we can rescue as well as strengthen *some* of the insights of their position when we appeal to the theory of extended cognition. As we shall see, when physical objects become integrated into cognitive processes in the right way, they should be considered parts of the mental state. Hence it is possible that if physically producing an artwork is integrated with a cognitive process, then that artwork may be a physical part of the artist's mental state. As a result, we can hold onto our intuitions that physical objects and their actual history of production are

what constitute artworks, whilst still embracing the possibility that artists are interested in expressing their emotional states.

Note that we still cannot embrace some of Collingwood's statements such as that artworks may exist solely in someone's *head*. In general, neither Collingwood nor Croce shows sensitivity to the possibility that mental states may not be entirely in someone's head. Yet as I argue in chapter seven, the extended cognition account allows the possibility of sharing an emotional state as closely as Collingwood ever envisaged. By locating mental states partially in the environment, we can bypass the need for any reproduction or replication of mental states, but instead allow people to collaboratively generate and thus physically share mental states.

### **Extended Cognition**

In order to explain how extended cognition applies to jazz improvisation it is necessary to provide some background on the theory. Its *locus classicus* is Andy Clark's and David Chalmers' influential paper 'The Extended Mind' (1998) (though they are indebted to externalist theories of mind generally). Their basic idea is that we often incorporate external objects into our cognitive processes, the active manipulation of which enables the completion of cognitive tasks such as recognition or search. So for example, if I do some long division step by step on a piece of paper, or if I rearrange Scrabble tiles in order to find a word to play, I am exploiting these information rich features of the environment to support and supplement my cognitive processes. Hence they argue that these kinds of action are 'epistemically active' in that they alter the world for the purpose of completing some cognitive task rather than simply to enact some pre-conceived change that the subject intends.

It seems that many of our behaviours can be interpreted in this light where the environmental resource is reliable, easily available and the subject automatically endorses the information it provides. A key part of arguments for extended cognition (and externalist views of mind generally) then involve a supervenience claim. If two identical cognitive processes or states internal to the body can nevertheless have different content due to differences external to the body, then the constitution of the mental state must partially supervene on those relevant aspects of the environment. So for instance, if the Scrabble tiles happened to have been arranged differently, I would have come up with a different word despite being in the same internal 'search' state.

Obviously, this difference would then have impacted upon my brain state when I recognised the word to play. The point is however that within a certain range, I would endorse *whatever* word the external process resulted in. My internal state, though an important source of control in *choosing* the word, is not doing all the relevant work in *finding* that word. Work is being done by objects outside my body, which in combination with my actions and brain state physically realise an extended cognitive system. As a final test we may ask whether if the same process were done purely in the head (say imaginatively moving scrabble tiles around), we would be happy to grant it cognitive status. If we would, then it seems to exclude the object as part of the cognitive process simply because it is located outside the body is nothing more than a conceptual prejudice.

Clark and Chalmers go on to argue that not only can processes like recognition be cognitively extended but also more definitively mental states such as beliefs. Here they use the example of Otto who has Alzheimer's disease and who uses a notebook to tell him the location of the museum. The notebook then functions as a part of Otto's long term memory, which coupled with Otto's disposition to consult the notebook when considering the location of the museum, forms his belief that 'the museum is on 53rd Street'.

We might worry that because Otto has to visually perceive the notebook in order to consciously hold his belief, it is quite unlike the direct and sub-personal activity of accessing such information from our biological memories, especially since the visual image has a distinctive phenomenology. However Clark and Chalmers argue that because the flow of information from notebook to brain is all part of the same cognitive system, it is not perceptual at all in the sense of "registering the impact of something outside the system" (Clark & Chalmers 1998: 20). Furthermore, the phenomenology of retrieving the information is not particularly relevant to its status as a belief. We could equally suppose that if a sound accompanied the introspective access of our biological memories, those memories would no less form a part of our beliefs.

This point is relevant when exploring the role of music in the extended cognitive processing of emotional states. However for now I would like to point out that an important aspect of extended cognition is that the extended mental state is not an internal representational copy of some external object that then mediates between subject and world in directing behaviour. Though the information in the notebook

typically impacts on Otto's brain, (though not always, see below) his belief is not constituted by an internal mental representation of the page in his notebook, or the disposition to form such a representation. Rather it is the active exploitation of the information the notebook provides (that happens to be visually received) with additional internal processes (such as subject's endorsement) that together complete a *system* realising the cognitive function.

So there is a general sense of representation that is distinct from the acquisition of internal representational states (cf. Wilson 2004: 183). In some cases, such as knowing how to get from one town to another, the knowledge of the route may not fully correspond to any internal representation. Rather the subject may simply have learned how to respond in the right way to reliable worldly features (such as signposts) and allow the world to be its own map. In this case we can still distinguish between someone who knows how to use the signs (for instance, they know when to look out for the sign that will point them in the right direction) and someone who just has a general ability to follow signs but has never taken that route before (cf. Clark 1997: 21-31). In general, having to mediate behaviour by fitting an internal map against the environment it represents is an expensive drain on cognitive resources. So there is an evolutionary reason to suppose that the brain is geared towards co-opting certain reliable aspects of the environment rather than developing a psychical distance characterised by codified representations. This is true of many bodily skills in which the direct manipulation of objects bypass the need for internal maps of those objects. And as I argue below, playing a musical instrument is a good example of such direct manipulation.

In eschewing internal maps, we should consider extended cognition as aligned with direct realist or disjunctivist theories of perception, in which internal ideas do not mediate between perceptual states and their objects. However, even if Otto's belief state *contains* a completely internal representational state, a visual perception, this does not entail that his belief state is also internal. Overall it is the *system* constituted by the perception of the notebook, the writing in the notebook, and the endorsement of the information it provides that constitutes his representational activity. Only this system as a whole functions to represent where the museum is. Another sort of system with only differences external to Otto, such as different words in his notebook would have a different representational content. Moreover, these differences would not necessarily have to impact on his internal brain state to do all the work that beliefs do. For instance, someone could ask Otto where the museum is, and Otto could simply show him the relevant page in his notebook without having to look at it himself saying, "*that* is where I think it is."

### **Jazz Improvisation as Extended Cognition**

Having briefly explained the theory of extended cognition we may now begin to see how it applies to the case of jazz improvisation. My overall strategy here is to show that it is possible for the improvising jazz musician to go through a process of emotional expression much as Croce and Collingwood described. Yet contrary to the standard account of expression theory, this act of expression is realised by the physical interaction between the musician, his instrument and the music. So forming the emotion by expressing it in music is a cognitively extended process. However, before we deal with emotional expression, it is necessary to first show how the musician can develop such a close relationship with the music. This is my first claim;

that playing the instrument cognitively extends the musician's creation of the music. Hence I begin with a universal claim that simply improvising, apart from any concerns about having a real emotion, is a form of extended cognition. It is this view for instance that music psychologist Eric Clarke seems to be endorsing when he says that "playing music is a concrete form of musical thinking, and the body is as much a part of finding out about music as it is a means for its actualisation" (Clarke 2002: 68).

If we consider the simple task of generating notes, it is clear that the musician's interaction with his instrument allows him to do this, not just in the uninteresting sense that the musician needs the instrument to produce notes at all, but also that the particular way in which the notes are formed are a matter of sensitively responding to the capacities and affordances of the physical object. It is not the case that the musician always has some music in mind (either as sound or in symbolic form) that he just enacts on the instrument. Moreover, even if the performer does have a rough idea of what sound he wants to produce, there are many features of that sound that go beyond his concept of it, yet which he would still endorse as the sound he wanted to play.

At every level of creative decisions the musician and his instrument form a single tightly coupled system. For example, certain qualitative features of the sound may just be a product of the way the instrument is set up. It might have a harsh or smooth tone for instance. So the instrument itself helps to decide the character of each note that the musician then endorses. The way the instrument is physically set up also constrains and encourages ways that notes can be strung together. I mentioned earlier

the way that some musical figures will fall more naturally under the musician's fingers. In addition, building up a vocabulary of musical figures is something done on the instrument by exploring its physical capacities. This development then allows the expansion of the musician's imaginative capacities in terms of what is possible to play.

In addition, within the immediate context of performance there is the sense in which the musician may only have a rough idea of what he wants to play, for instance a fast upwards sequence of notes, but it is the interaction of his fingers with the keys that ultimately determine what exact notes are played. Thus when completing the cognitive task of choosing what exact notes to play, the instrument is part of an extended loop between the musician's brain, the muscles in his hands or lips and the keys of the instrument. As in the Scrabble example, within a certain range, the musician automatically endorses *whatever* the external physical processes result in.

There is then a further level at which the musician's conscious interaction with the *music* determines what he will play next. Similar to the instrument, the music itself will present certain limitations and affordances to the musician in terms of what sounds could follow in a musically coherent or meaningful way. By listening to what has happened so far the musician develops an idea of the music's momentum. A good way to conceive of this momentum is as an experience of the pattern of the music, which entails a limited range of ways in which this pattern can be completed. Thus music can gain momentum in terms of harmonic tension that can be resolved by returning to the home key, or thematic continuity and development. Also in terms of maintaining tempo, rhythmic pulse, and note durations, as well as developing

general dynamic shapes like building to a climax or fading away. We should note that the musician must *perceive* these qualities in order to react to them, and their perception relies on certain standards he has acquired concerning harmonic or rhythmic development. Yet at the same time the musician may not have *planned* any of these qualities. They may simply be the unexpected consequences of other musical decisions he has made.

Many features of the *expressive* character of the music can equally be unplanned, and can generate their own form of momentum, in terms of maintaining an emotional character, or developing it to completion or exhaustion. As we have seen in chapter three, the expressive character of the music is conveyed by features at every level of detail, from the intensity with which individual notes are attacked, to the smoothness of the timbre, the rhythmic pulse, harmonic tension as well as the large scale formal structure. It is thus highly unlikely that the musician can anticipate and control all of these potential sources of expressive effect. In many cases he can only react to the precise expressive character the music suggests. Yet he can still intelligibly take responsibility for the exact expressive qualities of the sound. Though the music possesses a wide variety of subtle expressive qualities that the musician did not plan or predict; he can, like Otto and his notebook, still point to the music and endorse the expression of *that*.

So overall, these features of the music form an information rich environment that the musician relies on when deciding what will happen next. In this way the music and musician form a cognitively extended system where the music presents certain information about what sounds are available which combines with the musician's

endorsement of these possibilities to enable and enhance the cognitive task of generating musical sounds. The musician's endorsement of these possibilities relies on his musical *intentions*. So a dialogue is set up between his intentions and the musical momentum. The key difference between the two is that a musical intention is an idea that the musician has about how the music *should* sound in the future. The musical momentum in contrast, is an idea about how the music *will* sound in the future so long as it is not disrupted either by mistakes or contrary intentions. Also, where musical momentum is something actually heard and anticipated, musical intentions are imaginative and creative. So a central component in the experience of performing music concerns the relative alignment of musical intentions with musical momentum. In the usual case the performer experiences a tension between his idea of what the music should sound like (perhaps determined by some mood he is trying to capture, or some structural concept he has) and what possibilities the sound itself seems to suggest or allow. A musician can also choose to violate what he regards as the musical momentum for expressive effect.

All this, by the way, is equally applicable to the relation between a singer and his voice. On traditional internalist terms we would say that the mechanisms of the voice are external to the singer's mind. Yet the singer must still interact with their vocal apparatus to decide what music is created. So this highlights that there are two aspects to extended cognition: the interaction with the mechanical means of performance; one's body and the musical instrument, and then interaction with the sound produced. In general, the mechanical means of performance are more central to my universal claim of extended cognition. It is conceivable for instance that a musician uses his instrument to decide what to play yet ignore what sounds are

actually produced. In this case the musician does not care about the momentum suggested by the music. Meanwhile the physical act of interacting with their instrument could generate a certain momentum. For instance, one may find oneself striking the keys more and more softly. Yet the musician could not do the reverse and ignore the affordances of the instrument whilst reacting to the musical momentum. Not if he actually wants to produce the sounds that the music suggests.

Altogether then my claim that playing the instrument cognitively extends the musician's creation of the music works at three levels, some or all of which are present in *all* improvised performances: The physical interaction between the musician and the instrument performs the task of generating the detailed notes. The interaction between his intentions and the perceived momentum of the music determines the larger scale shape and style of the music. Finally the interaction between the musician's intentions and perceived *emotional* momentum determines the expressive character of the music, affecting the sound at all levels of detail. All of this is analogous to performing long division step by step on a piece of paper, or rearranging Scrabble tiles in order to find a word to play.

### **Emotion Generation as Extended Cognition**

Having justified my claim that creating the music is cognitively extended, we can now turn to my second claim; that the musician may use the music to cognitively extend the formation of his emotions. My explanation of exactly how this is done is mostly reliant on the accounts of expression in music that I have provided in previous chapters. The basic point is that an improvised performance can help generate an emotional state because the music can fulfill the same role for the

musician as the primary expression of emotion (such as crying or punching the air). In the second chapter I argued that what makes these primary forms of expression expressive is that they contribute towards the bodily pattern and thus the feeling of an emotional state. The way it feels to cry or shout or jump for joy is a distinctive part of what it feels like to be sad, angry or joyful. In the same way, it feels like something to produce sounds on an instrument. To strike the keys gently or aggressively directly arouses the corresponding bodily feeling.

More importantly however, the patterns of the music are mapped by the simulative capacity of the musician's brain as patterns of emotional feeling, just as in the case of the listener I described in chapters three and four. Since the simulative capacity is used to recognise the emotions of others we might suspect that the emotion expressed is not the real state of the musician. But here the relation between simulated emotions and direct bodily emotions should be made clearer. Both processes equally result in a neural map of bodily changes in the brain, which in turn generates the phenomenal experience of the emotion. In chapter one I described how simulation may anticipate actual bodily changes in a personal emotional state, so the mere fact that the simulative capacity is being employed is no barrier to the authenticity of the emotion. Moreover, the neural pattern of bodily changes on its own is *insufficient* to distinguish between a real emotion, an empathic reaction, the recognition of emotions in music or the extended cognition of emotions in music. The wider context in which the bodily pattern is generated is *vital* for determining which of these states the subject is in.

The key identifying condition is not whether the bodily pattern tracks real or simulated bodily changes, but whether the system as a whole represents a real situation of personal significance for the subject (a core relational theme), the expressed state of another or simply a fictional expression of emotion. In the case of the improvising musician, there is no reason that the feelings expressed by the music cannot track the real situation of the musician. The music could even be both the object of an emotional state (for instance how well one is playing), and self-reflectively the emotional response to that object. This is similar to having emotions about one's emotion, as in the case of panic attacks that I mentioned in chapter one. Yet there is no reason why the musician cannot also be responding to some other aspect of the environment, such as the audience reaction, or any thought or imaginative idea he might have at the same time.

A second condition concerns the *control* that the subject has over the emotion generating process, which determines whether the state is partly cognitive or purely perceptual. Since the musician is responsible for producing and endorsing the emotional content of the music, this entails that the musician is cognitively deliberating his emotional reaction, in the same way as deliberately engaging in expressive actions allows the cognition of emotions (see chapter two). The sense of control or agency over the emotional content of the music also encourages a conscious identification with its content, unlike the case of simply listening to expressive music, where the listener must submit to its flow.

So what of the actual bodily changes of the musician in this case? In the first place, the bodily changes of the musician help to generate the intentions of the musician to

appropriately match the way they feel in the music. The musician's sensitivity to the expressive qualities of the music can also arouse his bodily changes in turn, as I described at the beginning of the chapter. Thus a dialogue can be set up between the musical momentum in emotional terms, and the inner bodily changes of the musician. So consider how the musician starts his improvisation: He could just wait beside his instrument and contemplate his emotional feelings before launching into a solo that he feels reflects these feelings, or as seems more likely, he will just start playing in a very impulsive way and then allow his sensitivity to the music develop into a sense of emotion. Either way, there is a constant feedback between the musician's inner changes and the expressive character he perceives in the music.

However, if this was all there was to it, we might complain that the inner bodily changes of the musician already constitute his emotional state. The music merely *influences* that state, or is influenced by the state in turn. So whilst we could admit a close relation between the two we need not say that the music partially *constitutes* the musician's emotion. Yet this criticism can be met on two levels: First of all, by tracking the music with his aroused bodily changes, and then reacting to maintain or develop the *unplanned* emotional character of the sound, the musician is endorsing *whatever* emotional content the music suggests. The music has direct control over what the content of the emotional state is. This is analogous to endorsing whatever it says in the notebook. So where we conceive of the emotion as a cognitive process of deciding what to feel, the music itself plays a vital role in enabling and enhancing this process.

Recall that a cognitive state is one in which the subject manipulates representations. In this case the representations take the form of emotional states. The musician simply cannot manipulate his emotion directly. He can only send a signal from the brain, to his hands, to the instrument, to the music, which he then hears and interprets as feelings. So he simply cannot cognise his emotion without actually manipulating the music. Suppose the musician had an identical twin that was a brain in a vat; so with exactly the same brain state, sending out exactly the same signals and receiving the exactly same felt feedback. Yet in this case the music is absent or changes in the music are independent of his instructions. This twin would be undergoing a different mental state. The twin subject would not be cognising his emotion. He would be having a passive emotional reaction, since he would not be in control of manipulating his emotional representations at all. His feelings would be occurring quite independently of his attempt to control them. So given that changes in the music alone entail a different sort of mental state, the music must be a physical part of the cognition of feeling.

However we may draw a distinction between deciding what emotion to have and the emotion itself, just as we may draw a distinction between performing an expressive action and the bodily changes that this action generates. So secondly it is important to note that the music is also immediately incorporated into the musician's emotion via his simulative capacity. That is, the musician enjoys the feelings expressed by the music *alongside* his internal bodily changes as the total combined emotional state. The patterns in the music effectively play the same role as his inner bodily changes in relating to the overall bodily pattern. Now, I argued in chapter one that we should equally include bodily changes as a constitutive part of the emotional system. This

was partly for reasons of wide content of the kind I articulated above, and partly because they do vital work in generating and maintaining the emotional content. The musical patterns equally do significant work in generating and maintaining the emotional content where the musician's brain simply endorses that content. So equally we should include the music as a constitutive part of the emotion. The music constitutes part of the *content* that is registered in forming the overall representational state.

In this case the music constitutes a physical *elaboration* of the musician's felt state. Phenomenally the musician partially experiences his emotional state *in musical form*. Of course, the qualities of the music are then registered by the musician's brain. Yet actual bodily changes are equally registered by the musician's brain. In either case the emotion is physically constituted by the system that incorporates both the neural pattern and the bodily or musical changes that sustain it. Then in the case I have been describing, there is interaction between all these parts. So we may say that the emotion is constituted by the entire system of bodily pattern in the brain, bodily changes, bodily actions, the activity of the instrument and the patterns in the music. The vehicle of brain, body, instrument and music then supports this system for generating the content of the emotion.

In general, one of the main motivations for including an external object as part of a cognitive process is that it should *enhance* that cognitive process, and not merely enact the results of some internal process. In this case, when the music is used to extend emotional cognition, it enhances the complexity of the emotion the musician undergoes. Just as expressive actions can generate patterns far more fine grained than

inner bodily changes, so musical expressions can generate patterns far more fine grained than expressive actions. The musical expression of emotion can have increased complexity, temporal range, subtlety and force. It can also have long-term structural content as a result of the repetition and development of motifs. As such, the music enhances the emotional response of the musician. Their emotional response to the world can be that much more structured, subtle and intense. Track one of the recording accompanying this thesis is intended to provide an example of this kind of musical performance.

### **Music as the Core of Emotion**

So far I have justified a weaker version of my claim that the music partially constitutes the emotional state of the performer. The core of the emotional state is still the internal bodily changes of the musician, which the music then elaborates. However there are some cases in which the music appears to *replace* the bodily feelings of the subject as the main focus of the emotional state. These are cases of radical absorption within the emotional momentum of the music.

Earlier I described a dialogue between the intentions of the musician and the momentum perceived in the music. In most cases, the intentions of the musician are not perfectly realised. So where the intentions and momentum are experienced in emotional terms, the musician may feel a tension between his bodily feeling and the music's expressive content. Yet it is possible for the musician's intentions to align with the emotional momentum of the music in such a way that the intentions of the musician effectively disappear. There is a continuum between what the music dictates in terms of its momentum and what the musician intends. At one extreme the

performer has a detailed idea of exactly what he wants the music to sound like and he simply executes it. At the other extreme, the performer seems as passive as a non-participating listener to the momentum of the music. So alignment may occur either because the musician finds that his emotional intentions are immediately realised in music or because he gives up trying to control the music and simply affirms the progression of the music as it happens. This kind of alignment is rare, yet there is some evidence that this does take place when we look at what the musicians themselves claim. Note that these descriptions are clearly impressionistic and suggestive rather than conclusive demonstrations of my claims. Yet they seem best explained by the theoretical apparatus I have suggested:

There is such a thing as letting the music take you, if you are willing, or if you are open enough. (David Baker (trombone) quoted in Berliner 1994: 219)

And:

I feel that I'm at my best when I can free myself completely from the effort of trying to put something out and feel more like I am the instrument being played... (Ira Sullivan (trumpet and saxophone) in Spitzer 1972: 14)

Similarly:

If a solo is going well, developing, I let it go on its own. Then I've reached that place where I've gotten out of my own way, and it's as if I'm standing back and watching the solo play itself. (Jim Hall (guitar) quoted in Berliner 1994: 798 ft. 38)

And again from saxophonist Ronnie Scott:

[W]hat seems to happen is that one becomes unconscious of playing, you know, it becomes as if something else has taken over and you're just an intermediary between whatever else and the instrument, and everything you try seems to come off, or at least, even if it doesn't come off it doesn't seem to matter very much, it's still a certain kind of feeling that you're aiming for - and when this happens - inspiration - duende - whatever you like to call it - a happy conjunction of conditions and events and middle attitudes - it will feel good. It will feel that 'I should be doing what I am' kind of thing. (quoted in Bailey 1992: 52)

Finally trombonist Curtis Fuller describes getting 'caught up' in the music saying:

I dance with it. That's my emotional state when I play. That's my feeling of expressing my total self in the music. (Curtis Fuller (trombone) quoted in Berliner: 217)

In many ways these experiences seem to represent one of the peaks of musical achievement. This is because whilst it is not so hard to play the note you intended to

play a fraction of a second earlier, especially if you have modest ambitions, it is extremely difficult to temporally align one's intentions with the moment of actually producing them, or to suspend one's critical judgement of what one is playing. When this does happen we may say that the perception of the music fully absorbs the attention of the musician. So the music takes the lead, or becomes the locus of control for how the emotion of the musician will progress (though of course, in performing that music, the musician overall is still in control)

Yet if the musician 'loses' himself in the music, such that he is no longer aware of his internal feelings, what allows us to say that the musician is feeling any emotion at all? Is the experience of 'watching the solo play itself' a form of self-alienation? Ciarán Benson, finding precedence in Dewey's descriptions of "a rhythm of surrender and reflection" in aesthetic activity (Dewey 1971: 144) describes these states of absorption as a 're-centering' of experience (Benson 2001: 187). She diagnoses the phenomenon as the absence of self-predicated thoughts. When self-analysis disappears and one is more absorbed in the music, the musician does not feel embodied in the same way. My idea of extended cognition justifies the idea that the normal boundary of body-world has actually changed. But the musician has not 'lost' himself, he has literally extended himself. It is not unreasonable then that this should result in a different sense of agency and embodiment.

Yet is it still an *emotional* state? It would seem so, since the case of musical absorption need be no different to any case of normal emotional absorption where we are so caught up in the feeling that we forget ourselves in the process. It is fairly atypical, since part of the purpose of emotions is to direct the attention of the subject

towards certain aspects of the world. Yet the emotion can continue to represent the relation between subject and world even when this meaning is not attended to. Moreover, there is no reason to suppose that if the progress of the music has up to this point been dictated by the goal of emotional expression, then the music suddenly loses its expressive status just because the attention of the musician is now more fully focused on the music.

Hence overall, it seems that the music just more fully constitutes and dominates the development of the musician's emotion. And although the attention of subject is not a necessary part of the emotional state, we may say that the musician's *experience* of the music more fully constitutes his *experience* of the emotion. At the same time, we should recognise that being absorbed to such an extent is a fairly exceptional state of being. So emotionally it is likely to be a state of great euphoria, where the dynamic relation with the world that it tracks is a sense of one's life flowing like music, of living musically.

Finally, it should be noted that such a radical identification between the music and the emotion of the performer need not be recognised as such by the listener. As described in chapter four, if they understand the music as having emotional content, the listener need only attach some persona imaginatively to that emotion, not necessarily the performer's. Of course, if the listener were to be watching the musician as he plays, it may be abundantly clear that the music produced is a primary expression of his emotion, given other behavioural cues such as facial expressions or bodily movement which would presumably align in an appropriate way with the emotional content of the music. However, it is by no means necessary that the

musician provide any of these other cues. Apart from producing the music, there may be no other indication of the musician's emotional state. In the following chapters, I explore what it is that really forces the listener to hear the music as constitutive of the emotional state of the performer right in front of them. That is, by a more profound sharing of attention to the music and finally, in terms of sharing in the production of that music.

## Chapter Six: Joint Attention to Music

In the previous chapter I described a maximal form of engagement that an individual may have with a piece of music. There the music served to partly constitute the emotion of the musician. In this chapter I return to a more minimal form of engagement, passive listening, in order to begin exploring our experiences of music at a social level. The key phenomenon that I analyse here is joint attention, which we may initially characterise as when two or more people are mutually aware that both are attending to some object in the environment. So I am interested in what ways jointly attending to music can affect the way that we perceive it, especially the music's expressive properties. In particular, I want to examine whether joint attention to music can cause a convergence of responses, or a mutual fixing of the music's expressive properties. Yet I am not just interested in whether we can agree in our perceptions so much as whether we can listen to the music *as a group*. That is, whether our perceptual activities can be so integrated as to be called 'socially extended' in any way (in the sense of extension that I elaborated in chapter five).

Whether jointly attending to music involves group perception in the sense I have indicated relies primarily upon the nature of joint attention itself. The philosophical and psychological literature on this subject is mainly focused on infant joint attention, because engaging in this behaviour is often regarded as a landmark in the infant's ability to understand other minds. It is claimed for instance (e.g. Tomasello 1999) that joint attention provides the foundation for all manner of cooperative activities involving mind reading, most notably the acquisition and use of language. Yet exactly what notion of other minds (or their own) infants grasp at this stage is

unclear. When adults jointly attend to something, they are able to appreciate that others have beliefs that are distinct from their own, and that those beliefs may not accurately represent reality. Adults are aware of the ways that joint attention could go wrong, and thus the level of mutual monitoring that is required to ensure its genuine occurrence. So the mature case should involve a more sophisticated form of interaction that is more applicable to the joint attention to music.

Yet I think it is worth examining how infant joint attention functions because it reveals just how *minimal* the requirements for joint attention are. In particular, it is implausible that infants have much understanding of the distinction between their own intentions and the intentions of others. Instead infant joint attention seems to rely on a more automatic imitation of the adult's behaviour given a context in which they are 'tuned in' to the adult's intentions. Then as they get older, infants learn to check and direct the adult's behaviour in ways that gradually take on the sophistication of the mature case.

Despite these developments however, I argue that mature joint attention still relies on essentially the same stance of mutuality that infants are able to establish. Mature joint attention is merely a refinement of this stance. In particular, I argue that infants form 'plural subjects' of attention with their caregivers, where the task of looking is structured interdependently with the adult. It is forming this plural subject that allows people to mutually determine the character of objects in the world, to engage in more complex cooperative actions, and ultimately to share mental states such as intentions and emotions. Hence I argue that adults equally form plural subjects of attention, they are just more sensitive to what kinds of conditions can upset that state.

### **Infant Joint Attention**

Infant joint attention develops in several stages during the first 18 months of life. At 9 months, the infant simply follows the gaze of the adult towards objects in their common environment. By 12 months however, the infant is leading the focus of attention, either by vocalising or pointing towards the object of interest. In addition, genuine joint attention is signaled not only by the infant's ability to grasp that the adult's eyes are 'pointing' at something, but that they also return eye contact with the adult, seemingly to confirm that both are indeed attending to the same object. This confirmatory behaviour is important because it shows that joint attention is a truly shared activity, as opposed to just two people looking at the same thing at the same time. The confirmatory behaviour is also significant to the further development of joint attention. Between 12 and 16 months the infant will check the direction of the adult's gaze whilst they are pointing, and by 18 months they will check before they start to point at all (Franco 2005: 143). Using this strategy will help to ensure that the adult's attention is successfully directed.

The change in behaviour at 9 months can partly be explained by the infant's new ability to accurately follow the adult's line of sight. Prior to this age, infants will follow changes in the adult's direction of gaze but much less reliably. It may be that initially the infant simply mimics the head movements of the adult, at which point they become attracted to some object lying near to them. Joint attention could then develop because infants learn that generally following the head movements of adults will lead them to interesting sights. The ability to focus on eye movements within head movements then gets gradually refined into a geometric mechanism enabling

the infant to infer the location of objects from the adult's eyes more exactly (Butterworth 1995).

However, this geometric mechanism is insufficient to fully explain joint attentional behaviour. Firstly, infants gesture at objects for seemingly no other reason than to share the adult's attention. That is, they appear satisfied when the adult's attention has been focused on the object of interest. This indicates that the infant's behaviour is socially motivated rather than motivated by the object *per se*, at least by 12 months. Secondly, as Michael Tomasello points out, a whole range of social behaviours appear between the ages of 9 and 12 months including "joint engagement, gaze following, point following, imitation of instrumental acts, imitation of arbitrary acts, reaction to social obstacles, use of imperative gestures, and use of declarative gestures (including proximal gestures such as 'show' and distal gestures such as 'point')" (Tomasello 1999: 63). As such Tomasello argues that a more general understanding in the infant must be linking all these behaviours together.

Since the goal of the infant's behaviour appears to be simply to attend with the adult, Tomasello argues that the infant must recognise the adult's *intention* to attend with the infant. Equally, when the infant comes to initiate joint attentional behaviour, they should recognise that the adult engages in this action because they recognise the cooperative intent of the infant. So the infant must try to reveal this intention to the adult. Thus Tomasello claims that the infant grasps the thought, "You intend for [me to share attention to (X)]" (Tomasello, 1999: 102). If this were the case, infant joint attention would involve a structure of intentions similar to that required by Grice's (1957) account of meaning, where part of the reason for interpreting the content of

the intention in a certain way (in this case to jointly attend) is recognising the intention for it to be understood. Yet some theorists (e.g. Eilan 2005, Roessler 2005) have objected to this account on the grounds that it is implausible that infants have intentions with such a complex internal structure. Infants do not seem to possess the necessary conceptual sophistication that would allow them to recognise the adult's intention as an explicit basis for forming their own intention to engage in joint attentional behaviour.

Clearly the infant is interested in making sure that they properly target what the adult is doing, otherwise they would not turn back to check the adult's gaze. This seems to involve understanding the *goal* of the adult as distinct from his behaviour. It is reasonable to suppose that infants can grasp goals, since we know that by 8 months infants can distinguish goals from intermediary actions, such as pushing away an obstacle in order to reach the desired object. Then at around 11 to 14 months, infants will imitate the goal of unusual adult behaviours that could be produced by various means and will look around in anticipation of the previously observed result (Tomasello 1999: 82). Between 14 and 18 months they also become sensitive to the difference between successes and failures of the adult's intention, revealed by their imitation of the successful action, rather than the failure (Tomasello 1999: 81-83). Yet what all this indicates is the infant's increasing ability to keep a goal in mind (the adult's or the infant's) as an imagined state of affairs. It does not show that infants have a *higher order* awareness of the existence of their own intentions, or the intentions of the adult, especially in so far as those can be directed towards each others' intentions.

Apart from reading the adult's goal, the infant must also have a goal of his own to imitate the adult's goal. Yet this goal may not be deliberately formulated. Rather we can appeal to the general disposition that infants have to imitate others. In chapter two I described infants' pervasive and innate imitative tendencies in order to explain our ability to recognise emotions. We can then point to a gradual transition between these direct forms of imitation and the imitation of the adult's attention. The first development is from simple mirroring to more extended protoconversations, where infants will take turns with adults in reciprocal<sup>1</sup> actions such as looking, touching and vocalising (Tomasello 1999: 59). These interactional routines develop because adults assume that infants are intentional agents and relate to them accordingly, even before they can be truly classed as such. They are reinforced with positive expressions of emotion every time the infant responds in the appropriate way. In addition, adults recognise the infant struggling with a response and so simplify the interaction, standardising it with repetition, directing the infant's attention onto a key aspects of the task and so reducing the number of variables that the infant has to deal with. Here the adult is not just tuned into the infant's attention but helps to structure their expectations, gradually extending their attention towards more sophisticated tasks (Bruner 1995: 6).

The interactional routine scaffolds the development of infants from one stage of social understanding to the next, gradually drawing the infant into more sophisticated social games. From the basic foundation of direct eye contact and dyadic interaction, the infant then moves onto triadic interactions like joint attention. It is still important for the adult to respond in the same reciprocal way, but now objects have become

---

<sup>1</sup> Reciprocal actions match the intensity and temporal pattern of the behaviour rather than the exact movement.

integrated into the interactional routine. So the infant and the adult have already tuned in to each others' actions and emotions during their dyadic interactions. The infant already has experiences of himself and the adult acting together so can therefore recognise this state, and given that it is an emotionally positive experience, be motivated to establish this state once again. Moreover, apart from providing basic information about looking behaviour (either 'at me' or not), direct eye contact establishes a primary sense of mutuality that is a constant from the earliest cases of dyadic interaction. So it is likely that the importance of occasional eye contact within bouts of joint attention is to establish or confirm this primary sense of mutuality.

Thus I think the simplest explanation for the infant's ability to target the attention of the adult may be just that their imitation gradually targets more subtle aspects of the adult's behaviour. In chapter two, I argued that infant imitation is most likely grounded in the mirror neuron system, which recreatively processes the actions of others from a first person perspective. This system was important in grounding a variety of mental simulation in which by mirroring another's behaviour (either actually or at a neural level) one could come to be in the mental state that drives their behaviour. In this way, infant imitation of attention could be an important precursor to more sophisticated empathic or mind reading projects. Certainly it is worth noting that one of the symptoms of autism is an incapacity or unwillingness to engage in joint attention behaviours.

Even in adult life we have a strong automatic tendency to follow other peoples' gazes when they suddenly look around. Hence infants may not even form much of a deliberate goal to act in conjunction with the adult. They may simply acquire goals

by two means; either by desiring it themselves, or by imitating the intentions of the adult. The infant may automatically go through a simulative mirroring process and then simply fail to inhibit acting on that simulation. Perhaps some mechanism for inhibiting simulation takes a few years to develop. Equally infants may be so prone to emotional contagion because they are less able to inhibit the arousal involved in emotion recognition.

Then to explain how infants can read the goal of the adult at all, we should also supplement this simulative activity with a generalised understanding of behaviour. We know that as well as having an innate tendency to imitate, infants also have an innate tendency to look at drawings that resemble faces rather than other patterns. This selective attention to and imitation of humans rather than other kinds of objects indicates that infants can immediately distinguish between the two and that they functionally identify humans as entities like themselves. It indicates what Andrew Meltzoff has called a ‘like me’ stance (Meltzoff 1996, 2002), a disposition to grasp bodily and behavioural states in a *general* way as applicable to both oneself and others. In effect infants understand the behaviour of others the same way they understand their own (without necessarily having to make some kind of analogical inference). So they can potentially imitate any aspect of the adult’s behaviour that they are sensitive to in their own case. Hence the idea here is that once the infant is generally able to distinguish goals within their own behaviour (such as focusing on one object rather than another), they should be able to distinguish the same in others.<sup>2</sup>

---

<sup>2</sup> Equally the process should also go in the other direction. Having formed imaginative representations of the goals of others, infants should also be able to form imaginative representations of their own goals (Asendorpf 2002).

Once the triadic structure of interaction is fully in place, the infant can then use it to learn how to approach the world generally. One of the key ways this is demonstrated is when infants socially reference, looking to adults when meeting unfamiliar objects or strangers in order to determine what kind of emotional reaction is appropriate. For instance, infants will avoid objects or people that their mothers show fear towards, not just whilst the mother is present but also later when they are on their own. Emotional contagion alone could not account for this permanence of effect, but something more like emotional conditioning towards the object. It is also possible that if the child is looking to the adult when approaching the new object, it may simply be for reassurance rather than to resolve any uncertainty about how to deal with it (Baldwin 1995: 136). This is confirmed by the fact that infants will look to their mothers more when their mothers display fear rather than pleasure, indicating that the child is generally feeling more threatened. However, infants show a high degree of specificity in how they correlate the adult reaction to a particular object or person. Hence the infant must also be led by the properties of the object rather than just a general feeling of unease (Baldwin 1995).

Peter Hobson makes the stronger claim that the *main* purpose of joint attention is to share emotional reactions towards objects. He claims that one of the reasons autistic children do not engage in joint attentional behaviour is because they are generally not responsive to emotion sharing in dyadic interactions (Hobson 2005: 190). In a similar vein, Johannes Roessler argues that the most significant aspect of this emotion sharing in reference to objects is that it is a form of predication. That is, objects come to acquire properties other than just being worthy of attention, such as

being funny or scary. The importance of this is that emotional predication can be right or wrong (Roessler 2005: 245-6). So the infant may have one emotional reaction to an object that the adult transforms, or ‘corrects’ with a different emotional reaction. This may be the earliest case of a sense of objectivity. Hence Roessler (following Werner and Kaplan 1963) claims that the importance of joint attention is in establishing the sense of objectivity.

So rather than joint attention enabling the infant to gain a sense of mutual awareness, they may in fact enable the infant to make a more accurate *distinction* between his reactions to an object and the way that objects ‘really are’. At the same time, his different reactions to objects can establish a more accurate sense of distance between himself and the adult.<sup>3</sup> So it seems the importance of joint attention is to teach the infant how to attend to the world. The way we attend to objects and people fundamentally shapes the way we experience them, in particular the kind of emotional reaction that they engender. As such joint attention allows the infant to begin participating in the cultural world, understanding the purpose and meaning that certain objects have to offer on top of whatever natural ways they can be manipulated. Moreover, when infants learn to manipulate the adult’s attention in return it signals their capacity to *negotiate* the meaning of their common environment.

In this way infant joint attention provides a model for how our perceptual activities can be interdependently structured. Adopting Margaret Gilbert’s (2000) terminology,

---

<sup>3</sup> This notion is very similar to Davidson’s theory of triangulation, whereby three forms of knowledge; knowledge of our own minds, knowledge of other minds, and knowledge of the world are all interdependent (e.g. Davidson 2001a, 2001b).

the infant and the adult form a 'plural subject' of attention. Gilbert's idea here is that the two try to integrate as much as possible the way that they approach the world. In this case, the adult and infant integrate their attention to objects, both in terms of when and what they attend to, as well as the sorts of properties of that object they are interested in. Thus we shouldn't describe joint attention as the infant simply looking at an object plus recognising the adult do the same, but from the start *sharing the task of looking* with the adult. Equally, when infants initiate joint attention they are interested in first setting up the framework of attending together. For the infant, this behaviour is unlikely to be mediated by a cognitive model of how the adult's and infant's intentions are related to each other. Rather it is directly and unself-consciously entered into. As much as possible, infants seek to make their activities interdependent with the adult's, to form plural subjects of attention. And as I argue in the following section, this is a form of socially extended cognition.

### **Mature Joint Attention**

The joint attention that an infant has with an adult is not like that between equals. It would be better to say that the infant is immersed in the attention of the adult. This is because the infant trusts the adult more than is warranted to be responsive to their intentions, at least until their checking behaviour develops. Secondly they defer their emotional reactions to those of the adult. Thirdly they are not yet able to understand that adults may have false beliefs about the situation or the infant, and that they in turn could have false beliefs about the adult. Finally it is unlikely that infants have the same self-reflective understanding of attention that adults do. Attention for adults is a conscious relation between a self-aware subject and a world containing distinct and meaningful objects. The adult is able to consciously shift his attention to

different objects within the environment or to different aspects of those objects, as well as bring to bear various background knowledge that will enable them to conceptualise the object and recognise its functional affordances. So the adult can understand that even if an object lies within another's perceptual field, the other will not necessarily see the same object as they do, or focus on the same aspect of it, or understand it in the same way unless specifically directed to do so.

Yet despite these differences, I argue that adults still rely on essentially the same stance of mutuality, the plural subject, that infants are able to establish. They are just more sensitive to what kinds of conditions can upset that state. In particular, jointly attending to something is still a matter of having, or setting up a *framework* for attending to the environment, where the task of seeing is structured interdependently with the other person (cf. Currie 2007). The nature of this framework in the mature case is just more complicated. For instance, adults can manipulate each others' attention in various ways, such as which sensory organ is to be brought to bear on the world, or the degree of urgency involved. More significantly, mature joint attention is infused with a sense of the *normality* of the situation, where objects have conventional ways in which they are attended to, and certain conventional reactions are implied. For instance we normally look at clocks to check the time, and plan our actions accordingly. In this way normality helps to fix the content of what is jointly attended to.

On the other hand, if the object of joint attention is more unique, then a more sophisticated process of mutual alignment towards the object is required. As well as directing your gaze, I may verbally describe various aspects of the object to you, for

example saying “look at the way the light reflects from that building” perhaps implying an emotional response as well. By reciprocal behaviour on your part, we can then build up a *specific* framework for that object, the mutual negotiation of which gradually fixes the nature of the object, and confirms that we are jointly attending to the same thing.

In this respect the framework is similar to the notion of a shared cognitive environment that Dan Sperber and Deirdre Wilson (1993) appeal to as the background for successfully disambiguating the content of communicative intentions. It involves gradually building up a picture about what kinds of information are mutually available (and mutually known to be mutually available) to both of us. This can include not just ordinary facts about the physical world, but also obvious cases of natural meaning such as black clouds indicating the coming of rain, or most relevant to my interests, that certain behaviours indicate certain emotional states. Of course in most cases there will be a degree of uncertainty as to what facts we can reasonably assume other people to possess. Yet on a day to day scale, as we continue to interact with other people, we mutually gain all kinds of information about what facts are available to each other. Moreover being able to identify someone as belonging to a particular social group, such as a culture or a profession, will immediately entail all kinds of information about what kinds of facts are shared between normal members of these groups. For example, amongst musicians within the Western Classical tradition, certain generic styles of music will immediately signify certain generic emotional states. Hence even two musicians who have never met before should find it fairly easy to communicate ideas about what emotion a piece of music expresses without having first to explain the basis for those interpretations. Combined with

continuing interaction, their ability to recognise each other's musically expressive intentions can become extremely fluent.

So the shared framework is a way for the content of our attentional states to be mutually fixed. Yet establishing a framework is also a matter of mutually structuring the *activity* of attending to the world. First of all, we can see that as in the infant case, some form of mutual monitoring is required that causally influences the way we attend to the object. Though I may be motivated to attend to the object because it is an interesting stimulus; that I am attending to it at this particular time, for this duration, with regard to this or that particular feature (that may be socially meaningful) is because you are attending to it as well. For example when I direct your attention towards the clock, I may be interested in the time because I want to catch the train, but also because I want to synchronise my behaviour or emotional feeling with yours, or to help explain to you the reason why I am rushing around.

Monitoring the other can involve periodically exchanging eye contact or any other form of mutually reciprocal behaviour such as verbal exchange or touch, yet it can also be less overt. My awareness of the attention of the other does not have to be a particularly complex part of the activity. In normal conditions, I can track another's attention simply by being disposed to notice when he is no longer attending with me, for instance if he walks away or changes the subject. Hence joint attention is characterised by an ongoing *preparedness* to alter the way I attend to something should you direct me to it, as well as being self-consciously aware of the publicly available aspects of my own attention as they have the potential to lead your attentional focus. This preparedness and openness should be mutual.

In general different cases of joint attention will require different degrees of monitoring. The notions of preparedness and openness can also help us to identify a minimal kind of mutual monitoring. Suppose for instance that I happen to be walking along a road tunnel at the same time as a complete stranger when suddenly a burst water main sends a wall of water crashing towards us. We could both assume that we both see the water (and that we both know that we both see the water) without ever having to check each other's gaze. Suppose we then both turn and flee, still without directly looking at each other. Given no reason to think one's reaction is unavailable to the other (i.e. they aren't blind) we could still expect a degree of mutual sensitivity. For instance, if the other person looked back and then started to slow down, I would most likely notice this and assume that the danger had passed. Similarly, if I looked back and slowed down I would anticipate that he would notice this in turn.

In general, when confronting very sudden or salient events, particularly where they impact on everybody around, it is possible to assume joint attention without having to engage in overt monitoring.<sup>4</sup> The pervasive sense of normality also means that the joint attentional attitude is potentially *always on* to a minimal extent when we are out in public. It is when the situation is more unusual or subtle or of a more individual impact (for instance if one finds money in the street) that we need to check each other's reactions more carefully. We also expect certain normal reactions to situations, which when violated (for instance, if the stranger runs *towards* the water) motivates closer monitoring of joint attention. Yet in any case there is still a

---

<sup>4</sup> Due to an undeveloped sense of normality, it is unlikely that infants could jointly attend in this minimal manner. Hence the sophistication of the adult case both allows more complex mutual alignments (i.e. conversation about the object) as well as less overt monitoring.

preparedness to adjust our responses to the situation according to the responses of others and an assumption that others can respond in kind to us.

The point of all of this is that since this monitoring behaviour is embedded within the activity of attending to world, we should say that it is a *constituent* of the task of attending to the object rather than an additional factor. The mutual awareness involved in joint attention entails basic differences in our perceptual attitudes in terms of the preparedness and openness that I have described. So similar to the infant case, it is not the case that when jointly attending to *x* that we each have (i) perception of *x* plus (ii) mutual awareness of (i), but rather a perceptual state which is *of both of us perceiving x*.

This is also the view of John Campbell who argues that,

[J]oint attention is a primitive phenomenon of consciousness. Just as the object you see can be a constituent of your experience, so too it can be a constituent of your experience that the other person is, with you, jointly attending to the object. (Campbell 2002)

One consequence of this view is that it would not be possible for me to have this perceptual state if the other person were not in fact attending to the object. So for example suppose that I thought we were both sitting by the lake watching the ducks when in fact you slipped away some time ago. It might be argued here that if the visual experience of attending to the ducks was the same whether you were there or not, then your presence could not partly constitute the perceptual state. However

Campbell argues that we should take a disjunctivist perspective on this case, in the same way as disjunctivist accounts of perception distinguish veridical and hallucinatory perceptions. On this account, when veridically perceiving, the actual object is necessary to determine the content of the perceptual state. Hallucinations in contrast, make the perceiver *believe* that an object is part of the content of the state when in fact it is not. They are two fundamentally different sorts of state. In the same way when I falsely believed that we were jointly attending to the ducks, I took you to be part of the content of my perceptual state when in fact you were not. I was simply not having an experience of joint attention.<sup>5</sup>

It is worth clarifying this point about perceptual experience because Campbell's point is slightly ambiguous. We are assuming that I don't literally see the other person, so he is not part of my perceptual experience in this way, he is not one of its objects. Rather he is part of the subject, the person to whom the perceptual experience is happening. But again this is an odd thought. That person is presumably having a perceptual experience of his own, which may be qualitatively similar to mine, but surely numerically distinct. Hence how could he be part of the *subject* of *my* experience?

I think we can best characterise it like this: I am having a perceptual experience, and my experience represents a relation between the subject and the object. So part of it is an awareness of myself as relating to a thing in the world. However, in this case, my awareness of the subject is also an awareness of the person sitting next to me. I feel like I am one point on the base of a triangle that converges on the object we both

---

<sup>5</sup> Campbell also argues that only this view of joint attention can provide the common knowledge necessary to engage rationally in cooperative actions.

perceive. So the *directedness* of my representational state is different, in that it is directed from a wider spatial point than normal. The other person helps to shape the framework of my visual perception, literally the way I take myself to be approaching the object. This is not purely a spatial experience however. I also have ideas about what kinds of information are available to the other person, and to some extent what kind of person he is, the kind of information I could equally self-consciously think about myself. So my thoughts about the other person, combined with my awareness of what facts about me are available to him, are a *filter* through which the visual perception is interpreted. Thus I have a different experience of the object when I see it with others.

Then according to Campbell's general disjunctivist views, I could not be having this experience unless the other person was actually there, in space, next to me. We might concede this much at least. Yet I have only really referred to thoughts about the other person. This is all quite compatible with a thoroughly internalist conception of the whole state. And as long as we keep talking about other people being constituents of my *experiences*, the internalist can continue to talk about the *thoughts* of other people, not the people themselves. So instead I think we should take the extended mind approach, and talk about cognitive processes in non-phenomenal and functional terms.

So let us assume that my visual experience supervenes on my brain state. All information that I experience must be first processed by my brain. This is what allows the internalist to make his claims. However, we can instead say that the *task of seeing* the object is something mediated by the other person's task of seeing. So to

the extent that I am monitoring the other person, the direction of his gaze, his responses to the object and to me and so on, my task of seeing is sensitive to him, and affected by him. I look at something because he looks at it, I identify certain features because he does. I have certain emotional responses to it because he does and in all these respects he is equally affected by me. And if he were not there at all, (if I was hallucinating, or he had slipped away unnoticed) I wouldn't really be directing my seeing via his seeing. I wouldn't really be having what is called a joint attentional state. I would be having something else, a regular attentional state, though one that might seem a bit like a joint attentional one. So just as infant joint attention is not mediated by some cognitive model of their interactions with the adult, but is constituted by those interactions themselves, so mature joint attention is equally constituted by the actual system of interactions, the plural subject framework.

The *extent* of monitoring is an important condition here. It allows the tasks of seeing to be more or less integrated. This factor will be important when we compare the silent joint attention to music to the noisy joint attention to music, since depending on the situation, the mere fact of jointly attending (with no additional overt monitoring) can have various affects on the perceptual state. But first let us emphasise that the intrinsic content of my visual state is not partly constituted by the other person. Rather my task of seeing is partly constituted by the other person, what we might call the structure of the visual state. Sometimes philosophers talk as if the task of seeing were equally the visual state, and since perception is a dynamic process, continually tracking features of the world, filling out details, solving the binding problem, focusing, recognising and interpreting and so on, it is often fair to say that this whole process is the visual state. Given this proviso then we could say

the other person partially constitutes the visual state. Yet when we talk about visual *experiences*, the distinction becomes clearer. One could not be engaged in the same task of seeing if one were a brain in a vat. But perhaps one could have the same experience if one were a brain in a vat. (We could then make a content externalist argument about whether or not this was the case (e.g. Putnam 1981) but I do not have the space to go into this issue here.)

Overall, my claim is that joint attention involves establishing a plural subject of attention, in which a framework for perceiving the world is generated. This is a minimal case of socially extended cognition since the actual interactions involved define that state. Furthermore, just like the infant case, the main purpose of joint attention is to objectively fix the content of attention. So when jointly attending to the clock, there is one clock that constitutes our two perceptual states and in this way our perceptual states converge. At the same time however, various details of our perceptual experiences may differ, such as the specific orientation towards the clock, or whether one of us attends to the second hand and the other does not. Joint attention does not require the matching of perceptual experiences beyond a basic level. Rather we can both be attending to the same thing whilst ‘filling out’ the details of that experience in different ways. Yet what matters is that we are able to accommodate these differences within the shared framework of the joint attentional experience. Where these differences are expressed we can recognise that they apply to the same thing. This means that we can jointly keep track of the shared object even if whilst one person is looking at the clock, the other is monitoring my gaze or vice versa. At the same time, joint attention can imbue the object with additional social meaning, as something that signifies our relation to each other, as something

that affords a cooperative action or as something that arouses a social emotion in us. It is with this in mind that we can now turn to the case of jointly attending to music.

### **Joint Attention to Music - The Silent Case**

Now that I have shown that joint attention is a socially extended state we can now begin to explore the consequences of this on our social experiences of music. To start with I will look at cases of silent joint attention, which we may characterise with the following kind of scenario:

In a prestigious city concert hall, an audience of several thousand people chatter quietly amongst themselves. After a while the musicians of the orchestra make a ramshackle entrance, take their seats and begin to tune up. Finally the conductor strides in and the audience applauds him to his podium. The conductor ritualistically shakes the hand of the leader of the orchestra and the musicians lift their instruments in preparation. The audience gradually becomes completely silent and then at a moment of exactly his own choosing the conductor begins the performance. The disorder of life suddenly coalesces into perfect order. The actions of the orchestra are exquisitely balanced, beyond most other social activities. The audience too has become ordered, uniformly focused on the movements of the conductor and orchestra. During the performance they remain silent and immobile. They hardly even look at each other until the music is finished.

If when participating in audience listening of this kind, you take a mental step back from the music and instead think about those around you, you realise the enormous concentration of attention focused on the stage. Concert halls of this kind are

designed to direct all eyes onto the conductor (or soloist) at the centre. The audience is still visible, yet any kind of audience noise is considered a nuisance. So the listener is encouraged to ignore the other listeners and concentrate on the music as much as possible. Yet this is certainly a case of joint attention, since the listeners are aware of their mutual participation in a listening experience. However the minimal degree of mutual monitoring of each other's reactions throughout the performance will limit the extent that the aural perception of the music is integrated.

Yet even without the freedom to openly comment on the music during its performance, the audience are at least directed towards the same event and may nonetheless be having extremely similar experiences. As in other cases of joint attention there will be normal ways in which the music is perceived. At the most basic level the music is to be treated as a piece of music, performed largely for its own sake rather than for any other instrumental purpose. In addition, within the classical repertoire usually performed in concert halls, the audience can commonly expect that there will be a theme to follow, that the music will have a familiar large-scale form and that it will be emotionally expressive.

The problem of course is that pieces of music are multi-faceted objects, capable of sustaining all manner of different perspectives beyond the most superficial level. Throughout chapter four I outlined the various ways in which listeners may differ in their responses. One of the most significant factors was the different levels of expertise that listeners can bring to bear on their listening experiences. Some listeners will be familiar with the cultural context in which the piece was composed, some will have heard the same piece before and some will have greater theoretical

knowledge about how the music is produced. The result of this expertise is that some listeners will literally be able to hear elements of the music that others cannot.

Yet in chapter four I also described how given that the music relies on its natural resemblance to emotions for its expressive character, we should not expect radical differences in what emotions listeners recognise. The different responses that listeners have should for the most part be intelligible to each other and grounded in real features of the music. I also described above how the framework for the perception of objects that joint attention generates is one that can accommodate different ways in which different people fill out the details of the experience. So even if one cannot eliminate the different levels of expertise, joint attention still entails that there is a common target at which the listeners all aim. The mutual fixing of this target involves a willingness to allow one's perspective to be influenced by others and an inclination to have one's reaction cohere with the reaction of the group. So just as in ordinary cases of joint attention we should not separate out the listening to the music into (i) perception of the music plus (ii) mutual awareness of (i), but rather a perceptual state *of us all listening to the music*.

Though there is a basic alignment in our perceptual activities, exactly how much the *content* of our listening experiences can be coordinated depends on how much a listener can be made aware of the reactions of other listeners. The problem of course is that within a silent joint attentional situation, there simply isn't the opportunity to allow different responses to be expressed and thus negotiated. Yet even in the silent case of joint attention, there is still a sense in which the awareness of being part of a situation like this will impress itself in common ways upon the individual listeners

within an audience. When it comes to the emotional content of the music it is important to note that silent joint attention implies a tacit *acceptance* of that emotional content. Sometimes this can be a particularly uncomfortable experience. For instance, imagine listening to a highly sentimental and romantic piece of classical music in a room alone with your boss. Since the emotional content of the music must be commonly assumed to be equally obvious to both listeners, it generates a palpable sense of an emotional ‘atmosphere’, which given your background knowledge about each other may or may not seem appropriate. Thus due to its mutual availability, that emotional content of the music now gains social dimensions.<sup>6</sup>

Furthermore sitting in a huge concert hall surrounded by thousands of other people immediately lends an intensity to any event that occurs in that space. Apart from the vast range of sonic forces that can be achieved, part of the drama inherent in live performance is to witness an extraordinary act of human skill, comparable to watching a tightrope act with no safety net. It is unlikely that the same atmosphere could be achieved if there was only one listener present. The greater sense of tension is dependent on the increased concentration of attention that the mass audience generates. Every event in that space has massively increased social consequences.

Compare this to recordings, in which the signs of human presence are diminished. Because the music is no longer a *social event* in the same way, we may even claim

---

<sup>6</sup> It would be worth empirically testing the different emotional responses of listeners when made to listen to music whilst directly facing each other as opposed to facing the source of the sound. I would predict that an inhibition as well as intensification of response might result depending on the listener’s mutual familiarity or the degree to which they identify with each other. Moreover, the content of those responses may be interestingly aligned.

that recordings are less likely to be surprising or shocking, even when we hear them for the first time. Recordings remain the same in each playing. As such they encourage 'perfect' performances. In contrast, live performances can be risky. They can aim at the broad sweep of the music's character without worrying so much about mistakes that would be revealed on further analysis. Thus a sense of the spontaneous, rare, or unrepeatable qualities of a human expression are emphasised by the live performance and undermined by the recording. This impression can confront a listener who listens again to the recording of a live concert he attended. One listener described it as "a pale reflection of what I experienced in the hall, with people around me getting excited, and seeing musicians in animated mode" (Andrew Keener in Philip 2004: 54).

All this seems to cohere well with what Walter Benjamin described as the loss of 'aura' that he says must accompany the mass reproduction of a work of art. Benjamin defines aura as the "unique phenomenon of a distance however close it may be" (Benjamin 1968: section III) and describes it as what gives an artwork its authority and uniqueness. The term 'aura' implies the peripheral social context within which the work is embedded, and which Benjamin claims is essential to its singular presence and our awareness of it as a complete and distinct object. Where social rituals of music go to great lengths to preserve context, one of the essential functions of reproduction is to remove musical content from one social context and translate it into another. The listener may be encouraged to imaginatively reconstruct the original context (particularly if some signs remain such as coughs or applause). Yet recordings often strive to surgically remove any signs of the social context. As result the 'distance' between the listener and the music is reduced. As even signs of

the performer are diminished, so the listener gains more control over the content of their experience. As I mentioned in chapter four, this may encourage the listener to imagine the music as an idealised expression of his own feeling.

This rather solipsistic phenomenon misses out on the almost unique potential for musical performances to unite listeners in experience. Just how great this potential is can be appreciated when we note that the important difference between joint attention to music as opposed to joint attention to most other events is that music is so richly expressive of the inner character of emotion. For this reason to involve oneself in the music and allow it to dominate one's sense of stress and flow is to locate a source for the character of one's inner life that is common to the thousands of listeners around you. It is based on this consideration that phenomenologist Alfred Schutz claims that music has a special capacity to align listeners' sense of 'inner time'. Schutz here is drawing on Bergson's notion of inner time or 'duree' (in contrast to measured clock time). The reason that music can structure inner time is because unlike a mathematical proof, music cannot be grasped all at once, but must always be experienced as a gradual revelation stretched out in time.

Moreover, Schutz claims that the structure of the music captures the subjective stream of consciousness as "an interplay of recollections, retentions, protensions, and anticipations" whereby the listener is continuously re-organising the sounds that he has heard previously as well as anticipating what is to come. This view is similar to Meyer's theory that music arouses emotions by generating and resolving expectations. And just as in Meyer's theory we may worry that listeners' relative familiarity with the work will significantly influence what expectations they have.

Nevertheless, the point is that by being drawn into the same sequence of sounds, the listeners gain a sense of simultaneity with other listeners, in Schutz's terms of "growing older together" (1971: 175), even if they have differing perspectives on that experience.<sup>7</sup> As such Schutz draws a particular connection to the sharing of inner time and the state of joint attention:

[T]his sharing of the other's flux of experiences in inner time, this living through a vivid present in common, constitutes what we called in our introductory paragraphs the mutual tuning-in relationship, the experience of the "we," which is at the foundation of all possible communication.  
(Schutz 1971: 173)

Adorno similarly claims that great symphonic performances can "annihilate... the contingencies of the listener's private existence" enabling the communal elation of an audience. Accordingly he complains that playing symphonies on the radio has 'atomised' the audience and destroyed the traditional collective aspect of the symphony (Adorno 2002: 256-257). Schutz has a different opinion here. He believes that radio listening provides just the same sense of simultaneity as the concert hall, though there may be "variations of intensity, intimacy and anonymity" (1971: 174). Schutz is mainly interested in how the content of the music must be experienced in a particular temporal way. However he doesn't appreciate the extent to which private listening can significantly affect the sense of control over the music. The listener to a radio programme can certainly imagine all the millions of other listeners that might

---

<sup>7</sup> In addition, when performing together, Schutz says that musicians must align both their inner and outer time. This is comparable to the claims about sharing emotion that I make in the following chapter in that they can share their reactions to real events.

be tuned in to the same music, yet that fact does not impress upon him to such an extent that he won't happily turn down, change stations, or switch off the music, whistle or sing along, or walk into another room for a while. The important difference with the concert hall experience is that a basic level of coordination is retained such that we are committed to focusing on the music as long as everybody else is. In contrast, a private listener has far greater control over exactly what he listens to and how he listens to it.

Finally the concert hall listener will be self-consciously aware of his emotional responses in a way that the radio listener will not be. It is this awareness that others can observe one's reactions, and the taboos against any behaviour that might be considered distracting (including exaggerated facial expressions) that is likely to inhibit the concert hall listener's response to emotions in the music. This is especially the case if other listeners do not show such signs of arousal. If on the other hand, other listeners *do* show arousal, we can expect the listener to feel more validated in his own responses. In either case there will be a feedback effect. If other listeners show arousal, then the individual listener is more likely to, which will further encourage the response of others. Likewise the lack of arousal shown by others will dampen down the individual's response and the response of others as a result. Due to the inherent constraints of the concert hall listening scenario, we can expect the dampening effect to occur more often. Nevertheless there is the potential to reverse this effect, though the factors that encourage such a reversal are liable to be complex and subtle.

Overall the most significant effect of silent joint attention to music is to intensify the drama of performance. This may make the emotions expressed seem more vivid or intense, especially since they gain additional social meaning. As such the *potential* for audience arousal is increased. In this way our responses to the music's expressive qualities can be interdependently structured. However the potential for greater integration and convergence of responses cannot be fully realised whilst the audience is unable to openly respond to what they are hearing. It is thus in what I call the 'noisy' case of joint attention that the opportunity for music to generate a deep sense of community in an audience becomes most apparent.

### **Joint Attention to Music: The Noisy Case**

The situation of the concert hall immediately changes when the performance finishes and the audience bursts into applause. Suddenly, a listener can fully appreciate the excitement (or apathy) of his fellow listeners. If he has found the performance to be of high quality, he will enjoy the sense of agreement in the rapturous applause of thousands. So we have at least the beginnings here of what I have been calling the negotiation of the experience of music. Yet because the applause is separated from the experience of the work, it cannot track the moment to moment fluctuations of music itself. So any new perspective that the other listeners provide can only be treated like a single filter through which the experience of the work is affected in its entirety. For this reason I am mostly interested here in reactions that occur during the performance itself. This includes noisy reactions like applause, cheering, booing, whistling, singing or humming along with the music, finger clicking, emotional exclamations and verbal commentary. But in addition I consider clear observable behaviours such as foot tapping, nodding, dancing and explicitly communicative

behaviours such as exchanging looks, smiles, grimaces, or placing one's hands over one's ears.

An interviewee in a paper by music psychologist Alf Gabrielson describes a particularly strong example of the kind of noisy joint attention I am talking about:

The music began before the curtain rose, and you just stood there as semi-paralysed and screaming... Everybody in the audience is exciting each other to a stage next to a climax, and when the artist at last comes on stage he does not have to say more than 'hi' to trigger that climax. It is very much the atmosphere in the audience that gives this concert feeling... One feels so free somehow. At concerts one can dance, jump, scream and sing as much as one wants. You are like a part of it all, not just a spectator. Throughout the whole concert the audience was in total ecstasy. It was the only thing that mattered: the music! ... You don't think about what you are doing. You do what you feel like without even thinking about it. (interviewee quoted in Gabrielson 2001: 437)

There is clearly a massive difference between this situation and the concert hall ritual I presented earlier. When emotional responses are unconstrained to this extent, we are unlikely to have the sense of tension that silent joint attention generates. What we have instead is an orgy of emotional abandonment. It would be hard *not* to get caught up in such a scene. So both the perceived emotional content and impact of the music is virtually guaranteed to be mutually recognised by the audience.<sup>8</sup> In less extreme

---

<sup>8</sup> It could be argued that much of the audience reaction is directed towards the social situation itself. Yet the music is still the foundation and justification for all responses.

situations, it is still the case that enthusiastic or disapproving responses to particular moments in the performance will focus the attention of other listeners onto those passages in order to ascertain the cause of the response (assuming it is not obvious already). Hence an awareness of the responses of other listeners can allow joint attention to particular aspects of the music and its perceived content.

Given also that joint attention already motivates the listener to allow his reactions to be influenced by others, we can expect that when the reactions of others are apparent, that there will be a high degree of *agreement* concerning the character of the music. So it's not just the case that noisy joint attention enables mutual recognition of what others perceive in the music. It also encourages an inner 'endorsement' of certain reactions, that is, mutual agreement about the effectiveness of the music's expression of emotion. In this way, adult listeners have not moved significantly beyond infant social referencing. There are still strong social pressures to emotionally conform with others. Moreover, since the music strongly affects the inner character of experience, this sense of agreement will entail that listeners will be in similar states of arousal. So far this is fairly obvious. If two people are cheering loudly in response to a piece of music, the chances are that they both feel the same way about the music, or more precisely, that they feel the same way about the music as it is performed in that particular social context.

The major difference between listeners will involve just how aroused by the performance they actually are. As I discussed in the second chapter, those that engage in more expressive behaviours are liable to have more intense inner feelings, though some individuals will be internalisers, enjoying more intense ANS activity

with less expressive behaviours, and others will be externalisers, engaging in more expressive behaviours with lower ANS activity. Clearly those listeners who already happen to have similar emotional personalities will have closer emotional reactions to such a situation. In many ways, these similarities will be self-selected for because those who enjoy a particular genre of music and its attendant performance situations are already liable to have similar emotional personalities. We can also expect alignments along the introvert-extrovert divide.<sup>9</sup> Because music has such immediate connections with the character of emotions, with particular types of music tending to express distinctive temporal profiles and intensities of emotions, we can expect that those listeners who already enjoy or aspire towards such emotional profiles will be most likely to seek out such performance situations.

So in the noisy case of joint attention to music, we can expect not just agreement about the character of the music, but a strong mutual awareness of this agreement. This sense of agreement will both intensify the listener's enjoyment of the music and their sense of community with the other listeners. In fact what is less certain about the noisy case of joint attention is how much it could possibly allow *disagreement* about the content of the music. Certain responses will be socially endorsed where contrary responses are silenced. The noisy reaction of others will make it abundantly clear which emotion is socially endorsed. So even if one did not agree about character of the music, one would certainly be disinclined to voice such a sentiment. The most likely result would be a sense of alienation from the group.

---

<sup>9</sup> Stelmack and Campbell (1974) have shown that in response to music, introverts are more sensitive to low auditory stimulation and show progressively less sensitivity to higher levels of auditory stimulation (corresponding with their tendency to inhibit higher levels of arousal). Extraverts reveal the opposite trend (cited in Kemp 1997).

So noisy joint attention increases the tendency of listeners to have similar attitudes towards the music, fixing the characteristics that it is perceived to express. Then because the audience are converging on their interpretation of the music, and this in turn involves not just the sub-personal arousal involved in emotional recognition but also encourages a fully aroused response, the audience will also converge on aroused emotional states. The mutual awareness of this creates a sense of the emotional atmosphere of the situation. Yet although the perception of the music is socially extended, and this perception involves emotional interpretation, we cannot say that the audience are sharing their emotional states in the strict sense. The task of determining one's emotional state is to some extent socially extended, yet this is mostly a matter of either affirming or denying the expressive character of the music. It does not involve detailed control over the character of one's emotion. Moreover the arousal of each listener, though clearly observable by others is still individual to each. As in the case of joint visual perception, the intrinsic content of each listener's emotional state is not constituted by the other listeners' arousal.

But what of the musician during all of this? Certainly musicians are sensitive to the reactions of the audience. For instance the opera singer Caruso claimed to be only able to sing top Cs in the presence of an audience (Davidson 1997: 215). Similarly performer and music theorist Jihad Racy (1998) describes situations that apparently occur regularly in Arabic music where the performer is spurred on by the expressive signals of a knowledgeable and appreciative audience to attain a sense of 'musical ecstasy'. He defines this as an effortless mastery over one's instrument that manifests itself in the form of spontaneous creativity. Regarding improvised performances in particular, the potential for audience reactions to affect the content of the music is

significant. Music analysts Hazel Smith and Roger Dean argue that audiences appreciate when a performance is being improvised and will be all the more sympathetic in supporting the musician as a result (Smith & Dean 1997: 4). As in classical performance there is a risk that the musician may fail, but in improvised performances the audience can recognise their role in ensuring a successful performance such that their reactions can be immediately reflected and incorporated into the music.

Yet despite the ability for the improvising performer to respond to the encouragement of the crowd, there is still a significant division between the state of the performer and the state of the audience. The performer must inevitably take responsibility for the content of the music. He must make decisions, influenced perhaps by the audience response, but ultimately based on his own technical and imaginative capacities. In his position at the centre of attention, the performer can control the emotional mood of the audience. If the performer is interested in expressing his own emotional state, and he is skilled, then this sense of control can afford a sense of communicating his emotions to others, but that does not necessarily involve a sense of *sharing* the emotions of the audience. There are all kinds of considerations that bear differentially upon the performer's state of emotion, such as the technical and social pressures of the task, his capacity to respond more immediately to his emotional state, and his inevitably greater understanding of the productional and structural aspects of the music. As such listeners are more likely to feel the same as other listeners, not the performer.

Let us assume that the goal of the performance situation is to encourage as close a sense of mutual awareness and emotional community as possible. Does this necessarily require the audience to *reproduce* the emotions of the performer or composer as accurately as possible, by means of understanding and endorsing the emotional characteristics of the work? Collingwood argues that the composer expresses emotions on behalf of the audience and that the audience act as collaborators in bringing his expressive intent to fruition (Collingwood 1958: 311-315). He claims that it is important to the artist that many people can recreate his emotional state by engagement with the work, or else the artist cannot be sure that he has had a genuine aesthetic experience. Nevertheless, Collingwood does not mean that the audience and the artist should enjoy the emotion expressed *as a collective*. Rather the artist communicates with each member of the audience separately on an individual basis. Other individuals reproduce the artist's individual emotions.

Tolstoy takes a similar view to Collingwood, but seems more optimistic about the capacity of art to unite the audience in emotion. Adopting a theory of expression as emotional contagion he claims that, "every art causes those to whom the artist's feeling is transmitted to unite in soul with the artist, and also with all who receive the same impression." (1899: 163). Tolstoy tries to overcome the problem of differing expertise by demanding that art only express simple emotions such as merriment, pity, cheerfulness and tranquility or "feelings flowing from the perception of our sonship to God and of the brotherhood of man" (1899: 164). All other arts either unite some people at the expense of alienating others (such as patriotic arts) or are inaccessible to the masses. Apart from his blithe proselytising that all nations and creeds would be better off uniting under Christianity, one unfortunate consequence

of Tolstoy's view is that he rejects Beethoven's *9th Symphony* on the grounds that it is overcomplicated. Nevertheless I think that Tolstoy is right to single out the special capacity of folk art to engender emotional community. The reason for this however is not the simplicity of any emotions expressed therein but the fact that folk musics tend to involve far greater communal participation in the actual production of the music than the professionalised genres of music we currently enjoy in the industrialised world.

The fundamental limitation of joint attention to music is that it still involves a division between those who produce the music and those who listen to it. Although clear audience reactions can go some way towards overcoming the subjectively variable nature of individual responses to music, there can be no guarantee that listeners are in fact enjoying exactly similar emotions. So long as the model of sharing involved is that of *reproducing* the emotional state of the performer, or the emotions of other listeners, there are numerous ways in which such reproduction can be partial or distorted. Thus in the next chapter I argue for a radically deeper form of emotion sharing based on mutual participation in the *production* of music.

In the meantime however, it seems the best way for an audience to become intimately involved in the music and thereby with each other is through dancing together. This is because dancing not only allows the listener to physically synchronise with the music but also with other listeners. One particularly interesting example that music anthropologist John Blacking (1976) describes is the 'possession dance' of the Venda tribe. Here dancers can apparently achieve trance-like states of absorption in the music by physically coordinating their movements with the

drumming. Interestingly however, it can only occur when dancers are surrounded by members of their own cult, presumably people with whom they identify and trust. This highlights how a sense of community with others can enhance, as well as be enhanced by jointly participating in a musical event.

Blacking emphasises the role of physical movement in experiencing one's relation to others. He states, "I do not say that we can experience exactly the same thoughts associated with bodily experience; but to feel with the body is probably as close as anyone can ever get to resonating with another person" (Blacking 1976: 111). Again, Blacking's idea of emotional resonance is based on the reproduction of feelings, in this case an additional reproduction of the dancers' proprioceptive states. Hence although it is guided by a communal source, the rhythm of the drums, it is not as radical a case of sharing as that I explore in the final chapter.

Yet we can still recognise that the more listeners can actively respond to the music the more their individual responses should converge. By interdependently structuring listeners' perceptual activities, joint attention provides a solid foundation for such convergence. Moreover, when listeners can clearly observe, negotiate and agree on their emotional responses, a significant part of their experience will be an awareness that it is the same for others. This will help to validate their emotional reactions, most likely leading to an intensification of arousal. At the same time, by providing a basic 'we' perspective, joint attention allows many individuals to think of themselves in terms of a group identity. An individual listener in these situations can legitimately think of himself as part of an emotionally bonded group. So as long as he remains tuned in to the attitude of the crowd, he is entitled to prefix his descriptions of his

behaviour and attitudes with 'we did', 'we felt'. Though their *intrinsic* emotional states are not shared, joint attention to music defines a plural subject, which listens and responds to the music as a group.

## Chapter Seven: Shared Emotions in Performance

In this chapter I argue that groups can share emotions. By this I do not mean merely that groups can share an emotion type (by having similar although individual feelings towards the same object) but that they can share *token* emotional states. Naturally this will strike some readers as a highly counter-intuitive claim since emotions seem to be paradigmatically individual mental states. It would be hard enough to show that groups could share a single emotion type since emotions are sensitive to all the subjective differences to which mental states are prone. Moreover, whilst it might be more credible that groups could commonly possess an emotion where we conceive of emotions purely as judgements, I use the conception I outlined in chapter one of emotions as primarily kinds of bodily patterns.

However, we have in previous chapters extended our conception of emotions quite considerably. Firstly, I showed in chapter one that bodily patterns are able to represent states of affairs without conscious attention. This separated the intentional content and causal role of emotions from their phenomenal experience. Then in chapter two I showed that bodily or vocal expression can enable cognitive control over one's emotional state. In combination with the arguments from chapters three and four that music is able to directly simulate bodily patterns, this grounded my claim in chapter five that musicians can use music to extend their cognition of emotions. Once emotions have been externalised in this way, the possibility of sharing becomes more plausible. That is, if one person can externalise their emotion using music, and more than one person can get involved in the production of that music, then why can't more than one person possess the emotional state expressed?

Yet making this kind of move is far from straightforward. Within the extended cognition account, the individual still very much remains the locus of the emotional state. In most cases the music is an elaboration of their bodily patterns. So we need to show how music can *socially* extend these patterns such that two or more musicians mutually possess the emotion. In the previous chapter I explained how the task of perceiving an object can be socially extended. I also began to describe how we may socially extend some of our responses to music. But the main concern of this chapter is to show how socially extended cognition can apply not just to the overall way we direct our cognitive activities, but also to the intrinsic *content* of our mental states. In order to show this I first look at intentions to act, because a great deal of extant material already discusses whether these may be collectively possessed. In particular I analyse the position of Margaret Gilbert, who not only argues that groups can possess irreducibly collective intentions but also provides a general model for group states which is applicable across a variety of psychological domains.

When this model is applied to emotions, the claim is that two or more people are jointly committed to having an emotion *as a body* (Gilbert 2000). However whilst I support this basic model, I argue that Gilbert does not have a satisfactory way to explain how bodily patterns can be shared. It is the nature of music as well as the special circumstances of musical production that overcomes this difficulty; what we may describe as its high potential for blending and expressivity. So first some discussion is required of how musicians blend their musical activities such that they can collectively express a *simulated* emotion. Then once this is done I can bring in the considerations about using music to extend cognition of real emotions that I

raised in chapter five. This allows me to describe how one musician's emotional state may 'overlap' with another's. Yet not only are overlapping emotions possible, but also states where the group as a whole possesses an emotional state that is distinct from any state possessed by the individual members of that group. This is not to say that the group has its own consciousness, though I am not as bluntly dismissive of this notion as other writers seem to be (e.g. Searle 1995: 25). Nevertheless by having an emotion the group certainly displays a mental state of its own.

### **Collective intentions**

In chapter six I argued that both infants and adults can form plural subjects of attention in which the task of perceiving is structured interdependently with another person. Clearly it is also possible to *intend* to form such plural subjects of attention, as when we attract the other's attention by pointing to an object. However I did not show that the intention to form such a plural subject could be an intention of the group rather than the interacting intentions of the individuals involved. The intention to engage in an action is a mental state. It typically has the content "I will do *x*"; both representing what will be done, and determining who will do it. So to *share* an intention would require that this content (and causal power) is possessed by a group and not by the individuals involved.

Note that I don't mean we share the intention as we would share a cake, each person taking a portion of the totality. Rather I mean sharing the intention as we might share ownership of a house. Though each person may make individual contributions to generating that state (e.g. paying the mortgage) they only enjoy ownership as a group.

By shared possession of the content (and causal powers) of an intention then, I hope to show that a mental state can belong to a group.

Since it is quite possible to intend to jointly attend to some object without worrying too much about the other's intention, as when we forcibly attract their attention, the case of joint attention is not a very good example with which to demonstrate the possibility of collective intentions. Instead it is more instructive to find an example that could only be achieved with the aid of a collective intention. This is not to say that we cannot collectively intend to jointly attend to something, or that collective intentions aren't very common. We can and they are. The point is just that once having recognised the existence of collective intentions we can then appreciate how many group activities utilise a similar structure of interactions.

For instance, an action such as performing a symphony is clearly not something that I can intend to do, or in normal circumstances, force others to do. Hence we might suppose that only the orchestra as a whole can intend to perform the symphony. At this point however, it is reasonable to object that no single intention to perform the symphony is necessary. Rather each person involved need only individually intend to play his or her *part* in the symphony. However, as Raimo Tuomela notes, we can only make sense of the individual intention to play one's part given the proper context. It is only rational to do my bit if I can be fairly sure that everyone else will do their bit as well (Tuomela 1991: 15-16). I must be aware of what it is that we are all supposed to be achieving as a group. I must desire that this overall goal be achieved. Moreover, I must be reasonably sure that everyone else has this goal, either as the result of explicit agreement or by long standing convention. Without this, it

could only be fortuitous that everyone happens to perform the right actions at the right time to achieve the goal. Ultimately then, it is this goal about the collective action that explains and coordinates our individual actions to play our parts.

Yet as recognised above, I simply cannot intend to achieve this goal as an individual, since the individual actions required are entirely conditional upon the corresponding actions of others. Since the goal is only achievable by something that *we* do, I can as a result only intend that *we* do *x*.

At this point we can still understand the collective intention as something possessed by individuals rather than the group as a whole. The collective intention that *we* *x* could be a type of which the group members each possess tokens (cf. Searle 1995). Yet despite the fact that this intention would be irreducibly an intention *for the collective*, it is not enough to guarantee the performance of the group action. We still require a sense in which my possession of a we-intention is dependent on your possession of a corresponding we-intention and all of this is common knowledge between us. Otherwise, since my we-intention belongs to me alone, I could unilaterally rescind it. Yet if I decided that I didn't want to play the symphony anymore and walked out during the performance I would not have thereby negated my obligation to act according to the intention. I cannot remove the intention like this. In normal group actions, once I've agreed to do something together with others, I am committed to doing what I've agreed to do until we all agree to do something else instead. So there is a distinctive normative aspect of collective intentions that shows I am not completely in control of when and how the intention comes into force (cf. Gilbert 2000: 16-18).

Michael Bratman (1993) has presented an account that incorporates these mutual dependencies with additional refinements to ensure that any subplans we have as individuals to satisfy these intentions are also compatible. Yet whilst an intertwined web of individual intentions goes some way towards a system that is only realised by the group as a whole, we may object that it fails to properly capture the nature of intentions. In the sense of intentions that we are interested in here, I can only individually intend to do  $x$  if doing  $x$  is under my control and this intention goes some way towards actually bringing about the desired result (Velleman 1997). For example, if I desire to kill Smith, and order a taxi to take me to his house, but on the way the taxi happens to run over Smith and kill him, the result I intended has been satisfied, but I certainly haven't killed him intentionally. Rather the intention must represent *itself* as responsible for bringing about the action. The intention determines how the action is to be carried out. So for example, uttering, "I intend to go for a walk," resolves me to go for a walk for the very reason that I have intended to do so. Hence equally, any individual intention that *we*  $x$  cannot actually bring it about that *we*  $x$ . I cannot intend that we perform the symphony because my forming this intention cannot bring about the action, or represent itself as bringing about the action.<sup>1</sup> Only an agreement of some kind can do this. So we are only entitled to identify this *agreement* as the true intention.<sup>1</sup>

---

<sup>1</sup> In his article 'We intend that J' (1999) Bratman has responded to this objection by claiming that one can intend *that*  $x$  despite not being able to personally bring it about that  $x$ . Just as I can intend that my children go to college, I could intend *that* we  $x$ . However Bratman has been accused of changing the subject here. Normally, when an individual intends  $x$ , there is something he intends to *do* to bring about  $x$ . Hence similarly an intention for some group action should be an intention to *do* something as a group.

It is worth exploring the nature of agreement in some depth, because it is crucial to all kinds of socially extended states. For example, suppose you say “we should play Bach” and I say “I agree”. What exactly is happening here? For the moment, let us just concentrate on what is happening with me in this episode. I have adopted the intention to play Bach, and I have done it by means of finding out your intention. But how important is that process to identifying how my mental state is physically constituted? It seems to depend on just what my purposes are when I say, “I agree”. In some cases, I may just be interested in expressing my own intention. However in many other cases, I may also be specifically interested in matching your intention. If for instance, I acquired the intention to play Bach, but later found that I had misheard your utterance “we should play Bartok” I may be disposed to change my intention accordingly. We might say that my attitude here is rather weak willed. Yet my main motivation may just be to do something together with you.

So it seems that in many cases where I agree with you, I am concerned to track your intention as closely as possible. In these cases, I *defer* the content of my intention to you. In other words, I intend *whatever you intended* when you made that utterance. So part of the content of my intention makes an ineliminable reference to your intention. This is the initial motivation for the familiar externalist claim that part of my intention, its meaning or content, is partially located somewhere outside my head.

Now we might explain what is happening physically here with the aid of types. So we both have qualitatively similar though numerically distinct intentions in our respective heads. We might then say that this type does not physically exist in either

of our heads but somewhere in the interactions of our linguistic community. Yet this is not entirely satisfactory. My intention is basically a shadow of yours. When I say “I agree”, I am implicitly saying “I intend *that*”, whilst mentally pointing at your intention.<sup>2</sup> In some cases I may not even be aware of what it is that I’m agreeing to, so I don’t independently possess *any* token of the intention type. I may defer to your intentions to such an extent that when I want to find out what I intend, I have to get you to say it for me. Suppose for example that I’m on trial, and I want to decide what plea to enter. I ask my lawyer who says to me, “non est factum” which I accept, trusting his judgement despite the fact that I have no idea what ‘non est factum’ means or can even remember the words a few moments later. We should say here that I have decided what to plea *via* my lawyer’s intention.

This is the kind of extended cognition that Clark and Chalmers describe in ‘The Extended Mind’. Furthermore, like the waiter that they mention, who decides what food I shall eat for me, it is a case of *socially* extended cognition. Now we may complain that in the case I described, I simply have the more vague intention ‘I will plea whatever my lawyer says’. Similarly, when two people agree in their intention, we might explain that each really has an independent intention in their own heads. It’s just that those intentions *influence* each other. It’s certainly a part of the *causal* story that my intention is formed by finding out yours, but this cause need not be a physical part of the mental state itself. Just as for sunburn to really be sunburn it must be caused by exposure to the sun, but this does not mean that the sun is literally a physical part of my sunburn (cf. Davidson 1987).

---

<sup>2</sup> In addition, since my intentions can only settle what I do, and your intentions can only settle what you do, we cannot have mutually identical intentions, or tokens of the same type, since strictly speaking, the subject of those intentions differ (Stoutland 1997 cited in Tollefsen 2004).

One of the replies to this is the parity argument that Clark and Chalmers make. According to this, I have the particular intention ‘I will plea non est factum’ because it is completely specifiable when occasion demands. This intention is accessible, reliable, and it guides my actions. So there is a parity between it and any other intention I might have more independently. Now our intuitions may differ on whether this is a satisfactory reply. Some might say it’s just more parsimonious to locate the vaguer intention as what is really guiding my actions, where others might complain that that’s just an internalist prejudice.

I think that what decides this issue is whether *another* internal representational state of mine is always required to get at the content of the extended state, in order for it to do any of the work that mental states normally do. So if I always have to think, ‘I will plea whatever my lawyer says’ in order to get at the intention, and then once my lawyer speaks I must then form an internal intention “I will plea non est factum” in order to act on it, or alter any of my other mental states, then there seems to be no reason to admit the extended state as a physical part of my mental economy. However, this doesn’t seem to be the case here, because in court, the lawyer can stand up and say “he pleads non est factum your honour” and I can sit there approving the process without ever having to form the specific intention internally. The plea meanwhile is still recognised as my decision.

Now we might debate whether this is an accurate description of the situation. The lesson I want to draw is just that we should only call something a mental state if it has an independent and distinctive representational content in the way I have

outlined above. What I want to concentrate on is the *process* of agreeing with somebody. It seems fairly clear that as a case of *deciding what to think*, agreement is an example of a cognitive process. It's a kind of reasoning. It is also clear that the process of agreeing is an extended process. To illustrate we can adapt the example that was used in chapter five: Suppose I were a brain in a vat, and I began the process of deciding what to intend by way of finding out what you think. But rather than actually tracking what you think, I am given some algorithmic signal by a computer and this brute causal process is what leads me to think 'let's play Bach'. In this case, I haven't agreed at all. I haven't even gone through a decision process properly so called, because my intention to make up my mind using a particular method has not been achieved. Rather I have rather passively acquired a new intention by a quite different process.

As I mentioned in chapter five, the external part of the process (the agreement) typically goes on to impact the internal part of the process (the brain state). However it is unreasonable to pick out any particular part of this system as *the* mental state. The content of the mental state could change externally and at that point have a causal effect (it could be communicated to another person for instance), before that change has been registered by an inner component of the mental system. Yet I still remain responsible for that content (despite not knowing what it is exactly). In addition, the external stage could be read off by *another, distinctive* internal state (e.g. a belief about what I intend), and thus play a role in modifying another mental state. The same is of course true of the inner stages in the mental process. Any part of the system could be the 'point of contact' in a process of mental causation. But because of this parity, we should only identify the system as a whole as the true mental state.

So what I would like to emphasise is that agreeing is a mental process or system that is partly physically constituted by the thoughts of another person. And the mechanism by which it is extended is by partially giving up control over the mental process to another. In the process of specifying what it is I agree with, something outside of me is partly in control of what that is. Such that within a certain range, I would endorse *whatever* the outer process resulted in. In general, deferring control or authority over one's mental state is the key factor to *all* cases of externalism and extended cognition; whether control is given to a person or some other feature of the environment.

Now the reason this can happen is because control over a process is something that can be held to a greater or lesser degree by the various components of that process. But another interesting feature of control is that it is also possible for two or more people to completely share control over some matter, such that one could only say that the group controls it. So if control over a *cognitive* process is shared in this manner, it would entail that the group possesses that mental state.

In the legal case that I mentioned above, there was still a definite sense that one person remained the locus of control. When it gets down to it, *I* have to enter the plea 'non est factum' or not and I can unilaterally veto this decision if I so choose. It is also possible to agree with someone without their ever knowing about it, if for instance, I was hiding in the bushes listening to them soliloquise to the stars (cf. Campbell 2005). However, when we return to the case of playing Bach from the perspective of *both* people involved, we can only agree on what we will do *together*,

or at least agree on who should decide the matter for us. The other person relies as much on tracking my intention as I do on tracking his. As such only the interaction itself, the agreement, has the final authority over what is intended.

Overall, my notion of agreement is similar to Margaret Gilbert's idea of joint commitment, which she uses to talk about the possibility of collective mental states in general. J. David Velleman elaborates Gilbert's notion of joint commitment as standardly being formed by the joint conditional 'I will if you will... then I will'. He says:

Each of us places his behaviour under the joint control of both, by issuing an intention that's conditional on the other's intention... The result is that each of us conditionally settles, and is represented as conditionally settling, one and [the] same set of issues - namely, how both of us will behave - and we thereby categorically settle those issues together.

(Velleman 1997: 48)

Velleman goes on to say that we both make up both of our minds, though he qualifies that "this joint making up of minds is not the making up of a joint mind" (1997: 48). But then what exactly are we talking about physically when we say that the group has an intention? Velleman explains that an oral or written agreement can literally be an intention, since it represents the action to be performed and plays the right causal role in bringing about the action (Velleman 1997: 37-38). But if the agreement functions exactly like an ordinary internal mental intention to bring about an action, why does Velleman deny that it's a mental state? His denial seems like

nothing more than an ‘individualist’ prejudice because the agreement, the collective intention satisfies the conditions I gave above for being a mental state.

To explain: First of all, no additional individually based intentions are required for the collective intention to do its work. The reason for this is because the collective intention is not floating around in type-heaven, existing independently of the tokens possessed by individuals. Rather the collective intention is generated by, and *directly supervenes* on the conditionally intending states of the individuals involved plus the various relations between them. No change in the collective intention can occur unless some component of this system changes. So the collective intention can function independently of any *additional* individual intentions because the agreement already derives its representational and motivational properties from the neural states of the people who make the agreement!

Yet note that what I’ve called the conditionally intending states of the individuals are not complete intentions. On their own, the individuals’ brain states do not have the power to decide what the goal is and how it should be achieved. It is only once they stand in the right relations to each other that they (jointly) gain this content and causal power. So the collective intention is made from individual brain states that are like intentions in some ways, but only the extended system as a whole, the agreement, is the true intention. The agreement itself is the ultimate locus of control.

As a result, the individuals involved must defer to the agreement itself as specifying the *content* of whatever is decided on. Where problems arise, they must refer back to what was agreed in order to settle how to proceed. So individually, each person may

retain only a *partial* grasp of the content of what is agreed. Hence as well as being independent, the agreement itself has a distinctive representational content. In respect of both content and causal power, the collective intention satisfies my conditions for being a mental state. As such it is only an individualist prejudice to deny the agreement the status of a mental state.

What I would like to emphasise is the mechanics of this process. It is still the individual brains of the people involved that are doing most of the work in actually implementing the intention. The point is that the connections between these individuals, the way they interact, is doing some important work in structuring the process, and in specifying its content. Moreover, where the locus of control is the agreement itself, we are able to say that it is a mental process performed *by* the group *for* the group. Agreement defines a group subject that has ultimate authority over the state in question, and to which all members of the group must defer.

*Still* we might complain that it's just a combined set of individual intentions. But this is not a fair description because in collective intentions, the individual states involved are *blended* to form a distinct and *unified* authority over an issue. What do I mean by blended? Well, there are difficult issues concerning emergence here. But in general, a group activity or state is blended where it forms an effect that cannot be decomposed into parts whilst remaining that effect, such that responsibility for the parts could be attributed to the individuals involved. This is clearer when we consider the kinds of activities that can be intended in this way. For instance, if two people are pushing a car up a hill together at 5 mph, it is ludicrous to say that each pushes half the car, or that each pushes the car at 2.5 mph. It is true that each person

is contributing a measurable force towards the effect. And the movement of the car is entirely supervenient on the force they individually exert. Yet you cannot divide up the movement of the car in a simple temporal or spatial way, a way where each part could still be described as the moving of the car. So we cannot meaningfully ask which part of the car moving effect is the responsibility of which individual. Similarly, if two people own a house together and contribute towards a joint bank account that is then used to pay for various repairs, we cannot take any given repair to that house and ask which person paid for it. The owners paid for each repair as a group.

Music is a particularly good example of a blended activity in a number of different respects. Take for example a group of musicians playing a dissonant chord together. One musician plays an E, another plays a G and another an Ab. We might say that responsibility for the chord can be parceled out to the individual musicians involved since each plays a different note that in combination form the chord. Yet what about the dissonance or the particular harmonic quality that is generated? Can we divide up the dissonance and say that each musician is responsible for a part of that dissonance? Certainly they contribute different notes, but this is a different sort of description. The overall effect, the dissonance, is not something that can be meaningfully partitioned whilst each part remains a dissonance.<sup>3</sup> Similar arguments could be made about jointly producing a syncopated rhythm, the overall volume, and most significantly for me, the expressive quality of the music. It is not meaningful to ask which performer is responsible for which bit of the sadness of the music.

---

<sup>3</sup> The case of difference tones; where a third tone or vibration effect is generated when two similar tones are played together is a particularly good example of this.

Playing music together requires the utmost blending of individual actions. Equally when I listen to the music, I am not listening to the oboe part and the flute part and the timpani part and so on. I am listening to the music that they produce together. Accordingly there is one overall goal; to produce a piece of music. And it is only the blending of the individual intentions that is capable of achieving that goal. The collective intention has a causal power that no individual intention can have, though the causal power of the collective intention supervenes on the conditional causal powers of the individual states plus their relations to each other. The individual states involved must blend prior to having this effect. Their effect is mediated by needing to be agreed first. Thus there is one unified cause for one unified effect.

A further important point to make is that agreement can be achieved in a variety of mediums, not only words. When we understand agreement in a broad sense, we may observe that given the right sort of context, simply engaging in cooperative behaviour can signal agreement. For instance it is possible that several people could spontaneously help to push a car up a hill without any form of verbal negotiation. The fact that such cooperation would be obvious and fairly normal to the parties concerned would mean that as long as nobody protests, they could mutually assume to have tacitly agreed on what to do. Similarly a group of musicians playing together could agree on what to play purely by starting to cohere on a particular song. Of course, in different contexts what exactly counts as agreement will differ. As I discuss below, different groups of musicians may have different standards about what counts as a unified sound. The point is that agreement can be an ongoing process, and embodied in the actions of the participants. It doesn't have to be something that begins and ends with a prior verbal negotiation. What matters is that

control and content is deferred to whatever the medium of interaction is, be it actions or words.

So I think that the concept of agreement is a good way for us to characterise group states in general. Moreover, by recognising that one person asymmetrically agreeing with another is itself a socially extended state, we can see that group mental states are built on top of individual socially extended states. What distinguishes these states from group states is merely the locus of control. Hence by appeal to the locus of control, we can posit a single dimension or variable of mental states ranging from the traditionally internal, to the extended, to the *socially* extended and then finally to the group state.

#### **Dimension of mental state extension:**

Internal individual – Extended individual - Socially extended individual - Collective

#### **Gilbert on Group Emotions**

Having established that groups can share token intentions the foundation has now been laid for group emotions. Group intentions provide a model for the structure of group mental states generally. Margaret Gilbert generates such a model from her notion of joint commitment:

For the relevant psychological predicate “X” and persons P1 and P2, P1 or P2 may truly say “We X” with respect to P1 and P2 if and only if P1 and P2 are jointly committed to X-ing as a body. (Gilbert 2000: 19)

As in joint intentions, commitment here involves mutually and conditionally expressing willingness to engage in  $x$ , and taking on certain normative obligations as a result. Applying this to emotions, Gilbert claims that the nature of the participants' commitment is, as much as possible, to constitute by their actions and words a single body that feels the emotion (2000: 135). This, she argues, properly captures what it is we mean when we say that a nation mourns the death of their leader, or that a family is overjoyed by a daughter's marriage.

Gilbert is specifically interested in emotions that the group has regarding its own group actions, such as an army or nation feeling remorse over its conduct during a war. The focus on group actions is however unnecessary for group emotions, it merely helps to strengthen Gilbert's case. This is because a group emotion must be directed towards a content that impacts on the group as a whole. The nature of emotional perceptions requires a sense of the group as a definite subject that dynamically interacts with the world. Group acts are therefore good examples of such interactions, especially in so far as emotions motivate certain behavioural responses. However group emotions could also be legitimately directed at any situation in which the group as a whole finds itself in. There is a definite sense in which groups as a whole rather than their members can be in situations and can have goods independently of the goods of their individual members. For instance the philosophy department could flourish by working all its lecturers to death (though presumably not all at once). Similarly a company could go bankrupt although all its employees remain financially stable.<sup>4</sup>

---

<sup>4</sup> Keith Graham (2002) argues at some length that because groups can act, they can be regarded as distinct moral agents with the capacity to harm or to be harmed by others.

By focusing on group actions, the point that Gilbert is trying to establish is that just as individual intentions to do one's part cannot add up to a collective intention, so emotional feelings towards one's contribution to a group action cannot add up to a group emotion.<sup>5</sup> For instance, one's individual actions may have been entirely honourable, or one may have even protested against the group action, or simply have been unaware of it at the time. Yet it is still intelligible that a person feel remorse over the action of a group with which he is associated and legitimately say 'would that we had not done that!' The member of the group may feel no remorse on his own account, but only remorse *for* the group.

However, this case is equivalent to having individual intentions that *we x*, in that it belongs to individuals within the group rather than the group as a whole. It is not for that reason impossible to have such emotions, though they cannot impact on the group's future actions in the way that individual emotions can impact on an individual's future actions. It is just not a group emotion. In order to be a true group emotion, Gilbert says that there must be conscious agreement between the members of the group to endorse a certain emotional response. In order to justify this, she contrasts it with two cases that seem like less than group emotions. The first of these is where each member of a group feels an emotion (in Gilbert's example, remorse) towards an act that the group has engaged in, but keeps it secret from the others. This is clearly not a group emotion, since each member may believe himself to be alone in their feeling of remorse, so there is no sense in which the emotions of the members

---

<sup>5</sup> Equivalent to this would be having an emotion directed at one's personal situation within the groups' situation. For example, within the panicked reaction of a crowd to a gunshot, I may well fear only for my own safety (which is one of the reasons why panicking crowds are not good examples of group emotions).

are combined or coordinated in a way that intuitively captures the cases in which a nation mourns or a family is overjoyed.

In the second case Gilbert offers, there is common knowledge amongst the group that each member feels remorse over the group's act but this emotion is never publicly expressed. For instance, the participants may be already committed to expressing the opinion that the group can do no wrong and so all their statements to each other and the wider public may indicate that their actions were completely justified. This case is less clear than the first, since we might want to say that the group has an emotion but, as is possible in individual cases, it is 'in denial' about it. Yet just as in cases of individual denial where we might be reluctant to say that the individual really *knows* what emotion they are undergoing, it is unclear whether the members of this group could really be said to have common knowledge about the emotion they all possess. If the members of a group explicitly state to one another that the group action was justified, then the remorse of the members may not be sufficiently out in the open to qualify as common knowledge. The example could perhaps be fixed in this respect, or we might accept that there are borderline cases of group emotions. However the point is that unless the members' expressions of emotion are in some sense combined or coordinated, we cannot say that the group as a whole *represents* the emotional object.

The essentially perceptual status of emotions entails that they must represent their causes or objects. Within a group, unless this representation is done 'as a body' we can only say that the individuals all have separate but related emotions, i.e. individual tokens of a common emotion type. Such similar emotions are of course

possible, but it is not the kind of genuinely group emotion that we are interested in here. Generating a group emotional representation towards a particular state of affairs is vital for establishing that the emotion is truly the group's reaction to the group's situation.

That this representational state belongs to the group is more clearly seen where it differs from the representations of the members of that group. We might think that where the members of a group are reasonably aware of the group's situation, the content of individual and group representations may not differ very much. Yet as in cases of joint attention, the members of a group can mutually fix the target of an emotion and cooperatively fill out the details of the representational in ways that need not match with any member's individual representational state. More importantly, the role of the representational state may differ in terms of its capacity for directing group actions, especially since the group explicitly agrees on it. Group representational states may also differ in terms of vividness or intensity, temporal duration or complexity (which could be greater or lesser along each of these variables).

Yet this would not be just a combined set of individual representations because as in collective intentions, there is an agreement to which everyone's representations must defer. If for instance, it was found that something about the different participants' emotional representational states was in conflict. This would have to be resolved on pain of destroying the agreement. Again what counts as agreement or conflict depends on the context. For example, some groups may resolve apparent

inconsistencies by allowing that they feel happy about one aspect of the situation though sad about another.

In the cases that Gilbert imagines, the group representational state is achieved by public expressions of emotion and associated behaviours. Yet if all that is required are expressions of emotion, (which the individual may not even commit to individually) what is to prevent any kind of group activity from qualifying as a group emotion? For instance, if a wrecking crew destroys a building, does this qualify as a case of group anger because the various noises that they create combine to form an expressive representation of the situation, and moreover, motivates each person to destroy the building with more vigour? Or for that matter, need the participants even be human? Could a collection of machines engaged in the same activity express anger as a body?

I think we can resist these kinds of counter-examples by reminding ourselves of Dretske's theory of representation. According to Dretske, a representation must be *designed* either by evolution or learning to represent what it represents. It is not enough that the state just indicates its cause, the way tree rings indicate the age of a tree. An organism's bodily pattern represents the emotionally relevant aspects of its situation as a result of evolution, and within its lifetime, it can learn to associate the same core relational theme with different kinds of situations (such as imaginative or social objects). So even if the sounds that the wrecking crew makes are caused by its destructive activities and these sounds have some kind of behavioural consequence, they do not emotionally represent the group's situation because they weren't designed to do so either intentionally by the wrecking crew, or as the result of

recalibrating a pre-existing emotional response. Hence there is at least one important difference between individual and group emotions. The emotional nature of the group representation must be *derived* from the emotional nature of their individual representations. Collective representations must be acquired either as a result of explicit intention or gradual recalibration.

So far then, a group emotion involves an agreement to form an emotional representation as a group about a situation that impacts on this group. However, we have not yet tackled the most troubling problem of all regarding the existence of group emotions, which is the precise nature of these group representations. In ordinary cases, individual emotions represent core relational themes with patterns of bodily changes, but how can a group have its own bodily state, let alone its own phenomenal feeling? Gilbert to some extent sidesteps this issue by claiming that bodily feelings are not essential components of emotional states. However she does argue that individual ‘pangs’ of remorse can still be part of a group emotion, since they are responses to the group’s remorse. Thus she says,

Had the group not come to feel remorse, one might never have felt this way. And one’s feeling this way may not correspond to any judgements one has made in one’s heart with respect to the group’s act or to any associated act of one’s own. The pangs in question, then, may best be described as pangs of remorse associated with the group’s remorse or, more succinctly, as “pangs of group remorse.” (Gilbert 2000: 136)

Whilst I concur with Gilbert that these pangs may be *about* the group's act, and signal a commitment to the group's avowed emotion of remorse, they are emphatically *not* the group's pangs. Rather they are *my* pangs that I perhaps imaginatively contribute to the group.

This is precisely analogous to the individual intending that *we* engage in some group act. So for the very same reason that an individual intention that *we* cannot, on its own, be part of a genuine collective intention, a pang of feeling on behalf of the group cannot be part of a genuine group emotion. It is not *blended* with the feelings of the others. Hence whilst Gilbert's response may be more permissible from her perspective, given her cognitive conception of emotions in which bodily feelings are construed as *responses* to emotions rather than their central characteristic, it is completely inadequate for my favoured conception of emotions.

A principle of sharing must be properly articulated here: The only mental states that can be shared are those that are *embodied* in an agreement. Without this, we could only point to mutually influenced mental states. And the only way for a mental state to be embodied in an agreement is if its *content* is manipulated in the process of agreeing. This is what allows the agreement to have distinctive representational content and to be capable of playing a mental causal role without needing to be duplicated by some additional internal mental state. Now because Gilbert's examples involve the participants reciprocally forming their feelings (I feel remorse because you feel remorse), we can say that they truly share the *overall task* of deciding their emotion. This is just like the cases of noisy joint attention that I discussed in chapter

six. But they could never share the intrinsic content of their feelings, since these are never embodied in the agreement.

Of course most philosophers' main motivation for resisting the existence of group emotions will be because emotions seem to rely on bodily sensations so much. Gilbert tries to avoid this problem by focusing more on the evaluative concomitants of emotions. But I want to grasp the bull by the horns and argue that groups really can share emotions even though bodily changes in response to situations will always remain the central case of emotional states. It is for this reason then that we must return to the expression of emotion in music.

### **Agreeing in Classical Music**

Before showing how musicians can collectively possess their emotions in music, it is important to first explain what it means to agree in music. Although as I argued in chapters three and four, music can capture the intrinsic content of emotions, the musicians cannot *agree* on that content without significant mutual coordination of their actions and attitudes towards the music. So at this stage I only concentrate on musicians agreeing on a *simulated* emotion, without worrying about whether the music is directly integrated with the performers' actual emotional states. It is like the difference between a group of people collectively describing a fictional situation as opposed to collectively describing a real situation that confronts them.

To begin with, the grounds for thinking that musicians are mutually sensitive to what it is that they are performing are very good:

In a musical ensemble... the interaction between players is of such a nature that mis-timings even of a fraction of a second, minute hesitations, slight differences in intonation, tiny misjudgements of dynamics and so on are regarded as monumental blunders even among musicians with quite modest pretensions. Across a wide range of variables almost total co-ordination is required, and this necessitates an exceptionally high degree of interpersonal awareness. (Young & Colman 1979)

The view above expressed by psychologists Young and Colman is far from unique. There is wide agreement that ensemble performance requires an extremely high level of mutual sensitivity (e.g. Blum 1986: 14-15, Waterman 2003: 123, Goodman 2002: 156-157). However, the purpose of this sensitivity is not necessarily to establish a completely uniform approach to the music. Several analysts emphasise that part of the interest of a form like the string quartet is in the balance between diversity and uniformity. For instance, David Waterman questions whether if one player could somehow perform all the parts of a Beethoven string quartet, it would eliminate something vital from the work (Waterman 2003: 100). The struggle to achieve ensemble between four different players (or to subvert it) may be part of the drama of the music. Similarly the Guarneri quartet make the point that sometimes the instruments need to stand in relief of one another rather than blending *too* well. As such, they do not seek a uniform approach to things like vibrato or sonority, placing as they say “a high premium on variety” (Steinhardt in Blum 1986: 3, cf. 6).

Yet this allowance for diversity may belie a deeper underlying uniformity. When we look at the details of group performance, we find that contrasts are ultimately

grounded on a bedrock of solid ensemble playing. For example, it would be unthinkable for the musicians not to stick to a joint tempo. Any form of rubato playing (where the timing of notes is stretched or shortened) is usually limited to the lead instrument and must still adhere to the overall pulse. In addition, where the score indicates common timings amongst instruments for phrases and rhythms, special attention is usually made to attacking and finishing notes in perfect unison. Intonation between instruments must also be balanced, or else where the parts form chordal textures, the music will sound horribly distorted. Likewise dynamic balance must be carefully attended to so that whilst each instrument remains audible, the lead instrument at any given time stands out.

Ultimately contrasts in vibrato, sonority, and rubato are fairly subtle. And whilst they certainly enable the listener to pick out the different instruments, they by no means undermine the conception of the music as a single unified work. These details simply add to the complexity and contrast within that single work. So even if the listener is not encouraged to experience the music as the product of a single person, they should nevertheless understand that the four people are combining to produce a single piece of music. It is in this sense then that ensemble performance is an exceptionally blended activity. The musicians must collectively express the content or meaning of the work, and given that this cannot simply be ascertained from objectively identifiable aspects of the score, they are required to align their subjective interpretations of the music in order to do it.

Regarding the *mechanics* of establishing a collective interpretation, we may identify two important strands; negotiation within rehearsal and mutual tracking during

performance. Regarding rehearsal, modern quartets like Waterman's and the Guarneri quartet tend to favour a democratic approach to settling musical details, (in contrast to traditional groups where one musician, typically the first violin, would act as leader). Unlike other kinds of democratic groups however, the string quartet must come to *unanimous* agreements, at least for the purpose of any single performance of a work, and those agreements should hopefully add up to a consistent vision of the entire piece. Hence the crucial importance of rehearsal negotiation should not be underestimated. Quartets survive or perish according to their ability to make joint decisions. As such, the compatibility of personalities within the group is important such that they must feel able to take and give criticism as well as accommodate each others' way of understanding the music. For instance, some musicians may prefer a technical approach to analysing the music where others are more comfortable with intuitive or metaphorical directions such as "it sounds too hot" (Waterman 2003: 102).

The overall purpose of such negotiation is to generate 'cognitive directives' where the musicians agree on how they are to *think* about the music. The effect of this is to coordinate the way the musicians perceptually organise the sounds they are producing. That is, the way that audible events are lumped together as well as the way some musical characteristics are prioritised over others. This will significantly affect the manner in which musicians actually hear the music, and as a result how they play it. Cognitive directives can then function at several different levels, such as measuring the tempo in the same way (in 4 beats rather than 8 for example), or where phrases are divided, and eventually more overarching features such as expressive content.

Elaine Goodman (2000) has investigated the way that musicians negotiate the expressive quality of works during rehearsal. In one illuminating episode, she describes the discussion between a cellist and pianist when jointly deciding how to execute the climax of a duet by Chopin:

The pianist described the mood as 'hesitant passion', while the cellist depicted the growing intensity of her levels using an analogy: 'it's like you're asking a question, and then you start crying, and you ask it a bit louder, and then you scream the answer'. They resolved however, to 'make it sound more searching'. The searching metaphor mediated between the two formal conceptions (i.e. levelled versus continual build-up) and provided a way to unify their conflicting insights. (Goodman 2000: 114)

Agreements such as these will have a widespread impact on the execution of all kinds of musical details.<sup>6</sup> However whilst this kind of discussion is not uncommon, we may expect that musicians will mostly communicate on a more practical level, and it is only having reached a fairly deep level of analysis, perhaps motivated by more intractable differences in interpretation, that such emotional insights will be discussed. In addition, where Goodman analysed the interaction of musicians playing together for the first time, we may expect that long term musical partnerships, being

---

<sup>6</sup> Although we can expect that different instruments may well have a different role to play in generating the overall emotional content, for example one instrument may need to keep strict time in order to allow another more expressive freedom (cf. Waterman 2002: 106). If however each musician is responsive to the total musical product as opposed to only their contribution, they will still need to perceive the emotional content of the music in similar ways.

more familiar with each others' musical approaches, may generally require less verbal negotiation during rehearsal. Long-term ensembles may well have settled various basic issues over the years and established a group style. They are also likely to be more capable of intuitively anticipating each other and so more comfortable with allowing some details to be settled during the course of the performance itself.

This moves us then towards the second strand of collective interpretation, which is mutual tracking during performance. Mutual tracking is mostly achieved by close attention to aural cues, although the role of visual cues such as eye contact or hand and other bodily movements is also significant (cf. Goodman 2002: 156-159 and Davidson 2002). Here the focus is not on explicitly establishing that the musicians think about the music in the same way, but on reacting immediately to each others' performance in order to achieve a balanced and coherent musical product. From a functional perspective however, the consequence of this is that the musicians are indeed converging on a common idea of the music, at least as they *express* that idea. This is because by tailoring their musical productions to each others' performances, the musicians are mutually establishing one particular interpretation of the piece.

The Guarneri Quartet (having played together for over 20 years) particularly emphasise this form of coordination, regarding it as more spontaneous and hence more exciting for the performers and less likely to result in bland performances. They also see it as encouraging more genuinely communicative interaction in contrast to, for example, planning which musician will take the lead at any given moment thus:

The actual lead may be in the first or second fiddle, but, in fact, everyone feels it at the same time; everyone is thinking towards a central point; the start of a piece, a ritardando, or whatever it may be. We don't follow each other; we play together. (Soyer in Blum 1986: 14-15, cf. 16 & 19-20)

Similarly the viola player Michael Tree says (referring to John Dalley, 2nd violin), "I'd rather not play quartets at all than nail everything down in advance. There's rubato in every note; I have to try to climb into John's psyche" (Tree in Blum 1986: 6). In general, the members of the Guarneri Quartet rely on an intuitive understanding of each other as individuals as well as a mutual trust in each other's judgement. This corresponds to the more informed levels of empathy that I described in chapter two. Hence we can expect that where one musician expresses emotion in their performance, the other musicians will be particularly sensitive to the attitude of that musician. In addition, the musicians will not always respond to each other as individuals, but will respond to the attitude of the group as a whole. For instance violinist Arnold Steinhardt says:

If two people are clearly in accord, the others sense it, and there's a greater chance of unanimity. We've never talked about this, but it's nonetheless something we've always done. (Steinhardt in Blum 1986: 11)

So by means of mutual tracking during performance or establishing cognitive directives in rehearsal it is certainly possible (and probably very common) for groups to converge on a unified expressive content. Thus they collectively produce a simulated or fictional emotional state.

At this stage we may note that according to the evidence I presented in chapter two and three, in order to register such an emotional quality, the musicians must to some extent be aroused by that emotion.<sup>7</sup> So the group simulation may tend to become a real emotion for each individual musician involved. However as I mentioned in chapter five, where the musicians have to concentrate on accurately performing a score, they must to certain extent stay relaxed and emotionally inhibited. Overall if the musicians do not personally identify with the emotional content of the music, that it does not signify something of importance to them, then any synchronisation of their emotions is likely to be limited. Hence it will be when the musicians can use the music to extend the cognition of their actual emotions that we will see the greatest effects. As I argued in chapter five, improvised music is far more likely to enable the extended cognition of emotions. As such it is worth reviewing to what extent the insights gleaned regarding scored performance will carry over to improvised jazz ensembles.

### **Agreeing in Jazz**

The first thing to note about jazz ensembles is that given little or no musical referent the group will lack a significant resource to enable them to think about the music in the same way and thus converge on the music's expressive content. However a

---

<sup>7</sup> To assess the alignment of musician's emotional responses during performance Mitch Waterman (1996) instructed musicians performing a duet to press a button "when the music caused something to happen". Combined with post-performance interviews of the musicians he concluded that their emotional responses were not synchronised. This experiment however was flawed in both the way the responses were measured (i.e. having to push a button at the same time as concentrating on performing the music) and due to the fact that the musicians had no prior rehearsal together. As a result, the performers were far more likely to worry about the technical demands of playing together than to concentrate on the emotional content of the music. At any rate, we should certainly not expect that musicians automatically align in emotional arousal irrespective of the manner in which they attend to the music. Rather, convergence is more likely to occur over several rehearsals as well as over the duration of a musical partnership and according to joint attention during the performance.

common familiarity with the history of jazz, and potentially with other interpretations of the same piece will make up for this to a large extent. In addition, many of the same issues regarding balance of intonation, dynamics and tempo that applied to the string quartet apply equally to the jazz group. For instance, although we can obviously expect the jazz improviser to be freer in the way they time their phrases, they must still be highly sensitive to the overall pulse of the music. Likewise, the group must cooperate to generate an appropriate balance of dynamics and intonation. So in general we can expect that jazz rehearsals (where they occur) will involve negotiation of cognitive directives.

Of course, where a jazz piece involves improvisation, the actual performances will differ from the rehearsals to a significant extent. Hence we can expect that to a far greater degree than the Guarneri quartet, jazz groups must settle musical details during the performance rather than planning them beforehand. Miles Davis for instance was known to avoid analytic discussions of the music or giving too explicit instructions to his players, preferring instead a range of intuitive gestural signals during performance (Smith 1998). Yet as we have seen in the case of scored performance, this need not be a barrier to a genuinely collective interpretation of the music, and may even encourage a greater sense of intimacy between players. So as long as the musicians are geared towards understanding the music in emotional terms, there is no reason they cannot mutually establish a unified emotional interpretation of the music.

In his analysis of improvisational arts, Keith Sawyer identifies several features of group performance that can encourage successful collaborative improvisations. The

most general of these is the point I made above that the musicians should share a common musical background and a shared frame of reference (2003: 54). The second point is a cognitive direction. He says that performers should not 'write scripts in their heads' that is, plan too much where the music will go. Saxophonist and pioneer of free jazz Ornette Coleman describes this very process:

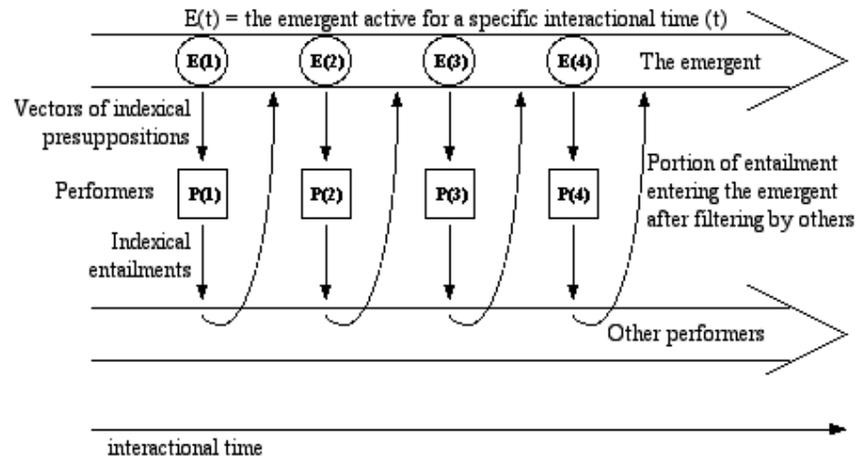
If I don't set a pattern at a given moment, whoever has the dominant ear at that moment can take and do a thing that will change the direction. Our group does not begin with a preconceived notion as to what kind of effect we will achieve. When we record, sometimes I can hardly believe that what I hear when the tape is played back is the playing of my group. I want the members of my group to play what they hear in the piece for themselves. (quoted in Hentoff 1962: 242)

Improvisers avoid too much planning because it can result in trying to force the responses of the other musicians in a direction that might not feel right for them. Alternatively when it fails to go as planned, the original musician may find his flow disrupted (Sawyer 2003: 9). Rather the musicians must be alive to the possibilities of the moment. As I described in chapter five, whenever a musician produces sound, he will thereby generate musical momentum, suggesting ways in which the music can progress. In reference to this Sawyer has three further points: Firstly that although these 'suggestions' will essentially limit the range of possible responses, musicians should try to leave a large range of possible developments (2003: 89). Secondly, the other performers should affirm whatever suggestions are made rather than ignore or deny them, that is, by harmonising with those ideas, imitating or developing them

(2003: 55). Finally, performances can be most successful where they suggest open-ended musical ‘problems’ (such as to make a musically successful transition between two distant keys) rather than a present pre-established ‘solutions’ that must simply be enacted (such as beginning a clichéd sequence) (2003: 105).

As in the case of mutual tracking during scored performances, jazz musicians may in fact have different perspectives on what they are playing, but what is most important is their ability to produce a coherent total sound. It is the collaboratively generated total music product itself that embodies the collective idea of the musicians. In support of this notion, Sawyer argues that for any given contribution a musician makes to the music, the full meaning or impact of that contribution can only be determined by the responses of the other musicians (2003: 6). In addition, the relative freedoms of the performers to develop what each other play mean that the range of possible outcomes multiplies with every moment. These expanding possibilities make it impossible to predict how the music will continue.

As I mentioned in chapter five, this makes the appreciation of what is happening in jazz music to some extent a backward looking process, where the expressive content of the music can only be appreciated in retrospect. Sawyer argues that these features of improvisation make it a prime example of an *emergent* process, where what happens is so reliant on the complex interaction between the parts that the music simply cannot be reduced to a description of what is happening at the level of these parts. Below is a diagram Sawyer provides to describe the emergent nature of improvised performance:



**Group Creativity.** No two-dimensional figure can adequately represent the complexity of real-time interactional this figure should be viewed as one of many possible visual representations. The horizontal axis, time, represents the constantly changing nature of the emergent. No representation of the structure of the emergent is implied or intended. (Sawyer 2003: 87)

Here, each musician listens to the total music product (the emergent), which for him will contextually suggest certain developments. The musician (P1-P4) then contributes his musical idea, which goes through a further stage of being filtered by the other musicians. At this stage the other performers will react musically to either support that idea, develop it, or ignore it. This then is what finally forms the total musical product, in a continual cycle of reactions.<sup>8</sup> This is all completely compatible with my assertion that performing music is a highly blended activity. As jazz musician Bobby Rogovin says, “you can only appreciate jazz if you listen to the whole group. The soloist’s part by itself is just one line in a whole painting” (quoted in Berliner 1994: 387).

<sup>8</sup> Sawyer claims that this process is analogous to the progress of art historically, where new contributions to that artistic field are filtered by other professionals within that field, and as such either denied admission to the field, supported, or developed in various ways (2003: 124).

In so far as this total musical product has expressive qualities then, it embodies a simulated or fictional emotional state. And responsibility for this total musical product can only be ascribed to the group as a whole. Thus the simulated emotion is a collaboratively generated, and collaboratively possessed state. It is not a type which the musicians possess tokens of, but a public representation that they all contribute to.

### **Shared Emotions in Music**

In order for the group's emotional simulation to actually function as a real emotional state it must be properly directed at the group's situation. Yet there are actually relatively few emotionally inducing situations that a group of musicians can jointly focus their attention on. They do not rely on each other for survival, or face common threats the way an early hunter-gatherer society might collectively face drought or war or disease. It may be possible for the group to consider some very general issues such as the threat of global warming or the impact of technology but these are rather abstract, distant concerns. A better option is to use the very act of performance as the source of an emotional response. For instance, the reaction of the audience could be something that the group fear or enjoy collectively. Putting on a good performance may be something that each musician regards as a personal success, but the musicians may also regard it as a success for the ensemble as a whole. They may even have a 'self-conscious' attitude towards their capacity to form a well-integrated ensemble. As Keith Graham says:

It may be just the rapport that is set up in collective playing, the monitoring of what is happening with other players and the adaptation to it and the reciprocity which that engenders which we value. Here it is the perfection of what happens *between* players that we value, rather than what happens to each of them considered individually. (Graham 2002: 124-5)

Having located some potential targets for the group's emotion, we can now address the manner in which the total musical product of the ensemble, which is supposed to embody its emotion, is related to each individual musician's bodily patterns. This is not necessarily a matter of each musician having bodily patterns at the same time as producing the music. Rather it requires that the musician's emotions are *recalibrated* in the right way to allow the music to stand in for those bodily patterns.

In the case of extended cognition that I presented in chapter five, I showed that a musician can use expressive musical patterns to function as his emotional state. In most of these cases the music functions as an *elaboration* of the musician's emotion, standing alongside his ordinary bodily patterns, and helping him to self-reflectively manipulate his overall emotional state. This is possible because musical sounds display a deep isomorphism with bodily patterns and because there are two channels to genuine emotional states; bodily changes and neurally simulated bodily changes. Due to the dynamic properties of music, and the pervasive intermodal capacities of the brain, music can hijack the simulational capacity and thus be treated by the brain as if it was a pattern of bodily changes. So the music stands in the same relation to his emotion as his bodily changes. The musician can then use this simulation in tandem

with his more direct bodily changes, or in some cases, as the musician becomes more absorbed in the flow of the music, the sounds can gradually replace his inner bodily changes as the dominant source for the content of his emotional state.

I noted that in individual cases of extended emotional cognition, changes in the medium of the emotional state (from bodily patterns to musical patterns) must always be a result of recalibration from the more central case of bodily patterns. Hence equally, a group of musicians could not simply choose to employ the music they produce as the group's emotion. Rather, it requires a further extension of the extended cognition of emotion described above. Of course, in a group of musicians there is not just one brain, and hence not just one simulative capacity to treat the music in the same way as emotion. What we have instead is the interaction between the simulative capacities of two or more musicians. This interaction is to be embodied in the music that they collaboratively produce with the help of their simulative processing. Most importantly, their simulative processing must not just target their own particular contribution to the music, but rather the music as a whole.

Yet even if the simulative capacities of the musicians interact using the music, this does not automatically entail that they are having a collective emotional state as opposed to a set of interacting individual emotional states. This is even given that each musician simulates the *total* musical product as the elaboration, or dominant component of his emotional state.

First of all we need to show how interacting in this way forms individually based socially extended emotions. Sue Campbell discusses cases like these when she argues that sometimes I may lack authority over the identity of my emotion (1997: 106-110). For example if I shrug intending to communicate impotence and you interpret this as contempt I might wonder if actually I am inadvertently but sincerely expressing contempt. In general we are not infallible judges of our mental states. Yet this need not be purely a matter of not *knowing* how I feel. Rather the interpretation of the other might actually change what my emotion is, transforming it into a more ambiguous or mixed emotion. This sort of change is especially plausible where our social interaction is what my emotion is about. For instance, I may shrug to express impotence about us losing a team game, which due to your interpretation becomes contempt about the value of the game as well. Thus Campbell says,

If the type of meaning we are trying to communicate is that of the personal importance of a shared occasion, the responses of others to our responses may become a part of the significance of that occasion.  
(Campbell 1997: 117)

This is especially applicable to ensemble performances of music where the very act of trying to achieve ensemble may be the content of the emotion. In addition, Campbell's description is entirely in line with the argument by Sawyer that I presented earlier with regard to collaborative improvisations, where the impact of each musical contribution can only be determined given the reaction of the other performers. Hence although the musical contribution of each individual performer extends his own emotion, this emotion now becomes *socially* extended. This is

because the way his expression is interpreted and developed in response by the other musicians, as they alter the total musical product, actually *changes* the identity of his emotion. In some cases the emotions constituted by the total musical product may be emotions that he could not have outside of that social context. This is for instance, how the following statement from trumpeter Terrence Blanchard seems to be best interpreted:

The amazing thing about playing with Art [Blakey] is that he has a way of tuning into inspiration that can draw an emotion out of you that may have never experienced before.” (quoted in Berliner 367)

Similarly one Chicago jazz musician interviewed by Sawyer claims,

The [ensemble] influences it the same way as if - we're having a conversation now... When I start talking about it, I start thinking about it, putting bits and pieces together, coming up with ideas on how I feel about things, and that way it helps me... when I do it, I'd find that there were these things coming out of myself, which I didn't even know were there. I'd never heard them, I didn't know where they came from... but playing with the others triggers it, so maybe consciously or subconsciously you'll hear that thing that you're trying to find... by listening to what other people have to say, and by talking to them about it, it's like talking about - really great music, it's guys getting together and talking about how sad or lonely they feel, or how happy or angry. (quoted in Sawyer 2003: 29)

Now at this stage it's still the case that other musicians *influence* one's emotional state rather than share in it. But because what (partially) constitutes their emotion is the *total* musical product, other musicians are directly contributing to the identity of that individual emotion. In this way, playing with others can lead an individual musician to have an emotion that is unique to that social context (cf. Wilson 2004: 299-302). So it's a socially extended emotional state.

From the point of view of the individual musician, the main difference between individually extended cognition using music and socially extended cognition using music is that he must react to and help to shape the *total* musical product rather than just the sounds that he produces on his own. This means that each individual musician can only have partial control over the development of his own emotion, which may undermine his sense that the music is an emotion that belongs to him as an individual. Then again, we rarely have complete control over our everyday emotions. So as long as the simulational capacity of that individual is directed towards the total musical product, and as long as his expressive acts are directed towards shaping that total musical product, there is nothing to prevent the musician from identifying the total musical product as his own emotional state.

So if the other musicians involved in the ensemble performance equally extend their emotions in this same manner, what we have is a case of *overlapping* emotions. The musicians dip into a common pool of music/emotional material to extend their emotions. Yet this same material could help to constitute *several different* emotional states. To explain: the musicians share the material for their emotions (the body) but they each remain the locus of control for what this emotional state represents.

Accordingly, they could each take the total music (in combination with their own bodily feelings) to emotionally represent some fact about their own lives. One person could regard the music as reflecting on the troubles of his home life, where another regards the music as reflecting on the hostile reaction of the audience.

However given this setup it is actually rather easy to switch from several overlapping socially extended emotions to one single shared or group emotion. Now that some of the intrinsic content of the musicians' emotional states has been embodied in the music we can now appeal to the mechanisms that Margaret Gilbert outlined for collective emotions and its corresponding basis in joint attention. All that is required is that the musicians *agree* to have an emotion as a body, to represent some situation that impacts on the group. The music can then define the agreement of the musicians. It becomes the locus of control for the emotional state. Like a written agreement in a collective intention, the music has the final word concerning what the content of the emotion is. Similarly, the agreement defines a group subject to which all the musicians must defer. As a result, the entire system comprising the patterns in the music and the bodily patterns of each musician realises a single coordinated emotion.

We can then see how this case satisfies my two conditions for being a mental state; that the system plays the normal causal role of an emotion and that it has distinctive representational content, *independently* of any *additional* internal states of the musicians involved:

The fundamental difference between the group emotion and the individual emotional states contributing to it is that the total system speaks authoritatively for the state of

the group in a way that no individually extended emotion can. This is a distinction both in terms of cause and content. The variety of other distinctive features of this mental state are then derived from this fundamental characteristic. For instance, the group state is obviously independently capable of expressing the emotion of the group to outside observers. Similarly, the players involved can also recognise the emotion of the group without having to engage in any additional empathic projects.

Furthermore, the system could have behavioural consequences on the group's actions. The musicians may stop playing if the emotion embodied in the music is too upsetting, or they may intensify their activities if it is joyful. In general the content of the music encourages the group to reflectively manipulate its emotional character just as in cases of individual extended cognition using music, where the music is used to reflect on, and control one's own emotional state. These behaviours are of course part of the activity of generating the shared emotion in the first place. Yet they are still occurring for *emotional* rather than purely musical reasons. Finally the system could have long-term effects on the solidarity of the group. Music can play a quite unique role in generating emotionally bonded communities, which as I noted in the introduction to this thesis, may be considered an evolutionary function of music.

As regards the distinctive content of the system, we can see that it may differ from the music produced by each individual in terms of vividness, intensity, temporal duration or complexity of emotional character. Yet it is also worth exploring the possibility that the emotion of the group may be of a contrasting emotional *type* to the emotions of the individuals involved (though it nonetheless supervenes on the emotional states of these individuals).

For example, in her analysis of group emotions, Margaret Gilbert argues that collective remorse need not involve personal remorse. I might be able to sincerely say ‘we are full of remorse’ whilst not having worked it out for myself yet. In addition, some group activities allow there to be non-operative members, such as when a member of a committee abstains from a vote. Yet group expressions of emotion don’t seem to leave room for non-committed members in quite these ways. As I argued in the case of collective intentions, group mental states are supervenient on extended individual mental states. In setting up the total musical product as an emotional state, each musician must contribute their simulative processing, which interacts with the music produced by the others to fix the content of the group emotion.

However, as we know from cases of empathy, one of the main functions of this simulative capacity is to detect the emotions of *other people*, not one’s own emotion. It even seems possible to simulate one emotional state (say the anger of another person) whilst simultaneously feeling fear for oneself (generated more directly by one’s bodily reactions). Hence it is possible for musicians to be involved in helping to generate the emotion of the group whilst simultaneously having their own private emotion, or just adopting the attitude that they do not identify with the emotion that they simulate. They could still jointly attend with the other musicians to the situation of the group and otherwise express a commitment to the group’s situation. Yet they could just not regard the emotion of the group as signifying anything of personal significance for them. We would not call this a shared emotion then so much as a ‘group only’ emotion.

These cases are unlikely to encourage the *sense* of agreeing with others that I report in the following section, yet the sense of agreement is a conscious experience. Group emotions do not require the consciousness of the performers to exist, only that a recalibration of their individual emotional patterns takes place, allowing the musicians to commit to the music as the emotion of the group. Nor do performers have to be playing at every moment. Though stopping will tend to separate the individual's emotion from the group's emotion, a legitimate contribution to the total musical product may be to occasionally leave space for the other musicians. One might even imagine a situation where joint deathly silence seems to the group like an appropriate emotional representation of the situation.

It is also worth considering situations in which the musicians may each individually feel that they are struggling to get it together, but an outsider could perceive this as the group struggling with itself, as if it were having a single self-conflicted emotion. The individual band members won't feel self-conflicted, they will just feel plain conflict (with the other musicians). Of course, the feeling of conflict with another person can be quite different to the feeling of self-conflict. Where conflict can be outward looking, even exuberant, self-conflict is typically inward looking, complex, ambivalent, fluctuating this way and that emotionally.

But could this really be a case of group only emotion, or merely something that an outsider might project onto what they are hearing? This will depend on whether the musicians satisfy the general conditions for group emotions. They must agree (either explicitly or conventionally) that the conflicted music emotionally represents the

situation of the group. So the group could ironically agree that the conflicted musical product expresses a genuine (group) self-conflicted emotion. The musicians may then contribute to this group emotion by each contributing clashing expressions of emotion. Moreover they may each really believe that their specific contribution helps to shape an appropriate representation of the group's situation. So they are not merely *pretending* to have a group emotion.

Overall, it seems possible for groups to express all kinds of emotions as a body, and play contrasting roles in bringing that group emotion to fruition. The main limitation is that because the musicians must concentrate very closely on the total musical product and use that to guide their emotional cognitions, they are far more likely to be infected by the group's emotion. Generating a group emotion need not demand that one has a similar emotion oneself. Yet sincere emotional commitment to one's musical productions greatly facilitates the generation of the group emotion.

### **The Phenomenology of Sharing Emotions**

We find some limited confirmation of the possibility of shared emotions when we look at the descriptions that musicians provide of their own experiences of coordinating with others. Clearly no amount of *individual* reports (even if there were measured in more controlled circumstances) could confirm a *shared* emotion. Nevertheless, we would still expect it to feel like *something* for the individuals involved. Now we should note that some of these reports could be taken to imply that there is a single token phenomenal experience that everyone shares. Yet I am *not* arguing for this possibility here. Rather, I am simply interested in showing that individual musicians have quite remarkable experiences when they take part in

generating shared emotions. This should help to confirm that sharing an emotion is a real possibility, because we would expect it to feel quite unusual to engage in such a state.

Although the cases presented in the published literature are rare, I have heard similar reports from most musicians I know. We should also note that musicians find the experience hard to describe. If they had the kind of conceptual apparatus that I have given in this thesis, they should hopefully find that this captures what they mean more exactly. However, what descriptions can be found are extremely compelling. First of all, we have this description from sociologist and jazz musician William Cameron:

If this collective enterprise succeeds, each feels the full warm response that comes from the wholehearted cooperation of the group. Indeed, such feeling is greatly intensified, because the immediacy of expression possible for a musician who has command of his instrument provides a more profound emotional release than almost any other kind of activity. It certainly provides the supreme emotional experience for the jazzman.  
(Cameron 1963: 124)

This sense of wholehearted cooperation is put in stronger terms by free jazz saxophonist Tim Berne:

... And I think I just decided it was time for me to rely more on my playing and then I just said you know “we’re all gonna improvise and just

open it up” because I think you know that’s where the magic happens because when you do all kind of *sense* that you’re agreeing on something, even though you’re not playing the same thing- it’s the most amazing feeling... (Tim Berne quoted in Marsalis 2004)

I claim that this ‘sense of agreeing on something’ is the main phenomenal characteristic (for individuals) of the experience of sharing emotion with others.

The other main variable is the sense of losing individual control over what is happening, which is most associated with the idea of *group flow*. Group flow is similar to the case of individual flow that I introduced in chapter five. It represents a peak of musical performance in which the musicians involved are absorbed in the momentum of the music. For instance saxophonist Franklin Gordon says,

Every jazz musician wants to be locked in that groove where you can’t escape the tempo... You’re locked in so comfortably that there’s no way you can break outside of it, and everyone’s locked in there together. It doesn’t happen to groups every single night, even though they may be swinging on every single tune. But at some point when the band is playing and everyone gets locked in together, it’s special for the musicians and for the aware, conscientious listener. These are the magical moments, the best moments in jazz. (quoted in Berliner 1994: 388)

Similarly Jerry Garcia of rock band ‘The Grateful Dead’ (who often performed extended improvisational concerts) has said,

In The Grateful Dead when we're playing very open with no structure, sometimes the sound level can speed a sensory overload of a kind which starts to become a physical experience rather than a musical one and that also has a certain kind of value... It's sort of stumbling into this area where there's a lot of energy and something happening and not a lot of control. So that the sense of individual control disappears and you are working at another level entirely. Sometimes this feels to me as though you don't really have to think about what's happening. Things just flow. It's kind of hard to report on but it's a real thing. I mean we've checked it out with each other and after twenty-five years of exploring some of these outer limits of musical weirdness this is stuff that we pretty much understand intuitively but we don't have language to talk about it. But it's reported back to us by people in the audience too so this is one of those things where we're sort of collecting data without really knowing quite where it's leading or what it's about but we feel a certain custodian relationship to it. (quoted in Bailey 1992: 42-43)

Similarly Trombonist Melba Liston says:

I don't know if I can describe it, but I know it when I feel it. Just one night, everybody can feel what each other is thinking and everything. You breathe together, you swell together, you just do everything together, and a different aura comes over the room. (quoted in Berliner 1994: 392)

Collective improvisation teacher Charles Ford attempts to articulate this collective phenomenon in more detail:

When collective freedom finds its voice in musical improvisation, the relationship between individual and collective becomes a static, though modulating unity. Individual freedom may well be lost, but what is promised is the most extraordinary union of minds in music, a union that dissolves and assumes ethics, pleasure and aesthetic experience into itself. (Ford 1995: section 2)

Finally one of the most compelling descriptions comes from jazz singer Carmen Lundy:

It's that freedom of expression and expressiveness that comes through from a feeling you have a musical rapport with other people. It's something that you really can't touch, but you know when you are sharing it with another musician. That's the same thing that I shared with the person next to me when everybody was participating in the service [in reference to gospel singing]. I can remember some unbelievable things from that time which I experience even now when I sing jazz. Sometimes, I really feel that I am just the vehicle, the body, and that something is really something through me, like I am not controlling everything that I am singing. The last time I sang, I thought to myself, "Gosh, I feel like something is just singing through me," That's what I mean by the spiritual thing. (quoted in Berliner 1994: 392)

In chapter five, I mentioned that individual flow states may be characterised by the absence of self-analysis or self-predicated thoughts. Since these thoughts will be private to each musician, losing them will enable the musician to concentrate more on the group representational state. Yet this absence would not make the group state more cohesive. It would just allow the individual involved to appreciate that group cohesion more.

However, I also argued in chapter five that this sense of flow is most likely caused when the intentions of the musician become better aligned with the momentum of the music. In emotional terms, the intentions of the musician correspond to his inner (and so individual) bodily feelings. Hence if these bodily changes have become perfectly aligned with the emotional momentum of the total music product, then it would indeed indicate a more intense coordination of the group on an emotional representation. This is because the private bodily changes of the musician have either become dominated by the patterns in the music (so they don't contribute much distinctive content to the emotional state) or, they have become directly expressed in the music. Either way, this represents a smoother alignment of the emotional system as a whole and a sense in which the music plays a greater role in generating and maintaining the intrinsic content of the shared emotional state.

This would not make the flow a group state however. In this case, individual flow has helped to align the group state. Rather *group* flow just seems to refer to a more perfect coordination of the musicians' activities. Perhaps the music can generate such a powerful momentum, a particularly driving rhythm, or a particularly intense

emotional character, that the individuals involved can do nothing but submit to it. So the sheer momentum of the group activity may cause the individuals involved to let go of their individual musical and emotional intentions and *simultaneously* experience coordinated flow states. We might call this contagious flow.

Keith Sawyer argues that group flow is more likely where the extent to which an extrinsic goal of the performance is specified is proportional to the amount of pre-existing structures the performers share, such as scored parts, conventions, or interactional routines (Sawyer 2003: 167). In jazz ensembles the lack of pre-arranged material may correspond to the lack of a specific musical goal. But group flow is equally possible for scored ensembles, where all the pre-planned materials required are available to produce a highly specific musical product.

For instance, the Guarneri quartet also describes exceptional cases of mutual coordination. Dalley says for example, “it’s true that there are occasions when we really do seem to feel and breathe as one player.” Steinhardt agrees, “when a performance takes flight I feel as if all four personalities meet somewhere in the air - maybe two and a half feet above the quartet.” In reference to Beethoven’s ‘Heiliger Dankgesang’, a hymnlike movement within the Opus 132 that demands especially blended performances, John Dalley claims that on one occasion:

I only felt that we were coming close to the essence of what Beethoven had in mind when wrote the piece: that kind of hymnlike obeisance to a higher power... I felt that all of us shared the same sense of something special taking place. (Dalley in Blum 1986: 169)

David Soyer agrees:

Of course, by its very nature such a movement demands that we have a completely homogenous sound. But as John says, it's sometimes more than our attempt to make it so: everything becomes concerted and blended and propelled as if by itself. The music seems to take over.  
(Soyer in Blum 1986: 169)

Again these descriptions are suggestive of the group sharing a single token conscious experience. I make no claims about the possibility of sharing the phenomenal experience of music here, though we should note that it could only be their *emotional* experience rather than their *auditory* experience that is shared. At any rate, we can see that musicians can feel an especially intimate connection with the music and by extension with each other. In states of group flow, the momentum of the music dominates the shared emotional state.

Now it should be noted that the descriptions I quoted do not focus purely on emotional feelings. However the expressive qualities of music are generally a given amongst musicians. Also it might not always feel appropriate to describe the emotions expressed in music in terms of everyday emotions like anger or sadness, yet a sense of the music capturing the personal significance of the situation, the sense of dynamic interaction with the world, the purpose and value of life being lived, are all fundamentally emotional states. Moreover a feeling of intense group coordination will almost by definition be a highly ecstatic or euphoric feeling with profound

impact on the individual's sense of life afterwards. For instance Trumpeter Herb Pomeroy has said:

It's an incredibly warm feeling that you have, one that you've shared with the other musicians and you've shared with the audience. And when the evening's engagement is over, you still retain it. It fills you up inside, and you feel it like there an aura all around you when you leave the club to go home. It's the kind of precious feeling that no other kind of career can give you. (quoted in Berliner 1994: 394)

### **How to Share an Emotion**

So if the experience of absorption into the shared emotion is so blissful, why don't performers get into this state more often? And what factors make it more likely to occur? Some musician's descriptions suggest that it is purely a matter of serendipity whether an ensemble manages to achieve group flow and thus a particularly intense sense of generating a shared emotion. Yet I think that understanding the nature of this group state allows us to identify several factors which will encourage shared emotions (again, I will focus on improvised performances here).

The first thing to note is that the group must collectively commit to the music as the emotional representation of their situation. So the musicians should all be mutually aware of this goal, and should pay continual joint attention during performance to the expressive nuances of the music and the emotional impact of the shared situation. Yet although collective improvisation *can* be guided by ideals of group convergence, it is certainly not the case that every group improvised performance has emotional

unity as its goal. Sometimes musicians may want to surprise or challenge one another or simply to explore the possibilities of the music.

However, let us suppose that the group has agreed to try and express an emotion ‘as a body’. Moreover, the greatest level of integration is achieved by reaching a state of group flow. So recall Sawyer’s claim that group flow is more likely where the specificity of the extrinsic goal is proportional to the amount of pre-existing structures the performers share. In the case we are focusing on here, the extrinsic goal is to emotionally represent the situation. Since the *sincerity* of this emotion, that it is a live response to a real situation, precludes planning it too explicitly, the pre-arranged material should not dictate any particular emotional response. At the same time however, since maximal coordination is desired, material should be provided that enables the musicians to interact and cohere fairly easily. These two factors are in tension with each other, so it is likely to be quite difficult to achieve the right balance between spontaneity and collectivity, between individual expressive freedom and submission to the emotion of the group. However, it is at least more likely to occur where these goals are deliberately sought, and this historically has been a common purpose within the free jazz movement. For instance musicologist James Collier writes of free jazz,

For some musicians, the whole point of the music was no longer the reaction of the audience, but the emotional interaction of the musicians themselves. It came to be a sign of success when a musician could say, “There was a lot of love going on up on the stand tonight.” Too much

soloing, said many of these players, smacked of self-aggrandizement...

(Collier 1978: 474-5)

As Collier notes however, “the free-jazz players were further hampered by an ideology that was at cross-purposes with itself, espousing both individualism and collectivism at the same time” (Collier 1978: 475).

Yet despite the fact that the goals of both collective and individual expression are often in tension, they need not be mutually exclusive. Some free jazz musicians have even attempted to articulate a synthesis between them. For example Lou Gare, a saxophonist with the free jazz improvisation group AMM has said: “It’s not the freedom to do anything you like. It’s the freedom to do what the music likes. And what the music likes happens to be what you like as well” (Gare quoted in Richards 1992: 65). So if the individual musician closely identifies with the music and by extension with the group, then perhaps the sincere expression of the group and himself may all amount to the same thing.

Yet no matter how it is interpreted, there is still a definite sense in which individual freedom as we commonly understand it is undermined. For instance saxophonist Evan Parker has said,

What makes (free music) relevant is that it’s a group activity. You have to look for this other organism, which is the group mind. You block access to that group mind if (your own) personality is too strong. (quoted in Smith & Dean 1997: 63)

Remember that contributing one's simulative activities to the group expression is greatly facilitated by being sincerely committed to that emotional activity. So the principle difficulty in achieving a shared emotion will be precisely the need to make a complete and sincere commitment to the emotion of the group. Since jazz music typically values virtuosic solo performances, and since our culture in general emphasises individual excellence, we can expect that this level of commitment will be hard to achieve. In addition, as dedicated musicians, jazz performers are also likely to be highly sensitive to their own and each other's mistakes. As a virtuosic form, jazz performance is also liable to be both technically very difficult, and highly competitive. These factors make it less likely that the performers can coordinate on generating group emotions.

Yet these limitations are not impossible to overcome. Sincere commitment to the group will be most easy where the musicians already share similar backgrounds and compatible personalities (cf. Sawyer 2003: 49 & Berliner 1994: 395). On a technical level as well, the musicians will collaborate most easily where they share common frames of reference in musical styles as well as similar levels of musical ability. Groups that rehearse and stay together over longer periods will also be more able to anticipate and react to each other's musical ideas, as well as share similar musical backgrounds. Furthermore, a long-standing group is more likely to provide a musically safe and supportive environment, allowing musicians to emotionally commit to their performances without worrying too much about the critical evaluations of others. For example Sam Richards (part of the avant-garde

improvisation group AMM) describes the advantage of this kind of situation, explicitly referencing the idea of a group mind:

[T]he individual was free to play whatever he was driven to knowing that the others were doing likewise. The permission given to each other to do this is what, paradoxically, produces a 'group mind'. This, at times, seems to function autonomously. Thus the space becomes safe despite its ability to sound abrasive or even violent. Prévost and Rowe have written in a sleeve note: 'The players could share a timeless immersion in a world of sound, while simultaneously being free to pursue their individual paths. It was not uncommon for the musician to wonder who or what was producing a particular sound, stop playing, and discover that it was he himself who had been responsible.' (Richards 1992: 64)<sup>9</sup>

When we get down to the details of the music itself, although musicians may wish to push themselves technically, music that is less technically demanding, or which only requires the musician to play at a level that suits them, is more likely to allow freedom from worries that distract from emotional expression, as well as to more easily enable coordination. With this in mind, I have in the appendix provided an example score that aims to maximally satisfy the various factors I have identified above. The principle goal of this score is to provide material that allows the musicians to cohere fairly easily, is emotionally nuanced though allowing for a range

---

<sup>9</sup> I find this idea that the musicians involved in a collective improvisation can lose track of their own contribution the music to be a striking and surprising effect (it also compares interestingly to the statement by Ornette Coleman that I quoted in the section on agreeing in jazz). The best way to understand it seems to be that by becoming absorbed in the group emotion, the musician stops attending to his own particular experience of generating the total musical product.

of expressive characters, and which is flexible enough to allow the musicians to react to one another spontaneously. In addition there are explicit instructions to treat the total musical product in emotional terms and various suggestions for how the material can be rehearsed and developed. The accompanying recording also provides an example of a jazz trio (myself and two friends) collectively improvising according to the rules specified by the score (tracks 2-5).

As such, although it is admittedly rather limited, there is potential to test whether group emotions can be deliberately invoked. However there will always be difficulty in verifying whether the musicians really have generated a shared emotion. This is because whilst we should hopefully find that individual subjective reports as well as physical measurements of their emotional responses would align, this is not the same thing as the emotional response of the group as a whole. It would be like measuring the different parts of a brain coordinating to produce a single mental state and expecting them to match. Even if they did match it need only show that they were in similar though numerically distinct states. In the end the best way to appreciate an emotional state is from a first person perspective, or failing that, to directly empathise with it. So although we might observe some alignments between members of the group, or observe resulting group behaviours or bonding, the best way to observe a group emotion is to listen to the music that embodies and expresses it.

## Conclusion

In this thesis I have explored and developed several general philosophical ideas: First of all I argued that emotions are primarily perceptual states, constituted by bodily changes and the bodily patterns that organise and register them. However, I also argued that emotional perceptions are of a more complex and evaluative nature than vision or hearing. This evaluative character is mostly a reflection of what it is that emotions perceive, which I have described as the dynamic relation between self and world. Emotional perceptions involve a sense of the self's changing circumstances, its integrity and powers, and the affordances that the world offers for directed action. Given this focus on the self, and given also their evolutionarily ancient nature, emotions are likely to be an important factor in the development of a reflective sense of personal identity in the first place. At the very least, emotions come in at the 'ground floor' of self.

In the second chapter of this thesis I then extended this theory of emotions further; arguing that primary expressive behaviours should be construed as constitutive parts of emotional states, since they help to generate the bodily pattern. This is what underlines their communicative function. It also entails that we can cognitively manipulate our emotional states by means of expressive behaviour. Establishing the nature of emotional expressions then allowed me to show that when we recognise the emotions of others, we get a direct sense of the way it feels for them by recreatively forming neural simulations of their aroused states, directly accessible to us by perceiving their expressive behaviour. This basic form of emotion recognition is often foundational to more sophisticated acts of empathy.

I also made general claims about the nature of music; in particular that the patterns embodied in music display a deep isomorphism with emotions, and that we may view the expressive capacities of music as real dispositional properties. So music is a pre-eminent form of artistic expression of emotions because it vividly captures the most intrinsic character and content of emotional states. Yet one of the central problems in the philosophy of music is accounting for how music may express emotional *experiences* despite not having a psychology. I argued that our ability to recognise emotions in music is a result of music hijacking our simulative capacity. Ironically, given that musicians may use music to extend their cognition of emotions, it may turn out that sometimes the music really *does* have a psychology (at least partially).

This brings us to most important general idea of this thesis; that the mind is often partly physically constituted by its interactions with the environment and other people. In particular I argued that a musician could use the music he produces to extend the cognition of his own emotional state. Here the music either elaborated upon his bodily changes or played a dominant role in generating his bodily pattern. Moreover, just as ordinary bodily changes are constitutive of emotions, so also musical changes can be constitutive of emotions. Within these cycles or systems of interaction whereby the emotion is generated, it is unreasonable to locate any final state in the brain as *the* mental state. Certainly the mind *can* function independently of the environment, as when we dream or imagine, but that is more likely to be a sophisticated feature of *human* minds rather than an essential feature of all minds, all the time.

Since things outside the brain may constitutively contribute to the character of a mental state, it entails that the mind is not essentially dependent on neural processes. Nor is the mind necessarily private to each person who has one. Rather it is possible for several minds to be integrated with one another in a variety of ways. All that is required is a medium of interaction that permits the contents of mental states to be fluently manipulated and coherently blended together.

In order to justify this point, I had to gradually develop an account of how we can integrate our mental states from the most basic level. First of all, I argued that the simple act of joint attention allows our perceptual activities to be interdependently structured. This behaviour, almost unique as well universal amongst humans, provides a foundation for more radical interactions. It defines a group subject that can allow people to mutually fix the character of objects and share their responses to those objects. The case of collective intentions is then built upon this foundation, allowing the intrinsic content and causal powers of a mental state to be collectively realised. I then returned to the case of the musical expression of emotion. I showed that musicians may use their total musical productions to constitute an emotional reaction of the group.

Due to its expressive potential as well as its capacity to allow the coherent blending of individual activities, music is inherently suitable for the sharing of emotional states. However shared emotions in music may be just one of a variety of possible shared emotional states. Scenes of mass rallies, mob riots, the roaring football crowd or the singing congregation also suggest the potential for shared emotions. Where

large crowds of people are mutually focused on a target (scoring a goal, the worshipped dictator) and mutually express their attitude towards that target, then it may be the case that a shared emotion is generated. This can be a terrifying as well as exhilarating prospect.

The key difference between shared emotion in music and these other possible cases is that music is almost certainly more likely to allow more sophisticated emotional reactions. It is claimed for instance (e.g. McDougall (1920), Le Bon (1896)) that mob emotions encourage the lowest common denominator of emotional type. Though the force and range of the mob emotion may be greatly extended, the fluency and complexity of the emotional interaction afforded by music is far greater than that afforded by ordinary emotional expression. Shared emotions in music may genuinely encourage an *enhancement* of our emotional cognitions, to bring out emotions that we have never experienced before. Most of all however, it can provide a profound sense of reaching beyond the normal boundaries of the self.

# Appendix

## Experiment in Shared Emotions No.1

**PIANO** **SAXOPHONE** **1<sup>st</sup> Movement: Uncertain**



The first movement section contains three musical staves. On the left, a grand piano staff is shown with a treble clef on the upper staff and a bass clef on the lower staff. To its right is a saxophone staff with a treble clef. Below the saxophone staff is a drum staff. All three staves are currently empty.

**PIANO** **SAXOPHONE** **2<sup>nd</sup> Movement: More Intense**



The second movement section contains three musical staves. On the left, a grand piano staff is shown with a treble clef on the upper staff and a bass clef on the lower staff. To its right is a saxophone staff with a treble clef. Below the saxophone staff is a drum staff. All three staves are currently empty.

**PIANO** **SAXOPHONE** **3<sup>rd</sup> Movement: More Relaxed**



The third movement section contains three musical staves. On the left, a grand piano staff is shown with a treble clef on the upper staff and a bass clef on the lower staff. To its right is a saxophone staff with a treble clef. Below the saxophone staff is a drum staff. All three staves are currently empty.

# **Experiment in Shared Emotions No. 1**

## **Guide to performers**

### **1. Goal**

The purpose of this piece is to enable performers to generate a shared emotion. This shared emotion is to be physically constituted by the music the performers collectively produce and its means of production (including the brains of the musicians). The goal of sharing emotion should be prioritised over the quality, originality or technical virtuosity of the music.

### **2. Basic Rules**

**2.1.** Everyone must agree to use the music to express the situation that confronts the group.

**2.2.** Each performer must always use his or her sense of feeling to generate sounds.

**2.3.** Each performer must pay close attention to the expressive qualities of the music produced by others and the group as a whole.

### **3. Set up**

**3.1.** It is conducive to the goals of this piece if the musicians involved share similar cultural backgrounds or attitudes, stylistic preferences, and compatible levels of technical proficiency. The musicians should hopefully be friends prior to approaching the piece or become friends during the rehearsal process.

**3.2.** It is recommended that performers warm up by engaging in cooperative forms of play (e.g. telling a story or describing a situation, taking turns to say one word each).

**3.3.** Performers may also warm up by trying to imitatively copy themes from one to the other, or coordinating their timing by gradually speeding up or slowing down on a single note.

**3.4.** Performers should spend time listening to each others' solo performance. They should mutually encourage a feeling of confidence in the music they produce.

**3.5.** It is recommended that performers spend a significant amount of time rehearsing together and exploring the emotional content of the score, so that they are completely familiar and comfortable with the expressive possibilities of the music.

**3.6.** In order to use the music to agree on the emotional impact of the situation, it is helpful to make a statement to this effect before starting, perhaps using a ritualistic formulation such as:

A: "A, do you agree to express the situation that confronts us?"

B: "I do so agree it!"

**3.7.** Performers may also pre-arrange certain signals or gestures that indicate agreement or not on the expressive content of the music.

**3.8.** The situation that confronts the group may include the physical surroundings, any audience response and even their ability to produce a coherent piece of music.

Alternatively performers can agree on a particular reference that impacts them all (such as global warming). Performers may wish to reflect on their feelings about the situation for a few minutes prior to performance (either silently or in conversation).

#### **4. The Score**

**4.1.** The purpose of the score is to stimulate the expressive imagination of the performers without dictating too much any particular emotional state.

**4.2.** The score should enable musicians of varying ability to play together without having to worry about the technical demands of the music.

**4.3.** As such, performers can treat the scored material *very* freely. In general the performers should treat the score as a starting point for more general improvisation.

**4.4.** In particular performers can skip or add notes, augment intervals, alter tempos and rhythms, repeat and loop material and so on according to what feels natural at the time. They may use material written for other instruments, or from any part of the score at any time.

**4.5.** For ease of coordination however, performers are encouraged to concentrate on the part scored for their instrument and on one movement at a time.

**4.6.** Performers can prepare details, such as themes, or ways in which the material can be looped. They should not, however, plan the overall progress of the music but should react immediately to the unique circumstances of the performance event.

## **5. During performance**

**5.1.** Although each performer must always use his or her own *sense* of feeling to generate sounds, the performer need not personally *identify* with those feelings. However it is strongly suggested that the performer should *only* play sounds that they can sincerely commit to as expressions of their actual occurrent feeling.

**5.2.** Performers need not play all the time. But if they feel the music does not accurately reflect the situation that confronts the group, they should try to change it.

**5.3.** Whilst it is expected that musicians will be heavily influenced by a stylistic genre (e.g. modal jazz), musicians should not allow the clichés of that genre to dictate what they play.

**5.4.** It is suggested that performers begin by taking turns, i.e. one musician plays a couple of phrases and other musicians then answer. This dialogue may then become more overlapping.

**5.5.** It is suggested that performers start very simply, perhaps by holding or repeating a single note.

**5.6.** It is suggested that performers imitate and harmonise with each other.

**5.7.** It is suggested that performers exchange eye contact frequently throughout performance.

## Bibliography

- Aou, S., Oomura, Y., Hishino, H., Ono, T., Yambe, K., Sikdar, S. K., Noda, T. & Inoue, M. (1983). 'Functional heterogeneity of single neuronal activity in the monkey dorsolateral prefrontal cortex'. *Brain Research* 260: 121-124.
- Adachi, M., Trehub, S. E. & Abe, J. (2004). 'Perceiving emotion in children's songs across age and culture'. *Japanese Psychological Research* Vol. 46, No. 4: 322-336.
- Adorno, T. W. (2002). 'The Radio Symphony'. Printed in *Essays on Music*. Leppert. R. ed. University of California Press, London.
- Asendorpf, J. B. (2002) 'Self-awareness, other-awareness, and secondary representation'. In Meltzoff & Prinz ed.: 63-73.
- Austin, J. L. (1979). 'Other Minds'. In *Philosophical Papers*. Oxford University Press, Oxford: 76-116.
- Bailey, D. (1992). *Improvisation: Its nature and practice in music*. The British Library Sound Archive, London.
- Baldwin, D. (1995). 'Understanding the Link Between Joint Attention and Language'. Printed in Moore & Dunham ed.: 131-158.
- Balkwill, L-L. & Thompson, W. F. (1999). 'A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues'. *Music Perception* Vol. 17, No. 1: 43-64.
- Balkwill, L-L. Thompson, W. F. & Matsunaga, R. (2004). 'Recognition of emotion in Japanese, Western, and Hindustānī music by Japanese listeners'. *Japanese Psychological Research* Vol. 46, No. 4: 337-349.
- Baron-Cohen, S. (1995). 'The Eye Detection Detector (EDD) and the Shared Attention Mechanism (SAM): Two Cases for Evolutionary Psychology'. In Moore & Dunham ed.: 41-60.
- Baron-Cohen, S., Tager-Flusberg, H. & Cohen, D. (1996). *Understanding Other Minds: Perspectives from Autism*. Oxford University Press, Oxford.
- Benjamin, W. (1968) [1935]. 'The Work of Art in the Age of Mechanical Reproduction'. In *Illuminations*. Schocken Books, New York: 217-251.
- Berliner, P. (1994). *Thinking in Jazz: The infinite art of improvisation*. University of Chicago Press, London.
- Blacking, J. (1976). *How Musical is Man?* Faber and Faber Ltd. London.
- Blacking, J. ed. (1977). *Anthropology of the body*. Academic Press, London.

Blood, A. J. & Zatorre, R. (2001). 'Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion'. *Proceedings of the National Academy of Sciences of the United States of America* Vol. 98, No. 20: 11818-11823.

Blum, D. (1986). *The Art of Quartet Playing: The Guarneri quartet in conversation with David Blum*. Victor Gollancz Ltd. London.

Bogdan, R. ed. (1986). *Belief: Form, content and function*. Oxford University Press, Oxford.

Bratman, M. (1993). 'Shared Intention'. *Ethics* Vol. 104, No. 1 (October 1993): 97-113.

Bratman, M. (1999). *Faces of Intention: Selected Essays on Intention and Agency*. Cambridge University Press, Cambridge.

Bruner, J. (1995). 'From Joint Attention to the Meeting of Minds: An Introduction'. In Moore & Dunham ed.: 1-14.

Buck, R. (1980). 'Nonverbal behaviour and the theory of emotion: The facial feedback hypothesis'. *Journal of Personality and Social Psychology* 38, No. 5: 811-824.

Budd, M. (1985). *Music and the Emotions: The Philosophical Theories*. Routledge & Kegan Paul, London.

Budd, M. (1995). *Values of Art. Pictures, Poetry and Music*. Penguin Books, London.

Butterworth, G. (1995). 'Origins of Mind in Perception and Action'. Printed in Moore & Dunham ed.: 29-40.

Cacioppo, J. T., Uchino, B. N., Crites, S. L., Snyder-Smith, M. A., Smith, G., & Berntson, G. G. (1992). 'Relationship between facial expressiveness and sympathetic activation in emotion: A critical review, with emphasis on modelling underlying mechanisms and individual differences'. *Journal of Personality and Social Psychology* 42, No. 1: 110-128.

Call, J. & Tomasello, M. (2005). 'What chimpanzees know about seeing, revisited: An explanation of the third kind'. Printed in Eilan, Hoerl, McCormack & Roessler ed.: 45-64.

Cameron, W. B. (1963). 'Sociological Notes on the Jam Session'. Printed in *Informal Sociology: A Casual Introduction to Sociological Thinking*. Random House, New York: 118-130.

Campbell, J. (2002). *Reference and Consciousness*. Clarendon Press, Oxford.

- Campbell, J. (2005). 'Joint Attention and Common Knowledge'. Printed in Eilan, Hoerl, McCormack & Roessler ed.: 287-297.
- Chanan, M. (2000). *Repeated Takes: A short history of recording and its effects on music*. Verso, London.
- Charlan, L. C. (1997). 'Reconciling cognitive and perceptual theories of emotion: A representational proposal'. *Philosophy of Science*, Vol. 64, No. 4 (Dec. 1997): 555-579.
- Clark, A. (1997). *Being there: putting brain, body, and world together again*. MIT Press. Cambridge, Mass.
- Clark, A. & Chalmers, D. (1998). 'The Extended Mind'. *Analysis* 58: 10-23.
- Clark, M. S. ed. (1992). *Review of personality and social psychology: Vol. 13, Emotion*. Sage, Newbury Park, California.
- Clarke, E. (2002). 'Listening to Performance'. Chapter 13 in Rink ed.: 185-196.
- Clynes, M. (1977). *Sentics: The Touch of Emotions*. Anchor Press & Doubleday, New York.
- Collier, J. L. (1978). *The Making of Jazz: A Comprehensive History*. MacMillan, London.
- Collingwood, R. G. (1958). *The Principles of Art*. Oxford University Press, London.
- Cook, N. (1990). *Music, imagination and culture*. Clarendon Press, Oxford.
- Cook, N. (1998). *Analysing Musical Multimedia*. Clarendon Press, Oxford.
- Cooke, M. & Horn, D ed. (2002). *The Cambridge Companion to Jazz*. Cambridge University Press, Cambridge.
- Corazza, E. (2004). 'Empathy, Imagination, and Reports'. Chapter 7 in *Reflecting the Mind: Indexicality and Quasi-Indexicality*. Oxford University Press, Oxford.
- Corkum, V. & Moore, C. (1995). 'Development of Joint Visual Attention in Infants'. Printed in Moore & Dunham ed.: 61-84.
- Croce, B. (1992). *The Aesthetic as the Science of Expression and of the Linguistic in General*. Trans. Lyas, C., Cambridge University Press, Cambridge.
- Currie, G. (2004). 'Anne Brontë' and the uses of imagination'. Chapter 9 in *Arts and Minds*. Oxford University Press, Oxford.
- Currie, G. (2007). 'Framing Narratives'. In Hutto, D. ed. *Narratives and Persons*. Cambridge University Press, Cambridge.

- Currie, G. & Ravenscroft, I. (2002). *Recreative Minds. Imagination in Philosophy and Psychology*. Clarendon Press, Oxford.
- Darwin, C. (1872). *The Expression of Emotions in Man and Animals*. Located at [http://pages.britishlibrary.net/charles.darwin3/expression/expression\\_intro.htm](http://pages.britishlibrary.net/charles.darwin3/expression/expression_intro.htm) accessed 3.7.05.
- Damasio, A. (2000). *The Feeling Of What Happens: Body, Emotion and the Making of Consciousness*. Vintage, London.
- Damasio, A. (2004). *Looking for Spinoza*. Vintage, London.
- Davidson, D. (2001a). 'The Emergence of Thought'. In *Subjective, Intersubjective, Objective*. Clarendon Press, Oxford: 123-134.
- Davidson, D. (2001b). 'Three Varieties of Knowledge'. In *Subjective, Intersubjective, Objective*. Clarendon Press, Oxford.: 205-220.
- Davidson, J. (1997). 'The Social in Musical Performance'. Chapter 11 in Hargreaves & North ed.: 209-228.
- Davidson, J. (2002). 'Communicating with the body in performance'. Chapter 10 in Rink ed.: 144-152.
- Davies, S. (1980). 'The Expression of Emotion in Music'. *Mind*, Vol. 89, No. 353 (Jan. 1980): 67-86.
- Davies, S. (1994). *Musical Meaning and Expression*. Cornell University Press, Ithaca and London.
- Davies S. (1997). 'Contra the Hypothetical Persona in Music'. Printed in Hjort & Laver ed.: 95-109.
- Davies, S. (1998). 'Artistic Expression'. Printed in Craig, E. ed. (1998) *Routledge Encyclopaedia of Philosophy*, Vol. 1. Routledge, London: 493-498.
- Deliege I. & Sloboda, J. ed. (1996). *Musical Beginnings: Origins and Development of Musical Competence*. Oxford University Press, Oxford.
- De Sousa, R. (1987). *The Rationality of Emotions*. MIT Press, Cambridge.
- Desrochers, S., Morissette, P. & Ricard, M. (1995). 'Two Perspectives on Pointing in Infancy'. Printed in Moore & Dunham ed.: 85-102.
- Deutsch, D. ed. (1999). *The Psychology of Music*. Second edition. Academic Press, a division of Harcourt Brace & company, London.
- Dewey, J. (1894). 'The Theory of Emotion. (I) Emotional Attitudes'. *Psychological Review* 1, (1894): 553-569. Located at [http://spartan.ac.brocku.ca/~lward/Dewey/Dewey\\_1894a.html](http://spartan.ac.brocku.ca/~lward/Dewey/Dewey_1894a.html)

accessed 7.11.06.

Dewey, J. (1971). 'The Act of Expression' from *Art as Experience* (1934). Printed in Hospers ed.: 72-94.

Dimberg, U., Thunberg, M. & Elmehed, K. (2000). 'Unconscious facial reactions to emotional facial expressions'. *Psychological Science*, Vol. 11: 86-89.

Dowling, W. J. (1999). 'The Development of Music Perception and Cognition'. Chapter 15 in Deutsch ed.: 618-619.

Dretske, F. (1981). *Knowledge and the flow of information*. MIT Press, Cambridge.

Dretske, F. (1986). 'Misrepresentation'. In Bogdan ed.: 17-36.

Dunham, P. & Dunham, F. (1995). 'Optimal Social Structures and Adaptive Infant Development'. Printed in Moore & Dunham ed.: 159-188.

Eilan, N. (2005). 'Joint Attention, Communication and Mind'. In Eilan, Hoerl, McCormack & Roessler ed.: 1-33.

Eilan, N., Hoerl, C., McCormack, T. & Roessler, J. ed. (2005). *Joint Attention: Communication and Other Minds*. Clarendon Press, Oxford.

Eisenberg, E. (1988). *The Recording Angel: Music, records and culture from Aristotle to Zappa*. Picador, London.

Ekman, P. (1977). 'Biological and cultural contributions to body and facial movement'. In Blacking, J. ed. (1977): 34-84.

Ekman, P. (1997). 'Should we call it expression or communication?' *Innovations in Social Science Research*, 10: 333-344.

Ekman, P. (2003). *Emotions revealed: Understanding faces and feelings*. Weidenfeld & Nicholson, London.

Fadiga, L., Fogassi, L., Pavesi, G. & Rizzolatti, G. (1995). 'Motor facilities during action observance: A magnetic stimulation study'. *Journal of Neurophysiology* 73: 2608-2611.

Fitzpatrick, D. (2003). 'Searle and Collective Intentionality. The Self-Defeating Nature of Internalism with Respect to Social Facts'. Printed in Koepsell & Moss ed.: 45-66.

Ford, C. C. (1995). 'Free Collective Improvisation in Higher Education'. First printed in *British Journal of Music Education* 12, (1995): 103-112. Located at [http://www20.brinkster.com/improarchive/cf\\_fci.htm](http://www20.brinkster.com/improarchive/cf_fci.htm) accessed 20.11.04.

Franco, F. (2005). 'Infant Pointing: Harlequin, Servant of Two Masters'. In Eilan, Hoerl, McCormack & Roessler ed.: 129-164.

- Fridlund, A. (1994). *Human facial expression: An evolutionary view*. Academic Press, San Deigo.
- Gabrielson, A. (2001). 'Emotions in strong experiences with music'. Chapter 19 in Juslin & Sloboda ed. (2001): 431-449.
- Gallese, V. & Goldman, A. (1998). 'Mirror neurons and the simulation theory of mind-reading'. *Trends in Cognitive Sciences*, Vol. 2: 493-501.
- Gilbert, M. (1989). *On Social Facts*. Routledge, New York.
- Gilbert, M. (2000). *Sociality and Responsibility: New essays in Plural Subject Theory*. Rowman & Littlefield publishers, Maryland.
- Goldie, P. (2000). *The Emotions, A Philosophical Exploration*. Clarendon Press, Oxford.
- Goldie, P. (2004). 'Emotion, feeling, and knowledge of the world'. In Solomon ed. (2004): 91-106.
- Goldman, A. I. & Sripada, C. S. (2005). 'Simulationist models of face-based emotion recognition'. *Cognition*, Vol. 94, No. 3, (Jan 2005): 193-213.
- Goodman, E. (2000). 'Analysing the ensemble in music rehearsal and performance: the nature and effects of interaction in cello-piano duos'. Ph.D. thesis, University of London.
- Goodman, E. (2002). 'Ensemble Performance'. Chapter 11 in Rink ed.: 153-167.
- Graham, K. (2002). *Practical Reasoning in a Social World: How We Act Together*. Cambridge University Press, Cambridge.
- Gregory, A. H. (1997). 'The roles of music in society: the ethnomusicological perspective'. Chapter 7 in Hargreaves & North ed.: 123-140.
- Greenspan, P. (1988). *Emotions and Reasons*. Routledge, New York.
- Grice, P. (1957). 'Meaning'. *The Philosophical Review*, Vol. 66, No. 3 (July 1957): 377-388.
- Griffiths, P. (1997). *What Emotions Really Are: The Problem of Psychological Categories*. University of Chicago Press, Chicago.
- Griffiths, P. (2003). 'Basic emotions, Complex emotions, Machiavellian Emotions'. In Hatzimoysis ed. (2003): 39-67.
- Hanslick, E. (1986). *On the Musically Beautiful*. Trans. Payzant, G. 8th ed. of 1891. Hackett, Indianapolis.

- Hargreaves, D. J. & North, A. C. ed. (1997). *The Social Psychology of Music*. Oxford University Press, Oxford.
- Harré, R. ed. (1986). *The Social Construction of Emotions*. Blackwell, Oxford.
- Hatfield, E., Cacioppo, J. T. & Rapson, R. L. (1994). *Emotional Contagion*. Cambridge University Press, Cambridge.
- Hatzimoysis, A. ed. (2003). *Philosophy and the Emotions*. Royal Institute of Philosophy Supplement: 52. Cambridge University Press, Cambridge.
- Hatzimoysis, A. (2003a). 'Emotional feelings and intentionalism'. In Hatzimoysis ed. (2003): 105-111.
- Heaton, P., Hermelin, B. & Pring, L. (1999). 'Can children with autistic spectrum disorders perceive affect in music? An experimental investigation'. *Psychological Medicine* 29: 1405-1410.
- Heble, A. (2000). *Landing on the Wrong Note; Jazz, Dissonance, and Critical Practice*. Routledge, London.
- Heelas, P. (1986). 'Emotion talk across cultures'. In Harré ed. (1986): 234-266.
- Hentoff, N. (1962). *The Jazz Life*. Peter Davies, London.
- Higgins, K. M. (1991). *The Music of our Lives*. Temple University Press, Philadelphia.
- Hjort, M. & Laver, S. ed. (1997). *Emotion and the Arts*. Oxford University Press, Oxford.
- Hobson, R. P. (1989). 'On sharing experiences'. *Development and Psychopathology*, 1 (1989): 197-203.
- Hobson, R. P. (2005). 'What Puts the Jointness into Joint Attention?' In Eilan, Hoerl, McCormack & Roessler ed.: 185-204.
- Hospers, J. ed. (1971). *Artistic Expression*. Meredith Corporation, New York.
- Hurley, S. (2005). 'The Shared Circuits Hypothesis: A Unified Functional Architecture for Control, Imitation, and Simulation'. Printed in Hurley & Chater ed. Vol. 1: 177-193.
- Hurley, S. & Chater, N. ed. (2005). *Perspectives on Imitation: From Neuroscience to Social Science*. Vols. 1 & 2. Bradford Book/MIT Press, London.
- James, W. (1884). 'What is an Emotion?' *Mind*, 9: 188-205. Located at <http://psychclassics.yorku.ca/James/emotion.htm> accessed 21.6.05.

Juslin, P. N. (1997). 'Emotional Communication in Music Performance: A Functionalist Perspective and Some Data'. *Music Perception*, Vol. 14, No. 4 (Summer 1997): 383-418.

Juslin, P. N. (2001). 'Communicating emotion in music performance: A review and a theoretical framework'. Chapter 14 in Juslin & Sloboda ed. (2001): 309-337.

Juslin, P. N. (2003). 'Five Facets of Musical Expression: A Psychologist's Perspective on Music Performance'. *Psychology of Music* 31(3): 273-302

Juslin, P. N. & Sloboda, J. ed. (2001). *Music and Emotion: Theory and research*. Oxford University Press, Oxford.

Kemp, A. (1997). 'Individual differences in musical behaviour'. Chapter 2 in Hargreaves & North ed.: 25-45.

Kenny, A. (1963). *Action, Emotion and Will*. Routledge, London.

Kieran, M. ed. (2005). *Contemporary Debates in Aesthetics*. Blackwell Publishers, London.

Kinsbourne, M. (2004). 'The Multimodal Mind: How the senses combine in the brain'. Located at <http://www.semioticon.com/virtuals/multimodality2/talks/kinsbourne.pdf> accessed 24.6.04.

Kivy, P. (1980). *The Corded Shell. Reflections on Musical Expression*. Princeton University Press, Guildford, Surrey.

Kivy, P. (1990). 'What was Hanslick Denying?' *Journal of Musicology*, 8: 3-18.

Koepsell, D. & Moss, L. S. (2003). *John Searle's Ideas about Social Reality: Extensions, Criticisms, and Reconstructions*. Blackwell Publishing, Oxford.

Lange, C. G. (1912) [1885]. 'The Mechanism of The Emotions'. Trans. Rand, B. Located at <http://psychclassics.yorku.ca/Lange/> accessed 21.6.05.

Langer, S. K. (1951) [1942]. *Philosophy in a New Key*. Mentor Books, Harvard University Press, New York.

Laird, J. & Bresler, C. (1992). 'The process of emotional experience: A self-perception theory'. Printer in M. S. Clark ed. (1992): 213-234.

Lawrence, A., Goerendt, I. & Brooks, D. (2006). 'Impaired recognition of facial expressions of anger in Parkinson's disease patients acutely withdrawn from dopamine replacement therapy'. *Neuropsychologia*, Vol. 44, June 2006 (Epub ahead of print).

Lazarus, R. S. (1984). 'On the primacy of cognition'. *American Psychologist*, 39: 124-129.

- Lazarus, R. S. (1991). *Emotion and Adaptation*. Oxford University Press, New York.
- Le Bon, G. (1896). *The Crowd: A study of the popular mind*. London: Ernest Benn, Ltd. Located at <http://etext.lib.virginia.edu> accessed 26.1.07.
- LeDoux, J. (1998). *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. Simon Schuster, New York.
- Levinson, J. (1990). 'Hope in The Hebrides'. Chapter 14 in *Music, Art & Metaphysics: Essays in Philosophical Metaphysics*. Cornell University Press, London.
- Levinson, J. (1997). 'Music and Negative Emotions'. In Robinson ed. (1997): 215-241.
- Levinson, J. (2002). 'Review of Aesthetics of Music by Roger Scruton'. *The Philosophical Review*, 109, No. 4: 608-614.
- Levinson, J. (2005). 'Musical Expressiveness as Hearability-As-Expression'. Chapter 12 in Kieran ed.: 192-206.
- Lewis, A. (1987). 'The social interpretation of modern jazz'. printed in White ed.: 33-55.
- Locke, J. (1996). *An Essay Concerning Human Understanding*. Winkler, K. ed. Hackett Publishing Inc. Indianapolis.
- Magrini, T. (1998). 'Improvisation and Group Interaction in Italian Lyrical Singing'. Chapter 8 in Nettl & Russell ed.: 167-198.
- Marsalis, B. (2004). 'It's a Jazz Thing'. Television programme aired on Channel 4, Saturday 13th November 2004.
- Martin, P. J. (2002). 'Spontaneity and Organisation'. Printed in Cooke & Horn ed.: 133-152.
- Matravers, D. (1998). *Art and Emotion*. Clarendon Press, Oxford.
- Matravers, D. (2003). 'The Experience of Emotion in Music'. *The Journal of Aesthetics and Art Criticism* Vol. 61, No. 4 (Fall 2003): 353-363.
- McDougall, W. (1920). *The Group Mind. A sketch of the principles of collective psychology with some attempt to the interpretation of national life and character*. Cambridge University Press, London.
- McEwan, I. (1998). *Enduring Love*. Vintage, Random House, London.
- Meijers, A. W. M. (2003). 'Can Collective Intentionality be Individualised?' Printed in Koepsell and Moss ed.: 167-183.

- Meltzoff, A. N. (2005). 'Imitation and Other Minds: The "Like Me" Hypothesis'. In Hurley & Chater ed. vol. 2: 55-77.
- Meltzoff, A. N. & Gopnik, A. (1996). 'The role of imitation in understanding persons and developing a theory of mind'. In Baron-Cohen, Tager-Flusberg & Cohen ed.: 335-366.
- Meltzoff, A. N. & Moore, M. K. (1983). 'Newborn infants imitate adult facial gestures'. *Child Development*, 54: 702-709.
- Meltzoff, A. N. & Prinz, W. ed. (2002). *The Imitative Mind: Development, Evolution, and Brain Bases*. Cambridge University Press, Cambridge.
- Meyer, L. B. (1956). *Emotion and Meaning in Music*. University of Chicago Press, London.
- Moore, C. & Dunham, P. ed. (1995). *Joint Attention. Its Origins and Role in Development*. Lawrence Erlbaum Associates Inc., New Jersey.
- Moore, C. & Dunham, P. (1995a). 'Current Themes in Research on Joint Attention'. In Moore & Dunham ed.: 15-28.
- Nettl, B. & Russell, M. ed. (1998). *In the course of performance: studies in the world of musical improvisation*. University of Chicago Press. London.
- Noordhof, P. (2005). 'Artistic Expression in Action'. Unpublished draft of a paper presented at the *Workshop on the Work of Kendall Walton*, Nottingham University, 1.7.05.
- Papousek M. (1996). 'Intuitive Parenting: a hidden source of musical stimulation in infancy'. Printed in Deliege & Sloboda ed.: 89-112.
- Peacocke, C. (2005). 'Joint Attention: Its Nature, Reflexivity, And Relation to Common Knowledge'. In Eilan, Hoerl, McCormack & Roessler ed.: 298-324.
- Peretz, I. (2001). 'Listen to the brain: a biological perspective on musical emotions'. Chapter 5 in Juslin & Sloboda ed. (2001): 105-134.
- Peretz, I., Gagnon, L. & Bouchard, B. (1998). 'Music and Emotion: Perceptual determinants, immediacy, and isolation after brain damage'. *Cognition*, 68: 111-141.
- Philip, R. (2004). *Performing Music in the Age of Recording*. Yale University Press, New Haven and London.
- Pickard, H. (2003). 'Emotions and the Problem of Other Minds'. In Hatzimoysis ed. (2003): 87-103.
- Pleasants, H. (1969). *Serious Music - and All that Jazz!* Victor Gollancz, London.

- Pratt, C. C. (1931). *The Meaning of Music*. McGraw-Hill, New York.
- Pressing, J. (1998). 'Psychological Constraints on Improvisational Expertise and Communication'. Chapter 2 in Nettle & Russell ed.: 47-67.
- Pressing, J. (2002). 'Free jazz and the avant-garde'. Printed in Cooke & Horn ed.: 202-216.
- Prinz, J. (2003). 'Emotion, Psychosemantics, and Embodied Appraisals'. In Hatzimoysis ed. (2003): 69-86.
- Prinz, J. (2004). *Gut Reactions: A Perceptual Theory of Emotion*. Oxford University Press, Oxford.
- Prinz, J. (2004a). 'Embodied Emotions'. Chapter 3 in Solomon ed. (2004): 44-58.
- Proust, M. (1996) [1913]. *Swann's Way*. Book 1 of *In Search of Lost Time*. Trans. Moncrieff & Kilmartin, revised by Enright, D. J. Vintage, London.
- Putnam, H. (1981). 'Brains in a vat'. Chapter 1 in *Reason, Truth & History*, Cambridge University Press, Cambridge: 1-21.
- Racy, A. J. (1998). 'Improvisation, ecstasy, and performance dynamics in Arabic music'. Chapter 4 in Nettle & Russell ed.: 95-112.
- Ramachandran, V. & Hubbard, E. (2003). 'Hearing Colors, Tasting Shapes'. *Scientific American*, May 2003.
- Richards, S. (1992). *Sonic Harvest: towards musical democracy*. Amber Lane Press, London.
- Ridley, A. (1995). *Music, Value and the Passions*. Cornell University Press, New York.
- Ridley, A. (1997). 'Not Ideal: Collingwood's Expression Theory'. *Journal of Aesthetics and Art Criticism* Vol. 55, No. 3 (Fall 1997): 263-272.
- Rink, J. ed. (2002). *Musical Performance. A Guide to Understanding*. Cambridge University Press, Cambridge.
- Robinson, J. ed. (1997). *Music and Meaning*. Cornell University Press, New York.
- Robinson, J. (2006). 'Gut Reactions: A Perceptual Theory of Emotion'. Book review in *British Journal of Aesthetics*. Vol. 46, No. 2, (April 2006): 206-208.
- Rochat, P. (2002). 'Ego function of early imitation'. In Meltzoff & Prinz ed.: 85-97.
- Roessler, J. (2005). 'Joint Attention and the Problem of Other Minds'. In Eilan, Hoerl, McCormack & Roessler ed.: 230-259.

- Roland, P. E., Larsen, B., Lassen, N. A. & Skinhoj, E. (1977). 'Regional cerebral blood flow increase due to treatment of somatosensory and auditive information in man'. *Acta Neurologica Scand. Supplement* 64: 540-541.
- Rosenthal, D. (1991). 'The independence of consciousness and sensory quality'. In Villanueva, E. ed. *Philosophical Issues*. Vol. 1, *Consciousness*: 15-36. Ridgeview Press, Atascadero, California.
- Salmela, M. (2006). 'True Emotions'. *The Philosophical Quarterly*, Vol. 56, No. 224, (July 2006): 382-405.
- Samet, J. (1996). 'Autism and theory of mind: some philosophical perspectives'. In Baron-Cohen, Tager-Flusberg & Cohen ed.: 427-449.
- Sawyer, R.. K. (2003). *Group Creativity: music, theatre, collaboration*. Lawrence Erlbaum Associates, London.
- Schachter, S. & Singer, J. (1962). 'Cognitive, social and physiological determinants of emotional state'. *Psychological Review*, 69: 379-399.
- Scheler, M. (1954). *The Nature of Sympathy*. Trans. Heath, P. Routledge and Kegan Paul, London.
- Scherer, K. R., Johnstone, T. & Klasmer, G. (2003). 'Vocal Expression of Emotion'. Chapter 23 in Davidson, R. J., Scherer, K. R., & Hill Goldsmith, H. ed. (2003). *Handbook of Affective Sciences*. Oxford University Press, London.
- Scherer, K. & Zentner, M. R. (2001). 'Emotional effects of music: production rules'. Chapter 16 in Juslin & Sloboda ed. (2001): 361-392.
- Schiffer, S. (1972). *Meaning*. Oxford, Clarendon Press.
- Schopenhauer, A. (1907). *The World as Will and Idea*. 6th ed., trans. Haldane, R. B. & Kemp. J. Kegan Paul, London.
- Schutz, A. (1971). 'Making music together: A study in social relationship'. In *Collected Papers II: Studies in Social Theory*. Brodersen, A. ed. Martinus Nijhoff, The Hague: 159-178.
- Scruton, R. (1997). *The Aesthetics of Music*. Clarendon Press, Oxford.
- Searle, J. R. (1995). *The Construction of Social Reality*. The Free Press, New York.
- Sibley, F. (1959). 'Aesthetic Concepts'. *Philosophical Review* Vol. 68, No. 4 (October 1959): 421-450
- Slawek, S. (1998). 'Keeping it going: Terms, practices, and processes of improvisation in Hindustānī instrumental Music.' Chapter 14 in Nettl & Russell ed.: 335-368.

- Sloboda, J. A. & Juslin, P. N. (2001a). 'Psychological perspectives on music and emotion'. Chapter 5 in Juslin & Sloboda ed. (2001): 71-104.
- Small, C. (1998). *Musicking, the meanings of Performing and Listening*. Wesleyan University Press, London.
- Smith, C. (1998). 'A sense of the possible: Miles Davis and the semiotics of improvised performance.' Chapter 12 in Nettl & Russell ed.: 261-289.
- Smith, H. & Dean, R. (1997). *Improvisation, hypermedia and the arts since 1945*. Harwood Academic Publishers, Amsterdam.
- Solomon, R. C. (1976). *The Passions*. Doubleday, New York.
- Solomon, R. C. (2003a). 'Against Valence ("Positive" and "Negative" Emotions)'. Chapter 10 in *Not Passion's Slave*. Oxford University Press, Oxford.: 162-177.
- Solomon, R. C. (2003b). 'Thoughts and Feelings: What is a 'Cognitive Theory' of the Emotions, and does it Neglect Affectivity?'. Chapter 11 in *Not Passion's Slave*. Oxford University Press, Oxford.: 178-194.
- Solomon, R. C. ed. (2004). *Thinking about Feeling. Contemporary Philosophers on Emotions*. Oxford University Press, Oxford.
- Sperber, D. & Wilson, D. (1993). *Relevance: Communication and Cognition*. Blackwell Publishers, Oxford.
- Stelmack, R. M. & Campbell, K. B. (1974). 'Extraversion and auditory sensitivity to high and low frequency'. *Perceptual and Motor Skills*, 38: 875-9.
- Storr, A. (1992). *Music and the Mind*. Harper Collins, London.
- Stoutland, F. (1997). 'Why Are Philosophers of Action so Anti-Social?' in Alanen, Heinamaa, & Wallgreen ed. *Commonality and Particularity in Ethics*. St. Martin's Press, New York.
- Stowell, R. ed. (2003). *The Cambridge Companion to the String Quartet*. Cambridge University Press, Cambridge.
- Strack, F., Martin, L. L., & Stepper, S. (1988). 'Inhibiting and facilitating conditions of facial expressions: A nonobstrusive test of the facial feedback hypothesis'. *Journal of Personality and Social Psychology*, 54: 768-777.
- Strawson. P. F. (1971). 'Intention and Convention in Speech Acts'. In *Logico-Linguistic Papers*. Methuen & Co. Ltd., London: 149-169.
- Tanila, H., Carlson, S., Linnankoski, I., Lindroos, F., Kahila, H. (1992). 'Functional properties of dorsolateral prefrontal cortical neurons in awake monkey'. *Behavioural Brain Research* 47: 169-180.

- Tollefsen, M. (2004). 'Collective Intentionality'. *Internet Encyclopedia of Philosophy*, located at <http://www.iep.utm.edu/c/coll-int.htm> accessed 26.5.06.
- Tolstoy, L. (1899). *What is Art?* Trans. Maude, A. Walter Scott Ltd., London.
- Tolstoy, L. (1933). *War and Peace*. Trans. Maude, A. Oxford University Press, London.
- Tomasello, M. (1995). 'Joint Attention as Social Cognition'. In Moore & Dunham ed.: 103-130.
- Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Harvard University Press, London, England.
- Tomes, S. (2003). 'Emotional Rescues'. *The Guardian*, Saturday November 22, 2003.
- Tuomela, R. (1991). 'We Will Do It: An Analysis of Group-Intentions'. *Philosophy and Phenomenological Research* Vol. 51, No. 2 (June 1991): 249-277.
- Trivedi, S. (2001). 'Expressiveness as a Property of the Music Itself'. *The Journal of Aesthetics and Art Criticism* Vol. 59. No. 4 (Fall 2001): 411-420.
- Unger, P. (1990). *Identity, consciousness, and value*. Oxford University Press, New York.
- Velleman, D. (1997). 'How to Share an Intention'. *Philosophy and Phenomenological Research* 57, No. 1 (March 1997): 29-50.
- Vermazen, B. (1986). 'Expression as Expression'. *Pacific Philosophical Quarterly* 67: 196-234.
- Violentyev, A., Shimojo, S., & Shams, L. (2005). 'Touch-induced visual illusion'. *Neuroreport*. Vol. 16, No. 10: 1107-1110.
- Walton, K. (1978). 'Fearing Fictions'. *Journal of Philosophy*, Vol. 75, No. 1: 5-27
- Walton, K. (1990). *Mimesis as Make Believe: On the Foundations of the Representational Arts*. Harvard University Press, Cambridge, Massachusetts.
- Walton, K. (1994). 'Listening with Imagination: Is Music Representational?' *The Journal of Aesthetics and Art Criticism* Vol. 52, No. 1 (Winter 1994): 47-62
- Waterman, D. (2003). 'Playing quartets: a view from the inside'. Chapter 5 in Stowell ed.: 97-126.
- Waterman, M. (1996). 'Emotional Responses to Music: Implicit and Explicit Effects in Listeners and Performers'. *Psychology of Music* Vol. 24, (1996): 53-67.

- Werner, H. & Kaplan, B. (1963). *Symbol Formation*. Erlbaum, Hillsdale, New Jersey.
- Wierzbicka, A. (1999). *Emotions across language and cultures: Diversity and universals*. Cambridge University Press, Cambridge.
- Wilson, R. A. (2004). *Boundaries of the Mind. The Individual in the Fragile Sciences*. Cambridge University Press, Cambridge.
- White, A. L. (1987). *Lost in music: culture, style, and the musical event*. Routledge & Kegan Paul, London.
- Wollheim, R. (1980). *Art and its Objects*. Cambridge University Press, Cambridge.
- Wordsworth, W. (1802). Preface to *Lyrical Ballads*. Located at <http://www.english.upenn.edu/~mgamer/Etexts/lbprose.html#preface> accessed on 16.9.05
- Young, V. & Colman, A. (1979). 'Some Psychological Processes in String Quartets'. *Psychology of Music* Vol. 7 (1979): 12-18.
- Zaibert, L. A. (2003). 'Collective Intentions and Collective Intentionality'. Printed in Koepsell & Moss ed.: 209-232.
- Zajonc, R. B. (1984). 'On the primacy of affect'. *American Psychologist*, 39: 117-123.
- Zajonc, R. B., Murphy, S. T. & Inglehart, M. (1989). 'Feeling and facial efference: Implications of the vascular theory of emotion'. *Psychological Review*, 96: 395-416.

